Identification of genetic loci underlying equine metabolic syndrome and laminitis risk

A DISSERATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Elaine M. Norton

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Molly E. McCue, Advisor

October 2019

Acknowledgements:

I would first like to acknowledge my amazing and brilliant advisor, Dr. Molly McCue. Molly, you are an incredible researcher and leader, and your innovation and constant drive are something I can only strive to achieve. You taught me how to constantly evaluate, and reevaluate a research question, and to quote you, "If it was easy, they wouldn't call it re-search." You are an inspiration and taught me more than I could ever express in words.

Second, I would like to acknowledge the chair of my committee, Dr. James Mickelson. Mick, you are an amazing mentor and someone that I strive to emulate. The major things that stand out for me is your constant patience and dedication to the heart of the science. The hundreds of hours you sacrificed for editing the numerous documents I sent you will be forever acknowledged in the pages of this thesis.

I would also like to acknowledge the two other members of my committee, Drs Ray Geor and Kent Reed. Thank you so much for your guidance, support, and motivation throughout the past several years.

I am incredibly grateful to have been funded by a National Institute of Health T32 Institutional training grant and Morris Animal Foundation Doctoral fellowship. I am incredibly grateful to the directors of the T32 training grant, Drs Cathy Carlson and Dave Brown for believing in me, and both the National Institute of Health and Morris Animal Foundation for providing me with the salary support throughout this Ph.D.

In addition, I would like to acknowledge the members of the equine genetics and genomics lab, both those currently in the lab and those that have already progressed in their careers. These include Dr. Annette McCoy, Dr. Nichol Schultz, Dr. Raffa Teixeira, Sam Beeson, Dr. Robert Schaefer, Dr. Sian Durward-Akhurst, Dr. Felipe Avila, Jonah Cullen, Kevin Murray, Dr. Lauren Hughes, and Dr. Sarah Jacobs. You all have helped me on sampling trips, learn how to write code, and understand a myriad of aspects of genetics. You all are incredibly talented, brilliant, and amazing researchers, and I am forever grateful to call you all family.

I also need to acknowledge my family who I love and am incredibly grateful to have in my life. Of utmost importance is my husband, Jacob Erdei. Jacob, you have made this dream possible with your support and willingness to follow me all over the United States from the beginning of this journey to the end. Also, to my parents who have always been there for me and encouraged me to follow my dreams. To both my sisters by birth, Stephanie and Becky who believed in me and have always been there when I needed them. To my brothers and sisters by marriage, I am honored to be a part of your family and am grateful for your support and encouragement. Finally, thank you to my two longest friends and sisters by choice, Nila Reese and Marliena Ellis. You both have been my shoulder and venting platform and made me laugh when no one else could.

Finally, I want to acknowledge all of the owners, horses and referral veterinarians who participated in this project. This research would not have been possible without your dedication to learning more about EMS and laminitis.

Dedication:

This dissertation is dedicated to my son Joshua and daughter Reagan. I love both of you so much and, as your mom, I pray that you will always find your happiness, follow your heart, and accomplish your dreams no matter how impossible they may seem at times. I am forever here to support and love you. I have been so blessed to continue to watch you grow, and you both will always be my greatest accomplishments.

Abstract:

Laminitis is a painful, debilitating disease of the hoof, often resulting in these horses being humanely euthanized due to uncontrolled pain. The most commonly cited cause of this life-threatening disease is a clustering of clinical signs resulting from a metabolically efficient phenotype, termed equine metabolic syndrome (EMS). While EMS is a commonly diagnosed syndrome, knowledge of the underlining pathophysiology is lacking and recommendations for diagnostic criteria are vague and inconsistent. EMS is thought to be complex disease, and identification of its underlying genetic risk factors and key gene-by-environment interactions will improve our understanding of EMS pathophysiology and allow for early detection of high-risk individuals and intervention prior to the onset of laminitis. *We hypothesized that major genetic risk factors leading to EMS and laminitis susceptibility are shared across breeds of horses, and that differences in the severity and secondary features of the EMS phenotype between breeds, or between individuals within a breed, are the result of modifying genetic risk alleles with variable frequencies between breeds.*

To test these hypotheses, my PhD thesis has consisted of using phenotype and genotype data on 286 Morgan horses and 264 Welsh ponies, two high risk breeds for EMS. Phenotype data collected on all horses included: signalment, medical history, laminitis status, environmental management (feed, supplements, turnout and exercise regimen), and morphometric measurements (body condition score (BCS), wither height, and neck and girth circumference). After an eight hour fast, an oral sugar test (OST) was performed using 0.15mg/kg Karo lite corn syrup. Biochemical measurements included baseline insulin, glucose, non-esterified fatty acids (NEFA), triglycerides (TG), adiponectin, leptin and ACTH; and measurements 75 minutes after the OST included insulin (INS-OST) and glucose (GLU-OST). For inclusion in the study, each farm had to have at least one control and one horse with clinical signs consistent with EMS under the same management. Single nucleotide polymorphism (SNP) genotyping was performed on all horses. Haplotype phasing and genotype imputation up to two million SNPs was performed on horses genotyped on lower density arrays using Beagle software. Quality control on the imputed data was performed using the Plink software package. After

genotype pruning, 1,428,337 and 1,158,831 SNPs remained for subsequent analysis in the Welsh ponies and Morgan horses, respectively.

In chapter 2, SNP genotype data from the Welsh ponies and Morgan horses were used to estimate the heritability of the nine EMS biochemical measurements. Heritability (h^2_{SNP}) was estimated using a restricted maximum likelihood statistic with the inclusion of genetic relationship matrix, which was corrected for linkage disequilibrium (regions of the genome which are not independent as they are inherited together). The confounders of age, sex and season were included in the model based on the Akaike information criteria. In the Welsh ponies, seven of the nine biochemical traits had h^2_{SNP} estimates with p-values that exceeded the Holm-Bonferroni corrected cut-off as follows: triglycerides (0.31), glucose (0.41), NEFA (0.43), INS-OST (0.44), adiponectin (0.49), leptin (0.55), and insulin (0.81). Six of the nine EMS traits in the Morgans had h^2_{SNP} estimates with p-values that exceeded the Holm-Bonferroni cutoff as follows: INS-OST (0.36), leptin (0.49), GLU-OST (0.57), insulin (0.59), NEFA (0.68), and adiponectin (0.91). Insufficient population size and high trait variability may have limited power to obtain statistically significant h^2_{SNP} estimates for ACTH (both breeds), glucose and triglycerides in Morgans and GLU-OST in Welsh ponies. These data provide the first concrete evidence of a genetic contribution to key phenotypes associated with EMS and demonstrate that continued research for identification of the genetic risk factors for EMS phenotypes within and across breeds is warranted.

Although heritability estimates provide valuable insight on the genetic contribution to a trait, they do not provide information on the number of contributing genes, specific genes involved, or where these genes are located within the genome. Genome wide association analyses (GWA) use SNP genotype data to identify those key regions of the genome that are associated with a trait. The objectives of chapter 3 were to (i) perform within breed GWA to identify significant contributing loci in Welsh ponies and Morgans, and (ii) use a meta-analysis approach to identify shared and unique loci between both breeds. For each trait, within breed GWA were performed from the imputed SNP genotype data using custom code for an improved mixed linear regression analysis. Prior to analysis, traits were adjusted to account for known covariates, with sex and age included as fixed effects

and farm as a random effect. GWA meta-analysis was performed with a random effects model using the Morgans and Welsh pony GWA summary data from the 688,471 SNPs that were shared between breeds.

To define the boundaries of the region, a pairwise comparison of linkage disequilibrium (LD) was calculated for all SNPs within the region. A custom code was used to identify regions where LD for all SNPs dropped below the LD threshold of 0.3 and spanned at least 100kb both 5' and 3' to the widest peak of LD within the window, which was used to define the boundaries of the ROI. An LD-region was identified as shared if it was within the boundaries of another LD-region and prioritized as described above for the fixed regions. Regions were prioritized based on whether they were identified as shared between breeds on meta-analysis (high priority), shared across traits (medium priority), or found in a single breed but not shared across traits (low priority). Prioritization resulted in 56 high, 26 medium, and 7 low priority genomic regions for a total of 1853 candidate genes in the Welsh ponies, and 39 high, 8 medium and 9 low priority regions for a total of 1167 candidate genes in the Morgan horses. Meta-analysis identified 65 of these regions that were shared across breeds. These data demonstrate that EMS is a polygenic trait with both across breed and breed specific genetic variants.

In chapter 4, we utilized imputed whole-genome sequencing (WGS) and linear regression analysis in order to fine-map selected high priority LD-ROI in both the Morgan horses and Welsh ponies. LD-ROI were fine-mapped if they contained at least 5 SNPs with one SNP exceeding the threshold for genome-wide significance. Five fine-mapped regions from each breed were further interrogated for predicted impact using variant annotation. Protein-coding genes containing non-coding or coding variants within the fine-mapping region were then further prioritized based on known function and biological evidence in other species utilizing the PubMed search engine. A total of 19 positional candidate genes were identified as having biological evidence for a role in EMS. These data provide support for the process of fine-mapping GWA ROI by increasing marker density and using biological evidence across species to further prioritize candidate genes.

In chapter 5, a missense mutation in the first exon of HMGA2 was identified as a putative functional allele for height and EMS phenotypes in Welsh ponies. It is well recognized

that ponies (short horses) are at high risk for developing EMS; and in humans shorter individuals have an increased risk of developing cardiovascular disease, type II diabetes and metabolic syndrome. We hypothesized that genetic loci affecting height in ponies have pleiotropic effects on metabolic pathways and increase the susceptibility to EMS. Pearson's correlation coefficient identified an inverse relationship between height and baseline insulin (-.26) in the Welsh ponies. Genomic signature of selection analysis was performed using a di statistic and identified a ~1.3 megabase region on chromosome 6, that was also identified on GWA. Haplotype analysis using HapQTL confirmed that there was a shared ancestral haplotype between height and insulin. This region contributed ~40% of the heritability for height and ~20% of the heritability for insulin. *HMGA2* was identified as a candidate gene, and sequencing identified a single a c.83G>A variant (p.G28E) in *HMGA2*, previously described in other small stature horse breeds. In our cohort of ponies, the A allele had a frequency of .76, was strongly correlated with height (-.75) and was low to moderately correlated with metabolic traits including: insulin (.32), insulin after an oral sugar test (.25), non-esterified fatty acids (.19) and triglyceride (.22) concentrations. For this allele, model analysis suggested an additive mode of inheritance with height and a recessive mode of inheritance with the metabolic traits. This was the first gene identified as having a pleotropic effect for EMS.

In conclusion, the results of my thesis are major steps forward in understanding the genetic contributions of EMS in two high risk breeds. Future directions include the continued identification of the specific genes and alleles contributing to EMS and could include prioritization of the positional candidate genes identified in aim 2 via (1) identification of biological candidate genes based on known gene function and evidence from other species; (2) use of whole genome sequencing and linear regression analysis to fine map regions; (3) use of intermediate phenotypes such as metabolomics or transcriptomics to identify shared regions; or (4) network analysis for identification of genes within similar, relevant pathways.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Literature Review

Laminitis is a debilitating and often career ending disease of the digital laminae in horses. Management of laminitis is one of the greatest challenges in equine practice and many horses are euthanized due to uncontrolled pain [1,2]. The most commonly cited cause of this life-threatening disease is termed equine metabolic syndrome (EMS), a clustering of metabolic disturbances including insulin resistance and derangements in fat metabolism [3,4]. Over the past few decades, our understanding of EMS has become more refined; however, the etiology, clinical consequences, and underlying pathophysiology of EMS are still largely unknown and remain under investigation.

History of defining the EMS phenotype:

In the 1980s associations between insulin insensitivity and laminitis in ponies were reported [5,6]. However, the term EMS was not coined until 2002 when the parallel between human metabolic syndrome (MetS) and what was being observed clinically in hyperinsulinemic, obese, laminitic horses was recognized [7]. In humans, MetS has been described as a clustering of risk factors leading to an increased risk of cardiovascular disease and type II diabetes mellitus [8]. Although several health organizations have published their own criteria for the diagnosis of MetS [9-12], in 2009 the International Diabetes Federation Task Force published a joint statement defining MetS as an individual with three or more of the following five risk factors: hypertension, dysglycemia, hypertriglyceridemia, decreased high-density lipoprotein cholesterol, and central obesity [8].

Shortly following the identification of EMS as a clinical syndrome in horses, several observational and experimental studies were published evaluating the EMS phenotype. Associations between EMS and insulin dysregulation [13-17], hyperglycemia [15], hypertriglyceridemia [13,14,17], elevated non-esterified fatty acids [NEFA] [15], hypertension [14,16], alterations in adipokines [15,16] or inflammatory mediators [18], and obesity and regional adiposity [13,15,16] were identified. However, several discrepancies defining the key components of EMS were present across the literature, particularly as to whether obesity, regional adiposity, hyperglycemia, elevated NEFA and triglycerides were consistent with the EMS phenotype [13-16]. These discrepancies likely reflected low

sample numbers, differences in diagnostic criteria or assays, as well as unknown variables such as between-farm variability, excessive nutrition, seasonal variation, different cohort characteristics (different breeds) and exercise regimens [19].

To address these issues, a large across-breed study of metabolic phenotypic variation in 610 horses and ponies, including 5 target breeds (Morgan horse n=293, Welsh Pony n=100, Quarter Horse n=59, Arabians n=64, Tennessee Walking Horses n=48), and 46 horses from 15 other breeds was performed [20]. Data collected included: signalment, history, laminitis status, morphometric measurements, body condition score (BCS), exercise regimen, diet (total intake and feed analysis of hay, concentrates, and supplements), and biochemical measurements including fasting glucose, insulin, ACTH, leptin, adiponectin, triglycerides (TG), and NEFA, and insulin and glucose levels 75 minutes after an oral sugar challenge (OST- 0.15mg/kg Karo syrup). To minimize environmental confounders, at least one EMS suspect and one age-, breed- and gender-matched control were sampled from each farm. A multi-level, multivariate, multiple regression model was constructed to assess the relationships between the eleven phenotypic responses and 16 individual and environmental explanatory variables.

Key results from this study included: (i) hyperinsulinemia, an exaggerated insulin response to OST, low serum adiponectin, and hypertriglyceridemia were the EMS phenotypes most strongly associated with laminitis; (ii) genetics (heritability) explained up to 55% of the variation in EMS phenotypes; and (iii) certain features of EMS phenotype varied between breeds, including fasting insulin, insulin after an OST, and adipokine concentrations. Further, TG and NEFA concentrations were higher in ponies than horses [20]. This work helped to redefine the EMS phenotype and provided the veterinary community with a better understanding of the complexity of this syndrome.

Current knowledge of EMS phenotype:

In 2019, the American College of Veterinary Internal Medicine published a revised consensus statement which defined EMS as a set of risk factors leading to endocrinopathic laminitis and identified the key component of EMS as insulin dysregulation [21]. The authors defined insulin dysregulation as derangements in the balanced relationship between plasma insulin, glucose and lipids, and could manifest clinically as baseline

hyperinsulinemia, an exaggerated or prolonged insulin or glucose response post oral or intravenous carbohydrate challenge, tissue insulin resistance, or hypertriglyceridemia [21]. Although obesity and regional adiposity are associated with EMS, several cases of insulin dysregulation have been identified in non-obese horses and this is no longer considered a key diagnostic criterion [21]. Further, the distinction between obesity as a risk factor versus clinical consequence of EMS are still being parsed out and will be discussed further in the *Risk Factors* section. Therefore, obesity will be discussed here in relation to prevalence and its impact on insulin dysregulation.

Epidemiology: Despite the relatively high number of horses diagnosed with EMS, there is little epidemiological data about the prevalence of this syndrome. Several studies have evaluated the prevalence of components of EMS [22] and reflect the variability between EMS phenotypic traits, breeds and geographical regions. Baseline hyperinsulinemia affected 23% and 27% of a population of 300 clinically normal horses in Ohio [23] and 208 Australian ponies [24], respectively. In these studies, hyperinsulinemia was associated with age [23,24], BCS [23], and other metabolic abnormalities including hyperleptinemia [24] and hypertriglyceridemia [24]. One study evaluating laminitis cases in a single referral hospital in Finland, determined a prevalence of 89% were defined as endocrinopathic laminitis, or laminitis induced by derangements in the metabolic system including both EMS and pars pituitary intermedia dysfunction (PPID). Of these cases, two-thirds were diagnosed with EMS and pony breeds were overrepresented [4].

As in humans, obesity in horses has been characterized as a growing epidemic [25] with the prevalence in over conditioned (Henneke BCS of 7) and obese (Henneke BCS≥8) horses ranging from 23-51% [26-31]. The highest prevalence of obesity was identified in the United States [30] and ponies have been found to have a threefold higher prevalence of obesity then horse breeds [28]. In addition, Rocky Mountain horses, Tenness Walking horses, Quarter Horses, and mixed breed horses had a higher incidence of obesity when compared to Thoroughbreds [30]. Season was also identified to affect obesity prevalence, with a higher prevalence at the end of the summer versus the end of the winter in horses housed outdoors [26]. Several studies also identified an association between obesity and the concurrent presence of regional adiposity (cresty neck) although they were not mutually

exclusive [28,30]. Carter et al proposed utilizing a cresty neck score (CNS) as measurement of neck crest adiposity, with scores ranging from 0 (no visual or palpable crest) to 5 (a crest so large it droops to one side) [32]. Evaluation of the prevalence of regional adiposity identified that 8.9% of Pura Raza Español horses had a score of 5 [33] and 33% of ponies from the United Kingdom had a score of ≥3/5 [34]. Unfortunately, the prevalence of horses with obesity and/or regional adiposity that also have insulin dysregulation is unknown.

Notably, these studies represent a limited picture of EMS components and additional studies are required to determine a more precise prevalence of this syndrome. Specifically, studies are needed to determine the prevalence of all the components of EMS, and the rate at which these components co-occur within individual horses, the prevalence of obese and nonobese horses with insulin dysregulation, and the prevalence within breeds considered high risk for EMS. A better understanding of the frequency of this syndrome is essential for determining its full impact on equine health.

Risk Factors: The identification of disease risk factors is imperative in order to accurately assess an individual's risk for developing a disease, as well as to identify environmental modifications that could substantially improve the overall outcome and reduce the severity of disease. Both environmental (diet, exercise, and season) and individual risk factors (breed, age, sex, and obesity) have been identified for EMS.

One of the most recognized risk factors for EMS is breed as specific breeds have a higher prevalence of EMS including, Standardbred horses, Morgan horses, Tennessee Walking horses, Andalusians, Paso Finos, and Arabians, with pony breeds being considered at the highest risk [7,35]. Although most breed predilections have been determined anecdotally through clinical observation, several scientific reports have supported difference in metabolic profiles between breeds. As early as the 1980s, published reports concluded that ponies were less insulin sensitive then large breed horses [5,36] and that there were distinct breed differences in lipid and glucose profiles [37]. Ponies and Andalusians have also been found to be less insulin sensitive when compared with Standardbreds; further, Andalusians had a lower disposition index indicating that the breed was less compensated for insulin resistance then ponies [38].

Metabolic syndrome is thought to be the result of a "metabolically thrifty" genotype, resulting in an advantageous adaptation to survive during periods of scarce feed and harsh climate conditions in undomesticated horses [39]. However, after the agricultural revolution, horses were presented with consistent feed, abundant soluble carbohydrates, and a more sedentary lifestyle. It has been hypothesized that this change in environment and diet has shifted advantageous thrifty genotypes to now result in susceptibility to the derangements in metabolism associated with EMS and obesity [7,36]. This theory was based on the thrifty genotype hypothesis in humans stating that obesity and metabolic syndrome are the result of genetic variants which allowed for human ancestors to survive during periods of poor nutrition by increasing adipose stores during the reciprocal period of food abundance [40]. See the *Genetics of Metabolic Syndrome* for additional theories in humans related to the thrifty genotype hypothesis.

Therefore, it is not surprising that initially obesity was thought to be the primary cause of EMS [3] and was identified as a risk factor in several studies [13,16]. However, this has been an inconsistent finding [14] and studies evaluating the effect of obesity and insulin dysregulation have had conflicting results [41-44]. In 13 Arabian geldings, dietary-induced weight gain over a period of four months led to a compensated insulin resistance with a 71% decrease in insulin sensitivity and corresponding ~400% increase in the acute insulin response to glucose challenge which was maintained up to 5 weeks after conclusion of the trial [42]. Notably, there were no significant differences in glucose, NEFA, or TG levels in this study [42]. On the other hand, increased weight gain from caloric intake at 200% maintenance for one to two years revealed elevated basal insulin levels in horses and ponies, as well as increased basal glucose and NEFA levels in ponies [41]. However, all values were within the reference range and the results from a combined glucose-insulin test (CGIT) were consistent with insulin sensitivity [41]. The findings in the latter study were similar to two previous reports which also did not identify a change in insulin sensitivity after dietary-induced weight gain [43,44]. Differences in the percentage of non-structural carbohydrates (NSC) fed to promote weight gain in these studies might explain the discrepancies in findings. While Carter et al's study utilized a diet with an NSC content of 34.7% of dry matter (DM) intake [42], while the remaining three studies had NSC contents at <20% of DM intake [41,43,44].

Diets high in NSC have been shown to decrease insulin sensitivity when compared to diets high in fat or fiber [44-46]. However, improved insulin sensitivity was identified in a group of adult and geriatric horses after adaptation to either a high starch (24.5% NSC DM intake) or high sugar (22.7% NSC DM intake) diet [47], as well as in a group of horses fed a diet containing 1.5g/kg glucose (~30% NSC on DM intake) once daily [48]. Consequently, the exact mechanism or threshold of NSC to induce tissue insulin insensitivity is unknown and likely reflects unrecognized interactions between insulin, glucose, lipid metabolism, and the gastrointestinal microbiome, which is further supported by the recognition of metabolically healthy obese phenotypes and metabolically unhealthy thin phenotypes in both humans [49] and horses [50,51].

Therefore, obesity as a risk factor or clinical consequence of EMS is still under debate and may represent distinction subtypes of EMS. Nonetheless, when present, obesity has been shown to exacerbate insulin dysregulation [52] and is considered a risk factor for the development of endocrinopathic laminitis [4,14,16,53]. Given the link between obesity and EMS, it is not surprising nutrition and exercise have also been identified as risk factors and will be discussed in further details under the *Management* section.

As mentioned above (see *Epidemiology*), seasonal variation has been identified for the prevalence of obesity in horses maintained outdoors, with a higher prevalence in the summer versus the winter [26]. This finding may reflect the ancestral adaptation to harsher climates and decreased food availability in winter, leading to the metabolic survival mechanism of increased fat storage during summer and hypometabolism in the winter [54-56]. This is further supported by the finding that even after maintaining a constant, controlled energy balance in a group of Quarter Horses, leptin, thyroid stimulating hormone, and total T4 levels were all found to be greater in the summer compared to the winter [57]. Interestingly, pony breeds have been shown to have a higher incidence of obesity in the winter versus the spring [58]. This may reflect that, even after domestication, ponies have maintained a more stringent metabolically thrifty phenotype with excessively suppressed metabolic rates and fat storage [56].

Given seasonally adapted changes in metabolic rates, it would be expected that most metabolic hormones would show circannual variation; however, a consensus on the effect

of season on several EMS components has not been established. ACTH, an adrenocorticotropic hormone, was found to be correlated with several EMS traits [20] but the biological significance in EMS has not been established. ACTH is commonly used to rule out pars pituitary intermedia dysfunction (PPID) in cases of insulin dysregulation and should be interpreted with caution during periods of stress or pain such as an acute laminitis episode. Notably, ACTH has been repeatedly shown to have circannual variations with increased levels in the late summer to early fall in the Northern and Southern hemisphere [59-63]. Interesting, geographical region has been found to affect the degree of this variation, with regions near the equator (smaller magnitude of change in day length) having a decrease in length and magnitude of the dynamic phase [64].

Findings assessing seasonal variation for EMS traits has been more variable. Several studies identified an increase in basal insulin levels and insulin resistance in the fall or summer months [14,61,62,65]; however, seasonal variation for glucose and insulin dynamics is not a consistent finding [60,66]. In one study, seasonal variation was found to be inconsistent, with the effect of season being significant on insulin in the first year but was not significantly associated during subsequent sampling of the same population the following year [66]. The authors also identified a large within horse and month-to-month variation in insulin dynamics [66]. Interestingly, when assessing the effect of season on the combined-insulin and glucose tolerance test, seasonal variation was identified for the glucose area under the curve and nadir but did not affect the overall interpretation of the test [67]. Further, in large breed horses, TG levels were found to be elevated in winter months compared to the summer; however, in the same study, TG levels were not affected by season in pony breeds although NEFA were found to elevated in the winter in this cohort [37]. This was in contrast to later findings in which TG were found to be significantly higher in the summer versus the winter in pony breeds [14]. Finally, when comparing inflammatory mediators in a group of previously laminitic and non-laminitic ponies, seasonal variation was observed for several inflammatory markers including fibrinogen, serum amyloid A, haptoglobin and interleukin-4 (IL-4) but was not observed for the anti-inflammatory marker adiponectin nor for plasma TG levels [58]. Conversely, in a group of Finn horses, adiponectin gene expression was found to be upregulated in subcutaneous adipose tissue at the end compared to the start of the grazing season [52]. Thus, seasonal

variation as a risk factor for EMS traits is variable and likely influenced by differences in geographical location and breed variations.

Age has also been identified as an individual risk factor with older horses having decreased insulin sensitivity [24,65,68,69] and lower adiponectin concentrations [70] compared to younger horses. Further, after weight gain, TNFα levels were increased in a group of older horses but not in a group of young horses challenged with the same weight gain [71]. The effect of sex on EMS is less understood as several studies found that sex was not a risk factor for EMS traits [58,65,72], although anecdotally stallions are considered more insulin sensitive. Further, stallions were identified as being 8 times less likely to develop pasture associated-laminitis compared to females in a group of 160 ponies [13]; although, sex was not considered a risk factor for recurrence of endocrinopathic laminitis in a later study [73]. Interestingly, sex does appear to influence inflammatory mediators although the effect of cytokines on EMS is still being investigated (see *Inflammatory Cytokines)*. In a group of ponies, geldings had a significantly higher concentration of plasma fibrinogen and serum amyloid A [58]. TNFα was significantly higher in females in a group of 110 light breed horses, whereas both age and being female were found to be associated with higher levels of IL-6. Consequently, neither inflammatory mediator was correlated with BCS or basal insulin concentrations although the authors did identify a correlation with serum amyloid A [18].

Although there are discrepancies between these studies, is clear that both environmental and individual risk factors affect the expression and severity of the EMS phenotype. However, these known risk factors have been shown to only explain 12.9-58.6% of the environmental and 9.6-36.3% of the individual variations between EMS phenotypes [20]. This led to the identification of an association between endocrine disrupting chemicals and EMS phenotypes, a previously unknown environmental risk factor [74]. Thus, the continued investigation of risk factors is necessary to fully understand the mechanisms underlying EMS as well as improving management options.

Clinical Consequences: Although metabolic syndrome may appear to be a relatively benign health concern, in both humans and horses, metabolic syndrome can lead to serious medical issues that have a major economic impact. In humans, individuals with metabolic

syndrome are 2 times more likely to develop cardiovascular disease and 4 times more likely to develop type II diabetes [8]. Horses with EMS have the highest risk for developing laminitis and have an increased risk of vascular dysfunction, reproductive issues, and a decreased immune response.

The development of laminitis is the primary clinical concern of horses with EMS, due to the painful and often career ending outcome of this disease. Although laminitis itself is not fatal, in the best interest of the patient, the severity and crippling pain often lead to the decision for euthanasia [75]. Although there is a plethora of inciting factors which result in laminitis, EMS is considered the leading cause [76]. Initially, laminitis associated with EMS was termed "pasture-associated laminitis" after a survey from the USDA identified that 46% of laminitis cases occurred from horses and ponies housed on lush pasture [77], which was later linked with insulin dysregulation [13,16]. The term endocrinopathic laminitis was proposed to encompass causes of laminitis due to dysregulation of the endocrine system, including EMS and PPID, and accounted for the development of laminitis in metabolic horses not housed on pasture [78].

Inflammation is a primary component of sepsis-associated laminitis and horses with EMS have been found to have higher levels of systemic pro-inflammatory cytokines; however, the role of inflammation in endocrinopathic laminitis has been questioned and the term endocrinopathic laminopathy has been proposed to reflect this distinct difference [79]. Histological evaluation of naturally occurring cases and experimentally induced models of endocrinopathic laminitis noted minimal neutrophil infiltration into the lamellar tissue despite the comparably large number of necrotic and apoptotic cells [76,80-82]. Immunostaining for calprotectin, a leukocyte marker, was positive in an euglycemic hyperinsulinemia clamp (EHC) model of laminitis, but to a lesser extent then what was reported in other models of laminitis [80]; further, there was no difference in calprotectin staining in ponies fed a high versus low carbohydrate diet [83]. Evaluation of the protein expression of inflammatory mediators in lamellar tissue post EHC, noted increased expression of toll-like receptor 4 (TLR4) and tumor necrosis factor α (TNFα) in clinical cases compared to controls and subclinical cases [84], while another study identified an upregulation of lamellar IL-1, IL-6, IL-11, COX-2, and e-selectin mRNA and

downregulation of COX-1 in EHC induced laminitis [82]. These reports suggest that the inflammation seen in endocrinopathic laminitis is a secondary response, potentially playing a role in the progression of disease, but not a primary inciting factor [80,83,84].

Further, histologically, endocrinopathic laminitis appears to have a different pathophysiology when compared to models of inflammatory laminitis; specifically, there is lack of global basement membrane separation and neutrophil infiltration [80,85]. In cases of experimentally induced exogenous or endogenous hyperinsulinemia, prior to the onset of clinical lameness, early disease progression revealed cell death, narrowing, and elongation within the secondary epidermal lamina (SEL) [81,86]. Progression of disease revealed further elongation and proliferative activity within the SEL as well as inflammatory cell infiltration. These changes likely reflect cytoskeleton disruption and cellular disorganization, leading to instability of the SEL [81,86]. Similar lesions were identified in naturally occurring cases of endocrinopathic laminitis, with marked apoptosis and elongation, tapering and fusion of the SEL with hyperkeratinized tissue [76]. One distinct difference between experimental and naturally occurring models was the lack of miotic cells, which the authors surmised reflected chronicity and cellular differentiation to hyperkeratosis [76].

Further, although the exact mechanism behind the development of endocrinopathic laminitis in not understood, experimental and field studies suggest that hyperinsulinemia is a primary inciting factor [14,16,73,85,87,88]. Experimental studies using the EHS in healthy horses or ponies were able to induce laminitis within 48-72 hours in all treatment groups [80,85,87]; however, these experiments required prolonged, supraphysiological levels of insulin to maintain euglycemia which may not mimic natural cases of endocrinopathic laminitis. Using intravenous glucose to induce hyperglycemia and endogenous hyperinsulinemia, investigators determined that horses in the treatment group, although not clinically lame, developed histopathological lesions consistent with laminitis and that insulin alone, or in combination with glucose, were inciting factors for endocrinopathic laminitis at an insulin toxic threshold of ~200µIU/mL [86]. Further, hyperinsulinemia and insulin resistance have been used to predict the development of laminitis [14,16], plasma insulin levels were positively correlated with laminitis severity

[89,90], and horses and ponies with basal levels of insulin >20μIU/mL had a higher risk of a recurrent laminitis episode within two year [73].

Several theories have been proposed as to the mechanism behind hyperinsulinemia induced laminitis including the activation of insulin-like growth factor 1 (IGF-1) and/or insulin receptor isoforms specific to lamellar tissues. Unlike the insulin receptor, IGF-1 receptor is expressed both in lamellar epithelial and endothelial cells [91], can be activated by insulin during periods of hyperinsulinemia via the mitogen activated protein kinase (MAPK) pathway [92], and IGF-1 induced pathology in cancer cells is similar to what is seen histologically in cases of endocrinopathic laminitis, including: increased mitotic rate, disruption of the basement membrane and cytoskeletal dysregulation [93,94]. Further, IGF-1 receptor was found to be downregulated in the lamellar tissue, without a concurrent upregulation in the circulation, of horses with experimentally induced hyperinsulinemia [95,96]. Further, two insulin receptor isoforms and hybrid have been identified in the lamellar tissue of horses [96]. Therefore, the role of these receptors are intriguing but further experiments are required to identify the effect and downstream signaling of hyperinsulinemia on both IGF-1 and the insulin receptor isoforms to understand if they have a role in laminitis pathology.

Additional theories include the role of vascular dysfunction; however, whether vascular dysfunction is an inciting factor or clinical consequence of endocrinopathic laminitis is still under investigation. Normal blood flow to the horse's hoof is critical to maintain healthy lamellar tissue and is regulated, in part, by the insulin-dependent signaling pathways phosphatidylinositol 3-kinase (PI3K; responsible for regulation of vascular glucose metabolism and vasodilation via the stimulation of nitric oxide from the vascular endothelium) and MAPK pathway (responsible for growth, mitogenesis and vasoconstriction via the stimulation of endothelial-1 from the vascular endothelium) [97,98]. Hyperinsulinemia has been shown to inhibit the PI3K pathway while overstimulating the MAPK pathway [99,100], leading to vasoconstriction or impaired vasodilation [91]. This was supported by *in vitro* experiments of laminar arteries and veins which showed vasoconstriction and increased endothelin-1 production [101] and MAPK mediated vasoconstriction/impaired vasodilation [102,103] after preincubation with

insulin. Further, laminar vessels from naturally occurring cases of laminitis had reduced vasodilation after stimulation with acetylcholine and increased vasoconstriction when exposed to phenylephrine compared to controls [104].

Vascular dysfunction is not limited to the lamellar vessels in horses with EMS. *In vitro* experiments revealed increased vasoconstriction and impaired vasodilation in the facial vasculature of horses with endocrinopathic laminitis [104]. Further, EMS horses and ponies were found to have higher resting heart rates then control horses [105,106], and EMS ponies had evidence of myocardial hypertrophy [105]. Although hypertension was identified in a group of prelaminitic ponies [14], this has not been a consistent finding [105,106]. Interestingly, both insulin insensitive and EMS horses subjected to an EHC had a limited response to insulin-induced changes in systolic, diastolic or mean blood pressure versus controls, with insulin insensitive and EMS horses showing a reduction in blood pressure to a lesser extent then controls [106,107].

Additional clinical consequences include a decreased cell mediated immune response to vaccination [108] and subfertility in mares and stallions with EMS [21]. Specifically, insulin dysregulation has been linked to altered estrous cycles [109], anovulatory follicles [110], and changes in the intrafollicular environment [111]. However, the distinctive, or overlapping, roles of EMS versus obesity in infertility have not been parsed out.

Thus, the clinical consequences of EMS have a major impact on the equine industry, with the primary cost occurring as a result of the development of laminitis. Extensive research into endocrinopathic laminitis has suggested that its pathophysiology and etiology are distinctly different from other forms of laminitis. This warrants additional study as a better understanding of underlying mechanism may lead to the development of more tailored treatment strategies for management of these cases. Further, although not as extensively studied, the additional clinical consequences discussed also impact the equine industry due to costs associated with subfertility, increased illness, and decreased performance secondary to vascular dysfunction.

Management: The primary management consideration in horses with EMS is a regimen focused on dietary modification and exercise. Initial studies assessing EMS identified obesity and/or regional adiposity as major criteria in classifying horses with EMS

[13,15,16]. Not surprisingly, diet and a sedentary lifestyle were identified as risk factors, and the mainstay of management focused on promoting weight loss by decreasing caloric intake and increasing energy expenditure. Although obesity is now considered a feature and not cause of EMS [21], diet and exercise modifications are still a major step in managing the obese EMS horse by improving insulin sensitivity and preventing obesity in the nonobese EMS horse.

Reduction of caloric intake has been shown to promote weight loss and improve insulin regulation [112]. Initial recommendations for weight loss include eliminating concentrated feeds and treats and limiting total dietary intake to a low NSC hay and ration balancer at 1.25-1.5% of body mass on DM basis, targeting a weight loss rate of 0.5-1.0% of body mass per week [113]. However, weight loss resistant individuals have been described, and may require a carefully monitored reduction in feed to as little as 1.0% of body mass on DM basis [113].

Regardless of obesity status, feeding a hay with <10% NSC content is recommended for EMS horses to minimize insulin response [114,115]. In order to achieve the desired percent of NSC, soaking hay is commonly recommended as this process was shown to reduce the total water soluble carbohydrate (WSC) content by 27-50% [116-118] depending on the type and cut of hay as well as the time the hay was soaked [116,119]. Further, horses fed soaked hay had a two-fold greater increase in body weight reduction per week compared to previous reports of horses fed the same quality of dry hay, which the authors surmised was due to the decreased provision of DM leading to a 23% decrease in dietary energy after soaking [113]. Based on the degree of nutrient leeching noted in this study, the authors recommended an adjusted pre and post soaking equation to ensure the horse is receiving adequate nutrient provisions [118].

Diurnal variation, season and environmental stresses can lead to pastures rich in NSC which has been associated with insulin resistance and peak occurrences of laminitis cases [13,120], and transition from pasture to an all-hay diet resulted in improved insulin sensitivity in a group of insulin resistant ponies [121]. Further, ponies with moderate or severe insulin dysregulation were found to have higher post prandial insulin responses, often surpassing the previously proposed toxic threshold of 200 µIu/mL, when compared

to ponies with normal insulin regulation [122]. Unfortunately, attempting to limit the access to pasture is not a feasible option. Although, horses housed on pasture can consume between 1.5-5.2% of their body weight in 24 hours [123,124], horses with restricted pasture access were shown to consume nearly 1% of their body weight during a three-hour grazing period [125]. Therefore, as part of initial dietary management, it is recommended that EMS horses should be removed from pasture.

Exercise has also been shown to improve insulin sensitivity in both horse and ponies, but these results are inconsistent and likely reflect differences in research methodology, animal heterogeneity, and failure to achieve high enough exercise intensity. A seven-day light intensity exercise program improved insulin sensitivity up to nine days after conclusion of the exercise program in both obese and lean mares [126]. Similarly, in Standardbred horses, a seven-day intense exercise protocol improved insulin sensitivity for up to five days after the last exercise session [127]; while the results were maintained for less than 24 hours in a separate study [128]. Further, after six weeks of moderate-exercise, insulin sensitivity was improved in a group of hyperinsulinemic ponies to the same degree as ponies which achieved weight loss based exclusively on diet control [129]. This contrasts with several studies that did not find an effect on exercise and improved insulin sensitivity. The use of a dynamic feeding system for three months in a group of obese ponies resulted in a 3.7 fold increase in daily movement and body fat loss of approximately 5% but did not improve insulin sensitivity [130]. Further, low intensity exercise for four weeks followed by moderate intensity exercise for an additional four weeks resulted in weight loss without a concurrent change in insulin sensitivity, leptin, or triglyceride concentrations in a group of obese Arabian horses [131].

Bamford et al evaluated the effect of both diet restriction and low-intensity exercise compared to a monotherapy of diet restriction in 24 obese horses and ponies [132]. After 12 weeks, both groups had similar improvements in adiposity, insulin, leptin and adiponectin concentrations; however, the combined group had improved insulin sensitivity as well as decreased serum amyloid A concentrations compared to the monotherapy group [132]. These results support the use of combined protocols in EMS management and are consistent with previous findings identifying an anti-inflammatory benefit to exercise

[133]. Further, individually tailored programs which incorporated both diet and exercise achieved targeted weight loss and improved insulin sensitivity in client-owned EMS horses [134]. Therefore, combined diet and exercise regimens should be considered as a holistic approach to EMS management in horses healthy enough to exercise. Although additional studies are required to assess the effect of exercise on insulin sensitivity as well as to determine the optimal exercise intensity and duration required to achieve desired results, current recommendations are low-to-moderate exercise (heart rate of 130-170 beats per minute for a minimum of 30 minutes at least 6 days per week) in EMS, nonlaminitic horses, and light intensity exercise (heart rate of 110-150 beats per minute for a minimum of 30 minutes at least 4 times per week) in previously laminitic EMS horses [21].

The addition of pharmaceuticals has also been investigated for the use in EMS management. Metformin is a biguanide in which the mechanisms of action in humans include (i) inhibiting hepatic gluconeogenesis, (ii) improving tissue insulin sensitivity, (iii) delaying the uptake of glucose within the small intestine while increasing enterocyte glucose utilization, and (iv) increasing uptake of glucose by adipose tissue and skeletal muscle [135]. In horses, pharmacokinetic studies calculated an oral bioavailability of ~7% in unfed horses and ~4% in fed horses, which was 10 times lower than the bioavailability of metformin in humans [136]. Further, 20-day oral administration at 15mg/kg twice daily resulted in a steady state concentration lower then therapeutic concentrations in humans [137]. Several studies evaluating the clinical efficacy of metformin in insulin resistant ponies and horses identified either no difference in insulin sensitivity [138] or a short term improvement which was reduced or mitigated after extended administration of the drug [110,139]. However, it has been proposed that despite the low bioavailability, metformin concentrates within the intestine, and its local action on enterocytes may provide clinical benefit in horses. Oral administration of 30mg/kg metformin immediately prior to an oral glucose challenge, reduced peak glucose and insulin response in horses with dexamethasone induced insulin resistance [140]. The authors surmised that even if insulin sensitivity is not improved, blunting postprandial insulin responses would decrease peak insulin levels and reduce the risk of endocrinopathic laminitis [140].

Although most EMS horses are euthyroid, levothyroxine is still commonly used in EMS management regimens based on evidence that supplementation will promote weight loss via triiodothyronine (T3) and thyroxine (T4) stimulation of carbohydrate and fat metabolism. Levothyroxine is a synthetic analog of T4, and although pharmacokinetic studies are lacking in horses, oral administration has been shown to increase serum levels of T3 and T4 in horses. Oral administration of levothyroxine in increasing doses (24, 48, 72, and 96mg) at 2-week increments, resulted in elevated levels of total T4 (tT4) that exceed the reference range with the exception of the lowest dose [141]. Long-term oral administration of 48mg daily for 48 weeks resulted in 1.5-2-fold increase in total T3 (tT3) compared to mean baseline values [142]. The largest increase in tT4 levels was observed at 16 weeks (5.4-fold increase) compared to 32 (4.0-fold increase) and 48 weeks (3.7-fold increase) [142]. Single dose of 240mg or 480mg resulted in mild elevations in tT3 and tT4 but results remained within the reference interval; however, daily two-week administration of 480mg resulted in a marked increase in tT4 values which gradually declined after cessation of the trial [143]. These studies reflect that levothyroxine is dose dependent and further research is needed to assess dose saturation, tolerance and paradoxical effects of this drug in horses.

Horses administered an increasing dose of levothyroxine over 8 weeks had significant decrease in body weight despite free choice access to feed [141]. Further, this group of horses had decreased plasma concentrations of TG, total cholesterol, and very low-density lipoproteins as well as a two-fold improvement in insulin sensitivity [144]. A moderate dose (48mg) of levothyroxine for 12 months resulted in improved insulin sensitivity and weight loss [145]. However, the highest percent of weight loss was achieved at week 16 (autumn with a 10% reduction) versus the conclusion of the trial (summer with a 5% reduction). No control group was included in this study so the effect of weight loss on the drug versus other factors such as season and nutrient content of the pasture/forage could not be determined [145]. Further, levothyroxine is perceived as safe in horses, as long-term administration of 48mg levothyroxine daily did not produce behavioral, cardiac, or systemic adverse effects [142]. Health complications were also not identified in horses administered high doses of levothyroxine [141,143], although one group did note increased level of anxiety amongst horses treated at a dose of 96mg daily [141]. Thus, the inclusion

of levothyroxine to promote weight loss is an appealing option for both the weight loss resistant individual or laminitic horses unable to exercise.

Thus, the current evidence supports diet and exercise modifications in order to promote both weight loss and insulin sensitivity as the mainstay of EMS management. Management strategies including the addition of pharmaceuticals should only be considered in conjunction with diet and exercise and not as a replacement. Additional management options have been proposed but currently are not yet commercially available or have little scientific evidence for efficacy in horses with EMS, including: magnesium supplementation [146], nutraceutical supplementation [147], incretin receptor antagonists (discussed further in the **Pathophysiology of EMS**: *Incretins*) [148], sodium-glucose linked transport-2 inhibitors [149,150], and mesenchymal stem cell therapy [151].

Pathophysiology of EMS

Much of the etiology and pathophysiology of EMS is still incomplete, and current knowledge has been extrapolated from evidence in humans and other model species. Given that horses are hindgut fermenters adapted to a high roughage diet, there are likely distinct differences between metabolic physiology in horses and other species that could have major effects on clinical outcomes and treatment strategies. This section will outlay the current knowledge of metabolic syndrome pathophysiology in humans as well as relevant literature in the horse.

Insulin Dysregulation: Insulin is a peptide hormone with roles in carbohydrate, lipid, and protein metabolism. Insulin dysregulation plays an important role in metabolic syndromes and has been defined as disruptions in the interconnected relationships between insulin, glucose and lipid metabolism [21]. In normal conditions, insulin is released by β-cells from pancreatic islets of Langerhans primarily in response to hyperglycemia, but other macronutrients, hormones, and neurotransmitters can also stimulate insulin release [152]. Pancreatic β-cells act as "glucose sensors," playing a critical role in glucose homeostatasis [153]. Glucose readily crosses the β-cell membrane via high capacity, low affinity glucose transporters (GLUT) [154]. Although GLUT2 is the primary β-cell transporter in the mouse, GLUT1 and GLUT3 appear to be more important in humans and the horse β-cell transporter has yet to be identified [155]. Once within the cell, glucose is phosphorylated

by glucokinase (the primary β-cell glucose sensor) to glucose-6-phosphate, initiating glycolysis resulting in pyruvate, ATP, and NADH generation [153]. Subsequent closure of potassium-ATP-dependent channels, depolarizing the cellular membrane, and activation of voltage dependent calcium channels leads to in an influx of intracellular calcium which results in insulin secretion via exocytosis [152,153]. In humans and rodent models, insulin secretion is biphasic with an initial rapid release followed by a more prolonged but less concentrated release [156]; however, it has not been determined as to whether insulin release is mono or biphasic in horses [157-159].

In the periphery, insulin mediates its effects by binding to insulin receptors located on the main target tissues: adipocytes, skeletal muscle, and liver. This binding activates intracellular insulin responsive substrates (IRS) via tyrosine phosphorylation. These IRS are responsible for insulin's metabolic, vascular and mitogenic effects. For example, IRS activation of the downstream effector pathway phosphatidylinositol 3-kinase (PI3K) promotes insulin's metabolic effects including: the translocation of intracellular glucose transporter proteins (GLUT4) to the cell membrane, stimulating lipid and protein synthesis, and inhibiting glycogenolysis and hepatic gluconeogenesis [160,161]. Once the glucose transporter is translocated to the plasma membrane, glucose is transported into the cell where it is phosphorylated to be stored as glycogen or utilized for ATP production.

Hyperinsulinemia and insulin insensitivity are mainstays of insulin dysregulation. Theories for the pathogenesis of this disorder focus on prolonged, sustained diets rich in carbohydrates and/or fats leading to a peripheral insulin resistance, which may be a consequence of insulin receptor downregulation and/or desensitization, reduction in insulin receptor protein levels, inhibition of GLUT4 translocation to the plasma membrane [162], or alterations in central neuronal control of regulatory pathways [163]. Experimental support for these theories is limited in horses. However, GLUT4 was recently shown to be decreased on the skeletal muscle cell surface in horses with naturally occurring insulin resistance without a change in total protein expression [164]. Further, experimentally induced hyperinsulinemia in horses led to a decrease in GLUT1, GLUT4, insulin receptors, fatty acid transporters and CD36 transcript abundance in adipose tissue [165] (see *Dyslipidemia* for further discussion of fatty acid transporters and CD36).

With insulin resistance, as blood glucose levels fail to return to normoglycemia, the pancreatic β-cells hyperfunction by increasing insulin secretion via β-cell hypertrophy and decreasing β-cell expression of glucokinase while increasing expression of hexokinase, shifting the insulin-glucose response curve to the right [166,167]. In addition, in normal conditions, the first portal passage through the liver is responsible for 50-60% of insulin clearance in humans and up to 70% in horses [168,169]; however, insulin insensitivity markedly reduced hepatic insulin clearance across species [169,170], contributing to hyperinsulinemia. Theories for the progression from compensated to decompensated insulin resistance include: (i) deterioration of the pancreatic β-cell glucose-sensory mechanisms [171], (ii) reduction in the conversion from proinsulin to insulin [172], and (iii) β-cell exhaustion and apoptosis [173,174].

In horses, hyperinsulinemia and/or an exaggerated response to an oral sugar or IV glucose challenge remain the most commonly used tests for diagnosing individuals with EMS [21]. However, although some studies have found glucose levels in EMS horses to be significantly higher than insulin sensitive horses, these levels remain within the reference range indicating compensated insulin resistance [46]. Progression to uncompensated insulin resistance is rarely reported in horses with few published cases of diabetes mellitus [175-177] and an occasional identification of hyperglycemia in individuals apart of larger EMS studies [13]. This unique mechanism of maintained compensated insulin resistance remains undiscovered and warrants further investigation.

Incretins: Postprandial stimulation of insulin release via the enteroinsular axis has been found to result in a greater insulin response versus intravenous or intraperitoneal glucose administration. This has been deemed the incretin response based on the effects of the incretin hormones glucose-dependent insulinotropic polypeptide (GIP) and glucagon-like peptide 1 (GLP-1). While incretin hormones are responsible for 50-70% of postprandial insulin release in humans [178,179], their role is likely smaller in horses [180]. Both GLP1 and GIP's have primary insulinotrophic roles by binding to G-protein receptors (GLP-1 receptors and GIP receptors, respectively) on pancreatic beta cells, upregulating proinsulin gene expression and stimulating insulin release in a glucose-dependent manner [181]. Disruption of the incretin effect has been implicated in obesity and insulin dysregulation.

This was supported by evidence in the mouse models which showed the disruption of Islet1, a transcription factor expressed in incretin producing intestinal cells, resulted in the loss GLP-1 and GIP and impaired glucose tolerance to an oral but not intraperitoneal challenge [182]. However, the roles of incretins in insulin dysregulation are debatable.

In humans, postprandial GLP-1 levels have been found to be decreased in insulin dysregulated individuals [183-185] and were negatively correlated to body mass index [184,186,187]. In a group of 30 horses, both active (aGLP-1) and total (tGLP-1) levels were decreased in horses with insulin dysregulation compared to controls [188]. However, these results are inconsistent as several studies were unable to identify correlations between GLP-1 and severity of glucose intolerance [189-191]. Further, a 2011 meta-analysis surmised that reduced GLP-1 concentrations were not a universal characteristic of insulin dysregulation and individual factors such as diet, age, NEFA concentrations, and obesity were confounding these results [192].

In horses, breed differences in GLP-1 concentrations were identified in ponies and Andalusians, two breeds considered high risk for EMS, which were found to have higher GLP-1 and insulin responses compared to Standardbred horses, potentially indicating a risk factor for insulin dysregulation [193]. Further, both active (aGLP-1) and total (tGLP-1) levels were decreased in horses with insulin dysregulation compared to controls [188]. However, other studies found that although GLP-1 levels correlated with insulin secretion, GLP-1 could not be used to differentiate between insulin dysregulated and insulin sensitive horses [122,194].

Studies evaluating the role of GIP in human metabolic syndrome have had similar conflicting findings, with some studies identifying decreased levels or reduced response of GIP in individuals with insulin dysregulation and others finding no difference [195-199]. However, the role of GIP in insulin dysregulation is further complicated by its counterintuitive glucagonotropic role. Unlike GLP-1 which inhibits glucagon release [200], GIP has been shown to enhance glucagon release [201] which exacerbates post prandial hyperglycemia [202] and may contribute to the development of obesity [203]. These hypotheses have been supported through studies showing that, when compared to thin individuals, obese individuals have higher fasting levels of GIP and increased early

phase responses to a meal [204]. Similarly, GIP was positively correlated with body condition and cresty neck scores in a population of nine mixed breed ponies [148].

An "incretin exhaustion," or impaired incretin insulinotropic potency, has been proposed to explain the phenomenon of normal GLP-1 or GIP levels in insulin dysregulated individuals [190,205]. This has been supported by a decreased expression of GLP-1R and GIPR in diabetic mouse models [206,207]. Further, hyperglycemia has been shown to promote endocytosis of both the GIPR and GLP-1R [208]. However, additional research is required to support this theory. Despite the uncertainty behind the roles of GLP-1 and GIP, synthetic GLP-1 and drugs targeted at inhibiting dipeptidyl peptidase 4 (DPP4), a serine protease that inactivates GIP and GLP-1, have been shown to increase glycemic control in humans with metabolic syndrome [209,210], and represent a potential target for EMS management [148].

Dyslipidemia: Fatty acids circulate within the blood bound to albumin as free fatty acids (NEFA) or esterified as components of triglyceride, phospholipids, and/or esterified cholesterol [211]. Circulating lipoproteins are characterized by central hydrophobic cores composed of triglycerides and cholesterol esters and hydrophilic membranes consisting of phospholipids, free cholesterol and apolipoproteins, with apolipoproteins serving major roles in ensuring the structural stability of the lipoprotein, acting as a ligand for the lipoprotein receptor, and activating or inhibiting enzymes involved in lipoprotein metabolism [212]. Depending on hydrated density and major lipid content, lipoproteins are divided into five classes: high-density lipoproteins (HDL), intermediate-density lipoproteins (IDL), low-density lipoproteins (LDL), very-low-density lipoprotein (VLDL), and chylomicrons (packaged dietary triglycerides). In horses, classes of lipoproteins have been characterized across breeds with distinct subclasses of lipoproteins and the most abundant fraction mainly being HDL [213-215]. Breed differences have also been identified, with Shetland ponies having higher triglyceride and VLDL concentrations compared to Thoroughbreds [214]. Further, Turkman horses were found to have higher concentrations of triglycerides compared to most breeds and higher LDL compared to Morgan horses [216].

Fatty acids are derived from dietary uptake or endogenous de novo lipogenesis or lipolysis. Insulin is the most potent stimulus for promoting lipogenesis and inhibiting lipolysis, and exerts its effect by (i) increasing the cellular uptake of glucose, (ii) activating lipoprotein lipase (LPL) in adipose tissue, and (iii) inactivating hormone sensitive lipase through phosphodiesterase 3B [217,218]. Glucose also plays a role in lipogenesis by acting as is a substrate for fatty acid synthesis via its glycolytic conversion to acetyl CoA, inducing expression of lipogenic genes, and stimulating the release of insulin from pancreatic beta cells as described above [217]. In concert, insulin induced stimulation of protein phosphatase activates acetyl CoA carboxylase converting the glycolytic acetyl CoA to malonyl CoA, which is then converted to palmitate by fatty acid synthase [219]. Dietary fatty acids and palmitate can be modified to produce multiple lipid species via endogenous elongase and desaturase enzymes controlled by the regulatory elements peroxisome proliferator-activated receptor alpha (PPARalpha), sterol-regulatory element binding protein-1 (SREBP-1), carbohydrate-regulatory element binding protein (ChREBP), and MAX-like factor X (MLX) [220]. Further, the family of transcription factors SREBP-1 and ChREBP also have fundamental roles in regulating the expression of fatty-acid synthase, acetyl CoA carboxylase and ATP-citrate lysase [221].

Lipid partitioning is important to insulin action, energy balance, and regulation of body weight [222], and is predominately mediated by lipoprotein lipase (LPL), which is synthesized by the paranchemal cells of muscle, heart and adipose tissue and bound to capillary endothelial by highly charged chains of heparan sulfate proteoglycans and/or glycosyl phosphatidylinositol, forming the glycocalyx [223,224]. Insulin is a major regulator of LPL activity in adipose tissue and acts by increasing LDL gene transcription and regulating its activity through both posttranscriptional and posttranslational mechanisms [225]. Insulin further promotes fatty acid uptake by promoting translocation of fatty acid transporters to the cell [226] and targeting apolipoprotein B for degradation reducing hepatic output of VLDL resulting in the preferential uptake of chylomicrons [227]. Glucose increases adipose tissue LPL activity via glycosylation of LPL, essential for LPL catalytic activity and secretion [225]. LPL activity is further regulated in a tissue specific fashion by several factors including: angiopoietin-like proteins, lipase maturation factor, and glycosylphosphatidylinositol HDL binding protein [228,229]. LPL, in

conjunction with its required cofactor apolipoprotein C-II, acts by anchoring the surface of chylomicrons or lipoproteins to the endothelium, resulting in hydrolysis of triglycerides and release of free fatty acids which can then available to bind to fatty acid transport proteins including the fatty acid translocase protein CD36 for cellular uptake [211,223]. LPL further facilitates the uptake of LDL via interactions with the low-density lipoprotein receptors, which is independent of its catalytic activity and in a dose dependent manner [230,231]. LDL receptors mediate the uptake of LDL, IDL and chylomicron remnants via endocytosis, with LDL receptor expression being controlled by SREBP based on the total cholesterol content within the cell. Once the fatty acids are taken up by the tissues, their fate is determined by the tissue and activity of hormone sensitive lipase, fatty acyl CoA synthase and glycerol-3-phosphate acyltransferase, ultimately, being oxidized or stored as neutral lipids.

Under physiological conditions, circulating fatty acids act as a monitor between insulin and hepatic glucose production, with elevations in serum free fatty acid concentration leading to impaired hepatic insulin function, decreased hepatic insulin clearance, and increased gluconeogenesis [232,233]. Direct inhibition of insulin action by elevated fatty acids eliminates the negative feedback on hormone-sensitive lipase, leading to further accumulation of triglycerides in the liver and other target tissues and reducing the cellular uptake of glucose. Further, via the Randle cycle, elevated fatty acids levels lead to a shift in skeletal muscle preference to oxidize free fatty acids as energy substrates over glucose, which is normally responsible for 80% of postprandial glucose utilization, resulting in a prolonged hyperglycemia [234,235]. Finally, chronic exposure of the pancreatic beta cells to high levels of free fatty acids results in an impaired insulin secretory response to glucose [232,233] via NEFA inactivation of the pancreatic transcription factor hepatic nuclear factor-1α which suppresses GLUT2 mRNA expression [154,236].

Given the extensive control mechanisms of insulin and glucose on lipogenesis and lipolysis, as well as that of fatty acids on insulin sensitivity and glucose homeostasis, it is not surprising that dyslipidemia (elevations in serum NEFA, TG, and phospholipid concentrations) and insulin resistance are interrelated. Dysfunctional adipose tissues, either due to adipose hypertrophy from overnutrition or lipodystrophy, have been shown

to lead to insulin resistance due to excessive amounts of fatty acids reaching the circulation postprandially and during lipolysis and dysregulation of adipokines (aka lipid-induced insulin resistance) [237]. The "portal theory" has proposed that lipid-induced insulin resistance is the result of increased visceral adipose tissue delivering a higher rate of free fatty acids to the liver via the portal vein, which increases hepatic glucose production and reduces hepatic insulin clearance, and ultimately leading to hyperinsulinemia [238]. Insulin resistance-induced obesity has also been proposed based on evidence in lean at-risk individuals (individual with a family history of type 2 diabetes) who had decreased function of the insulin receptor kinase activity and reduced plasma membrane insulin receptor concentrations on skeletal and/or adipose tissue [239-241]. Further, using a proxy for adipose-insulin resistance (calculated as free fatty acids multiplied by baseline insulin), adipose insulin resistance was higher in a small group of individuals with metabolic syndrome after adjusting for BMI and waist circumference [242]. The role of obesity as a risk factor versus clinical consequence of metabolic syndrome has already been discussed; regardless, disturbances to fatty acid metabolism are of importance to the pathophysiology of metabolic syndrome.

Notably, specific alterations in fatty acid uptake have also been linked to metabolic dysfunction and may be an inciting or contributing cause to elevated circulating levels of fatty acids in cases of insulin dysregulation. In humans, variants within the apolipoprotein C-II gene and/or LPL have been shown to lead to hypertriglyceridemia and increased plasma levels of chylomicrons [243,244]. Mice deficient in skeletal LPL were shown to have hypertriglyceridemia which ultimately leads to insulin resistance and obesity due to increased lipid partitioning in other tissues [245]. Interestingly, tissue-specific overexpression of LPL in mice skeletal muscle and liver lead to increased stores of triglycerides within these tissues but ultimately still resulted in insulin resistance [246]. LPL activity has been repeatedly shown to be reduced in humans with poorly regulated type II diabetes, and two variants within the LPL gene were associated with hypertriglyceridemia and decrease HDL level in 2328 Danish individuals [247]. CD36 knockout mice have impaired fatty acid uptake [248] and variants within the rodent CD36 were linked to hyperlipidemia and insulin resistance [249]. Humans with insulin resistance have been shown to have an increased protein abundance of skeletal muscle CD36 [250],

and insulin-sensitive horses with experimentally induced hyperinsulinemia were identified to have reduced CD36 adipose transcription abundance [165].

In horses, elevated NEFA and TG have been associated with EMS, although not consistently across studies. Treiber, et al and Carter, et al both found significantly increased TG levels in EMS ponies and hypertriglyceridemia was identified as a risk factor for the development of laminitis [13,16]. Bailey et al identified a seasonal elevation in TG levels in ponies with EMS [14], and Frank et al identified a significant elevation in serum NEFA concentrations, although there was no difference in serum TG [15]. Interestingly, while the former studies were evaluating EMS in pony breeds, the latter study consisted of six different large-breed horses possibly reflecting a difference in ponies versus horses. In one study comparing a small group of obese large-breed horses with insulin resistance to metabolically healthy horses, the plasma concentration of VLDL and HDL were found to be increased in horses with insulin resistance [15]. The elevation in HDL is opposite to what is characteristically seen in humans with metabolic syndrome and may reflect the absence of plasma cholesteryl ester transfer protein in equids [15]. Using continuous lipid profile measurements, subfractions of HDL, specifically HDL3a, were found to be significantly lower in healthy horses compared to horses with obesity, laminitis or both despite there being no difference in total HDL concentrations, possibly indicating a novel method of lipoprotein profiling [215]. Unfortunately, this study did not evaluate metabolic status and further studies would need to be performed to determine the utility for EMS [215].

Inflammatory Cytokines: Evidence supports that a large contribution to the pathophysiology of metabolic syndrome is related to the role of adipose tissue in promoting chronic, low-grade inflammation. Adipose tissue is a biologically active endocrine organ which secretes a myriad of substances including cytokines, eicosanoids, complement proteins, binding proteins, vasoactive factors, and regulators of lipid metabolism which are collectively known as adipokines [251] and will be discussed in the subsequent section. The exact mechanism behind adipose induced inflammation in metabolic syndrome is unknown, but it is proposed that "sick fat" is a result of adipose hypertrophy and hyperplasia secondary to excess nutrition leading to endoplasmic reticulum stress or

hypoxia from an insufficient blood supply [252,253], which results in macrophage and mast cell infiltration and the production of inflammatory mediators [253]. Additional hypotheses have proposed that endothelial stress and hyperactive platelets are the primary etiology behind adipose tissue inflammation as well as the low-level inflammation observed in other tissues [254]. Regardless, the increased production of inflammatory mediators has been shown to indirectly and directly contribute or exacerbate insulin resistance and dyslipidemia. The following section will describe the evidence and pathophysiology for TNFα, IL-1β, and IL-6 as these are the three cytokines that have been most extensively studied for obesity and metabolic syndrome in the horse. However, it should be noted that additional adipocytokines which have evidence as to a contributing role in insulin dysregulation include: C-reactive protein (CRP), fibrinogen, resistin, monocyte chemoattractant protein-1 (MCP-1), IL-8, WISP1, apelin, angiopoietin 2, omentin-1, chermin, dipeptidylpeptidase 4, and plasminogen activator inhibitor-1 (PAI-1) [253].

TNFα was found to be upregulated in mouse models of obesity as early as the 1990s [255] and is one of the most thoroughly studied adipocytokines. It has been shown to promote insulin resistance by (i) hindering insulin-stimulated translocation of GLUT4 transporters by inducing serine phosphorylation of insulin receptor substrate 1 (IRS-1), resulting in docking of PI3K, (ii) terminating insulin action by activating SH-PTPase, which removes the tyrosine phosphate groups from IRS-1 and focal adhesion kinase (FAK), and (iii) inhibiting insulin receptors via serine/threonine phosphorylation [256]. In addition, binding of TNFα receptors results in the activation of NF-κB and c-Jun amino-terminal kinase (JNK). JNK further contributes to insulin resistance by phosphorylating and inactivating IRS-1 [257]. NF-κB, which is also activated by toll like receptor 4 (TLR4) during periods of hyperlipidemia and hyperglycemia, further increases the expression and recruitment of inflammatory cytokines to specific tissues, including the pancreatic-β cells playing an important proapoptotic role in cytokine induced β-cell death [258]. NF-κB also upregulates protein tyrosine phosphatase 1B and suppresses the activation of the suppressor of cytokine signaling (SOCS3) in hypothalamic agouti-related proteins, both major regulators of insulin and leptin signaling [259,260]. Interestingly, NF-κB may also have a protective role against obesity and insulin resistance, as mice overexpressing NF-κB had

reduced weight gain and maintained insulin sensitivity after a high-fat diet challenge [261], indicating that the positive role of NF-κB may have a crucial tipping point to its proinflammatory actions.

Elevated concentrations of IL-1β and IL-6 have also been associated with obesity and insulin resistance [262]. In mice, the release of IL-1β from adipose tissue macrophages resulted in the binding of IL-1 receptors on the bone marrow myeloid progenitors, stimulating the increased production of monocytes and neutrophils [263]. Further, *in vitro* experiments suggest that IL-1β directly damages pancreatic β cells and inhibits insulin production and release [264]. IL-1 β also acts as its own positive feedback mechanism, upregulating both itself and IL-1 pancreatic receptors, further exacerbating inflammatory induced pancreatic damage [265]. Selectively blocking the trans-signaling pathway of IL-6 in mouse models of obesity resulted in a reduced number of macrophages in adipose tissue, indicating that IL-6 is a strong chemotactic for macrophage recruitment [266]. IL-6 has also been shown to inhibit adiponectin secretion (see *Adipokines*) and stimulates the synthesis of acute phase protein [267]. IL-6 may also promote insulin resistance by reducing the transcription or inducing the phosphorylation of IRS [268]. As with NF-κB, it is important to note that IL-6 may also have beneficial effects in preventing obesity and insulin resistance, as IL-6 knockout mice develop mature obesity and insulin insensitivity [269]. IL-6 has also been shown to activate AMPK which is known insulin sensitizer and infusions of IL-6 during EHC resulted in improved insulin sensitivity [270].

The interrelationship of inflammatory mediators in obesity and insulin dysregulation has also been investigated in horses with EMS but remains unclear. Histological examination of adipose tissue in obese, hyperinsulinemic ponies revealed a marked degree of hypertrophy [271,272] and macrophage infiltration [271] compared to the adipose tissue of obese ponies, indicating adipose dysfunction. Studies evaluating inflammatory cytokine mRNA or protein expression levels have identified significant differences between nuchal ligament, abdominal and subcutaneous fat, suggesting that adipose tissue depots have unique biological behavior [272-275]. Horses with EMS were found to have marked increases in TNFα, IL-1β, and CCL2 in both peri-renal and retroperitoneal fat depots [272], and IL-6 was significantly increased in the subcutaneous fat [271]. When comparing

insulin sensitive versus insulin insensitive horses, one study concluded that the nuchal ligament adipose tissue contributed the most to EMS pro-inflammatory profile, with an increased expression of IL-1β and IL6 in insulin resistance horses despite no difference in other adipose tissue deports or systemic inflammatory mediators [274].

Differences in serum cytokine levels have also identified statistically significant differences between obese and/or EMS horses but with conflicting results. In a group of 110 light breed horses, serum amyloid A concentrations were correlated with BCS and baseline insulin concentrations after correcting for age and sex, although correlations were not identified for TNFα or IL-6 [18]. Decreased concentrations of TNFα, IL-6, and IL-1 were identified in obese horses; however, endogenous cytokine levels were not associated with insulin levels [276]. These contrasted with several studies which identified (i) higher concentrations of TNFα in previously laminitic ponies versus non-laminitic ponies [277], (ii) higher concentration of TNFα and IL-6 in a group of obese, hyperinsulinemic ponies versus obese ponies [271], and (iii) an inverse association between TNFα and insulin sensitivity in 60 mares after adjusting for BCS and percent fat [278]. Interestingly, neutrophil oxidative burst activity was found to be markedly increased in obese, hyperinsulinemic horses. However, in the same study peripheral cytokine gene expression was lower for IL-1 and IL-6 and there was no difference between groups in cytokine response after an inflammatory challenge, potentially reflecting a lack of a direct inflammatory response due to obesity and hyperinsulinemia in horses [276]. Thus, the role of adipose tissue and inflammatory mediators in insulin dysregulation is still under investigation.

Adipokine Concentrations: Adipokines have both local and systemic effects and function as part of a complex set of physiological control systems with roles in regulating energy metabolism, cardiovascular functions, reproduction, inflammation and immunity [279]. Through the dysregulation of adipokines, adipose tissue contributes to insulin dysregulation and the pro-inflammatory state associated with metabolic syndrome via an increase in leptin and decrease in adiponectin.

Leptin, a proinflammatory cytokine, is predominantly produced in adipose tissue with circulating levels closely correlating with body mass index, making hyperleptinemia a

reliable marker of obesity across species [280]. In horses, hyperleptinemia has been associated with measurements of obesity and/or weight gain and has been shown to decline with weight loss [18,32,42,132,281-284]. Although the improvement in leptin concentrations was not significantly different between 24 obese ponies who were subjected to a weight loss regimen of diet or a combination of diet and low-intensity exercising [132], in a group of 10 obese horses, weight loss induced by moderate exercise improved plasma leptin levels and insulin sensitivity over weight loss induced by diet alone [281]. In a population of 127 Andalusian horses with normal basal insulin levels, leptin concentrations correlated with BCS and ultrasound measurements of subcutaneous fat at 75% of neck length [285]. Interestingly, both adiponectin and leptin were shown to be differentially expressed in fat depots with the highest levels of leptin identified in the nuchal ligament adipose tissue [273].

Leptin is coded by the *Ob* gene which is regulated by a variety of metabolic and inflammatory mediators. The proinflammatory cytokines, TNFα, IL-1β, and IL-6, increase gene transcription within adipose tissue [280], and binding of glucocorticoids or peroxisome-proliferator-activated receptor-γ to the *Ob* promoter region increases expression of leptin mRNA [279]. Two of leptin's primary functions are to decrease energy intake and increase energy expenditure via the hypothalamic satiety centers. Binding to the leptin receptors leads to activation of the Janus kinase (JAK) tyrosine kinases which activate the signal transducer and activator of transcription 3 (STAT3) [286]. The activation of these pathways inhibits the orexigenic factors neuropeptide Y, galanin, galanin-like peptide, orexin, and agouti related proteins while stimulating the anorexigenic factor α-melanocyte stimulating hormone via activation of pro-opiomelanocortin [287]. Leptin also serves important roles in angiogenesis, suppression of apoptosis, modulating insulin sensitivity and regulation of reproduction, inflammation, and immune functions.

Under physiological conditions, high levels of leptin suppress the appetite; however, obesity leads to leptin resistance, or the reduced ability for high levels of leptin to regulate energy homeostasis, which is selective for leptin's metabolic functions. This leads to weight gain and anorexic resistance and contributes to several obesity related co-morbidities including cardiac, renal and vascular dysfunction due to leptin-mediated

sympathetic nervous system hyperactivity, decreased nitric oxide generation and increased endothelin production [287,288]. In obese horses, evidence for leptin resistance has been suggested based on proportionally higher levels of leptin in comparison with BCS, as well as a lower anorexic effect in individuals with higher plasma leptin concentrations [289].

Several pathways for the development of leptin resistance have been described, effecting either the transport of leptin into the central nervous system, impaired leptin signal transduction on target tissues, downregulation of leptin receptors, or alterations in leptin-induced downstream signaling mechanisms [287,290]. Normal uptake of leptin into the blood brain barrier is due to leptin transporters within the brain capillary epithelium and choroid epithelial cells; prolonged hyperleptinemia due to obesity has been proposed to lead to saturation of these transports. blunting additional leptin uptake into the central nervous system [291]. Increased circulating fatty acids and TNFα has also been shown to lead to an over expression of occludin, resulting in the increased cell-to-cell adhesion of the hypothalamus tight junctions, inhibiting paracellular transport of leptin into the brain [292]. Further, endoplasmic reticulum stress has been shown to markedly reduce leptin-induced STAT3 activation, hindering the satiety Ob-Rb-STAT3 pathway [293]. Additional mechanisms which have been proposed include: increased expression of SOC3 and PTP1B within the hypothalamus and alterations within the hypothalamic phosphodiesterase-3B-cAMP and Akt-pathways [294].

As discussed in the *Management* section, weight loss in obese individuals has been shown to improve hyperglycemia and insulin sensitivity, and these beneficial results were originally thought to be due to the direct effects of decreased adipose mass; however, it has been proposed that leptin signaling can influence glucose regulation independent of this mechanism [288]. Both *in vitro* and *in vivo* mouse studies have provided strong evidence that leptin regulates glucose homeostasis via (i) interaction with the agouti-related protein neurons which are GABAergic neurons within the central nervous system [295], (ii) the adipoinsular axis in which insulin stimulates leptin production in adipose tissue and, in turn, leptin inhibits insulin secretion via central and direct actions on pancreatic β-cells while promoting hepatic glucose synthesis [296], (iii) inhibition of glucagon secretion [297], and (iv) promoting glucose uptake and utilization in skeletal muscle [298].

Importantly, the direct versus indirect effects of leptin on target cells or tissues is still unclear with conflicting results for many of these mechanisms across studies and are still under investigation [288].

Leptin's role in the innate and adaptive immune response has also been described. Leptin promotes the activation of neutrophils, monocytes and macrophages, which stimulates the production of TNFα and IL-6 [299]. Leptin further upregulates TNFα, IL-6 and IL-10 via activation of the P38 and JAK2/STAT3 pathways [300,301] and there is evidence that leptin has a role in activation or production of c-reactive proteins [302].

Given leptin's correlation with obesity and the controversial role of obesity in metabolic syndrome, leptin as a biomarker for insulin dysregulation is still under debate. In humans, leptin levels were positively correlated with HOMA-IR and BMI in young Algerian adults with metabolic syndrome [303]. Leptin levels were able to differentiate between obese cardiovascular disease patients with or without type II diabetes. In this population, individuals with a concurrent diagnosis of type II diabetes had significantly higher leptin levels compared to those with an exclusive diagnosis of cardiovascular disease [304]. Evaluation of 123 Egyptian patients with metabolic syndrome and 123 controls revealed lower adiponectin, higher serum leptin concentrations, and higher leptin: adiponectin ratios in individuals with metabolic syndrome; however, the leptin: adiponectin ratios were more sensitive, and the authors advocated for its use as an early biomarker for metabolic syndrome [305]. In horses, leptin levels have been found to be higher in horses with EMS and associated with insulin dysregulation [24,272] and was identified as a risk factor for the development of laminitis [16]. However, the role of obesity and insulin dysregulation were not parsed out in these studies. When comparing a group of 15 obese ponies, leptin levels were found to be similar in both the insulin resistant and insulin sensitive groups and both groups had equivocal improvements in leptin reduction after weight loss [306].

Adiponectin is produced almost exclusively by mature adipocytes and is one of the most highly expressed genes in white adipose tissue [279]. Gene expression of adiponectin is regulated by the transcription factors C/Eps, sterol regulatory element binding proteins, and PPARy and has been shown to be upregulated during periods of starvation and downregulated during periods of overnutrition [307-309]. In contrast to leptin, adiponectin

is negatively correlated with BMI/BCS and has been shown to increase with weight loss in both humans and horses [132,310].

One of adiponectin's primary roles is as an anti-inflammatory adipokine and hypoadiponectinemia associated with metabolic syndrome is thought to contribute to its proinflammatory state [311]. Adiponectin acts as an anti-inflammatory by inhibiting the production of TNFα from macrophages and restricting movement of monocytes into the subendothelium by reducing the expression of endothelial cell adhesion molecules [312]. It also promotes signaling through the endothelial cAMP-PKA-dependent mechanisms which prevent the activation of the NF-κB [313] and attenuates hyperglycemia-induced production of reactive oxygen species [314]. Further, studies have shown that adiponectin stimulates IL-10 production, a potent anti-inflammatory cytokine, and increases the production of tissue inhibitor metalloproteionase-1 [315,316]. In contrast, both TNFα and IL-6 negatively regulate adiponectin expression [311].

Adiponectin also serves a primary role in insulin regulation. In mouse models, adiponectin deficient mice showed insulin resistance, glucose intolerance, and dyslipidemia [317], and adiponectin supplementation reestablished insulin sensitivity in KKAy models of metabolic syndrome (mice expressing high levels of agouti proteins fed high fat diets) [318]. In rhesus monkeys, hypoadiponectinemia was shown to parallel insulin resistance prior to the onset of type II diabetes [319]. Proposed mechanisms behind adiponectin induced insulin sensitivity includes adiponectin receptor1 (AdipoR1) activation of AMPK in skeletal muscle, liver and adipocytes. Activation of AMPK enhances glucose uptake by promoting translocation of the GLUT4 transporter to the cell membrane surface, and fatty acid oxidation via the inactivation of acetyl CoA carboxylase [279,320]. Further, binding of adiponectin to AdipoR2 enhances the hepatic effect of insulin and inhibits hepatic gluconeogenesis, presumably through the AMPK and PPAR-α ligand pathways [321,322]. Notably, high molecular weight (HMW) adiponectin has been shown to the be the primary biologically active protein, representing a more sensitive biomarker for metabolic dysfunction over total adiponectin levels [323]. It has been proposed that hyperinsulinemia contributes to hypoadiponectinemia by promoting activation of serum reductase which triggers the dissociation of HMW adiponectin to low molecular weight adiponectin [324]

as well as contributing to adiponectin resistance by reducing the expression of AdipoR1 and AdipoR2 via the phophoinositide 3 kinase/Fox01 dependent pathways [325].

In humans, hypoadiponectinemia has been associated with the development of metabolic syndrome and type II diabetes [242,326], and has been observed in a variety of states frequently associated with insulin resistance including dyslipidemia, cardiovascular disease, and hypertension [320]. It has also been found to be positively correlated with HDL and negatively correlated with fasting insulin, LDL, and triglycerides concentrations [310,327]. Similarly, in horses, adiponectin levels were highly correlated with insulin resistance independent of obesity [283,306]. In a group of large breed horses, adiponectin levels were significantly lower in horses supplemented with a diet high in carbohydrates compared with those fed a high fat diet, despite both groups showing an increased BCS and total fat mass. The horses fed the cereal based diet also had the largest degree of insulin insensitivity, indicating a role for adiponectin in the development of insulin dysregulation independent of obesity [48]. Further, low levels of adiponectin in conjunction with hyperinsulinemia were found to accurately predict the development of laminitis in a group of 446 obese ponies at 1, 2, and 3 years post sampling [328]. These results suggest that adiponectin has a distinct role in insulin dysregulation independent of obesity and represents a valuable biomarker for insulin dysregulation.

Thus, the pathophysiology behind EMS is not fully understood but likely due to a combination of complex downstream interactions and feedback mechanisms between inflammatory mediators, and fatty acid, insulin, and glucose metabolism.

The Genetics of Metabolic Syndrome

Population predilections, high familial incidence, and identification of high-impact genetic variants in individuals with severe, metabolic phenotypes have all supported the hypothesis that metabolic syndrome is a complex trait with a strong genetic basis [329,330]. A metabolically thrifty genotype has been hypothesized in the horse, which proposes that genetic variations in metabolism allowed for an advantageous adaptation for survival during periods of scarce feed and harsh climate conditions in undomesticated horses [39]. However, after the agricultural revolution, horses were presented with consistent feed, abundant soluble carbohydrates, and a more sedentary lifestyle, leading to derangements

in metabolism associated with EMS and obesity. This theory was based on the thrifty genotype hypothesis in humans stating that obesity and metabolic syndrome are the result of genetic variants which allowed for human ancestors to survive during periods of poor nutrition by increasing adipose stores during the reciprocal period of food abundance [40]. Extension of this hypothesis include the thrifty epigenotype hypothesis which advances the notion that all individuals have a thrifty genotype but that variations in phenotype expression are due to in utero epigenetic modifications resulting from the influence of environmental risk factors [331]. However, the drifty gene hypothesis suggests that genetic predisposition to obesity is not the result of positive selection but predominantly due to random genetic drift based on current population prevalence of obesity and MetS [332]. Based on the complexity of mammalian evolution, it is likely that the genetic origins of metabolic syndrome are the manifestation of a combination of these hypotheses and additional factors including social pressures, pleiotropic effects, and microbial influences [333].

The identification of the specific alleles underlying metabolic syndrome will allow for a better understanding of the fundamental pathogenesis across species. Moreover, the promise of precision medicine, or tailored treatment and management regimens based on an individual's unique combination of genetic and environmental risk factors, has instigated a drive toward identifying the genetic risk factors of many complex diseases, including metabolic syndrome. However, unlike Mendelian (simple) traits where the genetic variation can be explained by a single gene with a well-defined mode of inheritance, as a complex trait metabolic syndrome is likely the result of the combination of dozens to hundreds of genetic alleles with variable allele frequencies, penetrance, environmental influences and gene by environment interaction. Further, although Mendelian diseases are often the consequence of high-impact variants within protein-coding genes, alleles contributing to complex traits have variable effect sizes and are primary non-coding variants that presumably affect gene regulatory elements [334-336]. Therefore, the identification of the genetic variants of complex traits requires large genetic and phenotypic data sets ("big data") and statistical approaches that can account for the numerous variables influencing these traits. Recent advances in sequencing technology have provided researchers the cost-effective, high throughput means required to collect

genomics data in a large number of individuals, allowing for the genetic risk factors underlying both human and equine metabolic syndromes to start to become unraveled.

Heritability of Metabolic Syndrome in Humans: Often, one of the first steps in the investigation of a complex trait is determining how much genetics is contributing to that trait by estimating heritability, which is the percentage of phenotypic variation that can be explained by genetics. Typically, heritability is estimated by calculating narrow sense heritability (h2) which is the proportion of additive genetic variance over the total phenotypic variance [337]. For example, a trait with an h2 of <20% is considered to have low heritability and indicates that other factors such as the environment are having a larger impact on that trait; whereas, traits with a moderate (21-40%) to high (>40%) h2 indicate that genetics is having a large impact on that trait [338]. This information is imperative for justifying the continued investigation of the genetic risk factors of complex traits, as well as ensuring the studies are appropriately powered and designed for identification of the specific alleles contributing to low, moderately, or highly heritable traits.

In humans, the heritability of metabolic syndrome has been extensively studied using pedigree data, with most studies estimating the heritability of several factors including biochemical measurements and/or components of metabolic syndrome such as the homeostatic model assessment of insulin resistance (HOMA-IR; proxy for insulin resistance [339]) and the metabolic syndrome score [340]. Across studies, the range of heritability estimates for MetS as a binary trait (typically defined based on the presence of three or more components of MetS) was 0.11-0.38 [341-344], and the ranges for quantitative traits were 0.43-0.51 for fasting insulin, 0.14-0.81 for fasting glucose, 0.17-0.24 for glucose post oral sugar challenge, 0.38-0.48 for HOMA-IR, 0.15-0.29 for MetS scores, 0.39-0.68 for BMI, 0.27-0.46 for waste circumference, 0.42-0.62 for total cholesterol, 0.42-0.63 for high-density lipoproteins, 0.58 for low-density lipoproteins, 0.17-0.60 for triglycerides, 0.28-0.55 for leptin, 0.51 for adiponectin, and 0.12-0.38 for diastolic and 0.16-0.28 for systolic blood pressure [341-352]. Differences in population predilection were supported by Musani et al who calculated the heritability of metabolic syndrome from three different study cohorts: the Jackson Heart Study (n=1404 African Americans), the Take Off Pounds Sensibly Study (n=1947 Caucasians), and the Princeton

Lipid Research Study (n=229 African Americans and 527 Caucasians). The authors reported a separation in heritability estimates by race with larger heritability estimates for MetS among African Americans compared with Caucasians. However, when evaluating systolic blood pressure, although African Americans had higher values, the correlation coefficients were approximately half of the Caucasians and the h2 was 0.09, indicating there was a lower genetic contribution and higher environmental influence for systolic blood pressure in this group [353].

However, it is important to note that although pedigree analysis has been historically used for calculating heritability, the estimates have been shown to overestimate h2 as a consequence of (i) small populations of highly related individuals often confounded by a shared environment; (ii) assortative mating; (iii) pedigree errors; and (iv) ascertainment bias (selection of pedigrees that have a high proportion of affected individuals) [354]. Furthermore, to achieve an unbiased estimate of genetic variance, the data must be representative of the general population and include all potential confounders [337]. The availability of algorithms to calculate h2 using single nucleotide polymorphism (SNP) genotype data has allowed for a more precise estimate of h2 by using data from a large group of unrelated individuals across multiple environments [355].

Comparing both pedigree and SNP based approaches, the MetS score was estimated to have a heritability of 0.27-0.34 using a pedigree approach and 0.24-0.25 using a SNP based approach in the Ogliastra population; consistent with what would be expected between pedigree and SNP based h2 estimates [347]. The authors also calculated the difference between SNP based approaches from unrelated versus closely related individuals in two separate populations and concluded that the additive genetic variation was a major contributor to MetS score but that common sibs-household effects had a moderate impact on trait variation, providing further evidence for the bias in estimates amongst closely related individuals with shared environments [347].

Importantly, SNP-based heritability estimates are limited to the genetic variability that can be explained by the common SNPs present on the genotyping arrays, but cannot account for causal variants that are not inherited together (in linkage disequilibrium) with these SNPs; nor can it include other genetic variations contributing to the disease phenotype such

as insertions, deletions or copy number variants [356-359]. For example, Vattikuti et al used a linear mixed-effects model with SNP data to estimate the heritability of metabolic syndrome using data from the Atherosclerosis Risk in Communities and Framingham Heart Study. By comparing the SNP based heritability in both related and unrelated individuals, the authors determined that ~40% of h2 was explained by the common SNPs, which was a larger percentage than previously identified for metabolic phenotypes. Overall, they estimated the heritability of BMI at 0.34, basal glucose at 0.33, basal insulin at 0.23, triglycerides at 0.47, high-density lipoproteins at 0.48 and systolic blood pressure at 0.30. These results were consistent with the h2 estimates noted above for the pedigree analyses albeit at the lower end of the range for most traits [360].

Direct comparisons across studies must be taken with caution given that heritability estimates are highly dependent on the represented population since the effects of environmental variance and additive and non-additive genetic variances are population-specific, and each study represented differences in ethnic group, study design, diagnostic criteria, and pedigree versus SNP based analysis. However, key points that can be extrapolated are that (i) within study populations, several phenotypes were under considerable genetic influence (moderately to highly heritable) while other phenotypes seemed to be affected by unmeasured nongenetic factors such as the environment (low heritability), (ii) differences in heritability estimates between ethnic populations likely represents differences in the genetic risk factors and metabolic profiles between these groups, and (iii) metabolic syndrome is a complex trait with a large portion of the phenotypic variance being explained by genetics.

It should also be noted that although heritability estimates provide valuable information on how much genetics is contributing to a trait, they do not provide information on the number of genes involved, the interaction or penetrance of these genes, nor the mode by which these genes are inherited. Therefore, additional analyses are required to identify the specific alleles contributing to metabolic syndrome.

Family-Based Linkage Studies: Genetic linkage analysis has been used to identify broad regions of the genome that harbor disease risk alleles in related individuals. Genetic linkage occurs when two loci are transmitted together from parent to offspring more often

than expected under independent inheritance and typically extends over large regions of the chromosome [361]. The probability that linkage is occurring between two loci can be estimated using the recombination fraction. This calculation is based on the likelihood that the segregation between two loci would have been broken during meiosis (the farther two loci are from each other on a chromosome, the more likely a recombination event will occur between the loci). Family-based linkage analysis is based on the rationale that if a trait is occurring at high frequency within a family, then it is likely that affected individuals will share haplotype(s) that are identical by descent in the region(s) harboring the disease-causing allele [361,362]. Thus, linkage analysis can be used as an initial step in the genetic investigation of a trait to find broad, chromosomal regions associated with that trait.

Linkage analysis has been used to identify numerous loci contributing to MetS and its components. For example, in 2,209 Caucasian individuals representing 507 nuclear families, a QTL on chromosome 3 (3q27) was strongly linked to BMI, waist circumference, hip circumference, weight, insulin and insulin:glucose ratio. A second QTL was identified on 17p12 linked with plasma leptin concentrations. Within these QTL were contained the biological candidate genes *GLUT2* and the Catalytic α polypeptide of PI3K [363]. In addition, linkage analysis using data from the National Heart, Lung and Blood Institute Family Heart Study (2467 individuals representing 387 three-generation families) identified a pleiotropic locus on chromosome 2 that was linked with BMI, waist-to-hip ratio, TG, HDL, and HOMA-IR [364]. In a study evaluating 250 German families, a locus on chromosome 1 (1p36.13) was linked with a diagnosis of MetS [365]. This region was also identified in Mexican Americans with gallbladder disease in which 46% were diagnosed with type II diabetes. After correcting for type II diabetes, this locus was no longer significant [366]. Further, in a group of 566 nondiabetic Mexican Americans representing 41 extended families, 1p36.13 was linked with body-size adiposity [367], providing further support for a metabolic role at this locus. Additional QTL identified by linkage analysis for MetS include: (i) 15q in 707 individuals from the Quebec Family Study [368], (ii) chromosome 2q12.1-2q13 in Caucasians and 3q26.1-3q29 in 53 Mexican Americans using data from the GENNID Study[369], and (iii) 10p11.2 in 456 Caucasians and 1p34.1 for 217 African Americans from HERITAGE family study [370].

Although linkage analysis has been successful in identifying loci contributing to MetS, this type of analysis has several limitations. First, linkage analysis identifies significant loci but subsequent analyses to discover the specific risk alleles can be hindered by the large size of the chromosomal regions shared by family members [371]. Second, this type of analysis relies heavily on multigenerational pedigrees with a large number of affected individuals, which is often confounded by shared environments, gene-environment interactions, and social economic status. Third, linkage studies have been shown to have a significant loss of power in the presence of genetic heterogeneity, thus hindering the analysis of polygenic complex traits [372]. Finally, risk loci may be family specific and not relevant to the general population. Therefore, although family-linkage studies have provided useful information in several genetic loci contributing to MetS, the identification of specific risk alleles from this data has been sparse and replication has been limited [373]. This is likely partially due to differences in case definitions and the fact that, based on the time of publication, many studies were using obesity as a primary factor for diagnosing MetS.

Genome-Wide Association Analysis: Genome-wide association analyses (GWA) have been used to overcome the limitations in family-based linkage studies by using SNP genotype data from a large number of unrelated individuals to identify key regions of the genome that harbor risk alleles. The SNPs present on genotyping arrays are considered common (ancestral) SNPs and are typically neutral polymorphisms with no effect on the trait studied. However, these SNPs can be used to tag the causative risk alleles through linkage disequilibrium (SNPs which, in the population as a whole, are found on the same haplotype more often than expected [361]), generating an association between the region of the genome and the trait of interest.

The underlying rational for GWA is based on the common disease, common variant hypothesis, which proposed that a significant proportion of the genetic variation in common traits could be explained by allelic variants that are present in more the 1-5% of the population [374]. Although the allelic architecture of some diseases follows this pattern, most common variants have been found to contribute only a small portion of the phenotypic variation in complex diseases, an observation that has been termed the missing

heritability [359,375-377]. Several explanations for missing heritability have been proposed including (i) insufficient power in GWA to detect variants of small effect, (ii) the common disease, rare variant hypothesis which stated that complex traits are highly polygenic and affected by a large number of rare variants, (iii) overestimation of heritability in twin and family due to confounders such as shared environment or gene-environment interactions, and (iv) the omnigenic model which states that virtually all active genes contribute to complex traits; however, variants with moderate to high effect were more likely to be enriched in specific genes or pathways that play a direct role in disease whereas the low effect SNPs, contributing the most to heritability, were more likely to be spread across the genome and not near genes with disease specific functions [359,375-377]. However, identification of common risk variants with moderate to large effect, despite not explaining all of the heritability of a trait, are still important for understanding the underlying pathophysiology of a trait, as well as finding potential therapeutic targets that would benefit a larger percentage of individuals.

Further, an important consideration when assessing the validity of a GWA study is replication of results within independent populations. Mixed linear models are one of the most common statistical methods used to perform GWA as they account for population stratification and relatedness within the cohort by including a genetic relationship matrix (GRM) as a random effect [378,379]. Within these models, each SNP is tested individually for an association with the phenotype of interest and significant associations are determined after a correction for multiple testing. However, it has been shown repeatedly that in the initial association study, the estimated SNP effect is inflated as only modest correlations have been identified between estimated effects sizes from the initial and subsequent association analyses of the same trait [380,381]. Many factors have been implicated in contributing to this lack of reproducibility including population stratification, unaccounted for covariates, phenotypic heterogeneity, selection biases, or the phenomenon known as the "winner's curse" where associations with the strongest effect are overestimated typically due to a small population size [382,383]. Notably, it is important to recognize that true differences in allele effect size can exist between populations and it is essential to assess for differences in populations between the initial and replicate studies [382]. Therefore, it is imperative that due diligence be met with validation of identified candidate

regions and risk alleles prior to the marketing of genetic testing, especially with complex traits.

Despite some of these caveats, over the past few decades, GWA have been used to identify thousands of associations between SNPs and complex traits across species [384] and has provided important information about the genetic architecture of a trait by identifying the number of loci contributing to the trait and estimating their effect size and allele frequencies [385]. In humans, GWA has been used to start to identify the genetic contribution of metabolic syndrome across multiple populations and environments. As with heritability, studies have defined their outcome variable for metabolic syndrome as a binary trait or components of MetS as quantitative traits. Although the extent of GWA articles published evaluating individual components of MetS, or those primarily assessing obesity, is beyond the scope of this review, several key studies evaluating MetS as an entity or as a combination of traits will be discussed.

A two-stage GWA was performed in a population of Indian Asian men to identify common genetic variation for MetS risk. During stage one, the authors genotyped 2700 individuals with MetS, as characterized by the International Diabetes Federation, for 317,000 SNPs. Based on results from Bayesian association analyses, 1500 SNPs were chosen to be genotyped in an additional 2300 individuals for stage two. For components of MetS, the authors identified two loci associated with HDL metabolism which contained variants within cholesteryl ester transfer protein and lipoprotein lipase genes and been previously identified as associated with MetS, as well as five novel loci. However, they did not identify regions shared across traits which they concluded indicated little evidence of a common genetic basis for MetS traits [386]. This was similar to findings from a study evaluating four Finnish cohorts comprising of 2,637 MetS cases and 7,927 controls, in which the authors found little evidence for pleiotrophy across traits. In addition, using a GWA meta-analysis, the authors identified a known lipid locus, including the *APOA1/C3/A4/A5* gene cluster, in all four study samples; serum metabolite analysis further supported this region as associated with VLDL, triglycerides, and HDL metabolites. An additional 22 known loci were identified, the majority associated with lipid metabolism.

The authors concluded that lipid metabolism pathways have key roles in the genetic background of MetS [387].

The lack of pleiotropic loci identified in these studies contrasted with findings in three later studied which could reflect population differences, the limited quantitative traits evaluated in the previous studies, how the authors defined a MetS case, or the statistical techniques used to analyze the data. The first study analyzed data from a population of 1,427 Africans from Ghana and Nigeria, and then followed-up with 2,475 samples from ARIC study. The authors performed GWA for six metabolic syndrome traits and identified two loci unique to individuals of African-ancestry, one of which was considered an at-risk locus and the other a protective locus, as well as a non-African specific loci located near *KSR2* which had a pleiotropic effect on triglycerides and measures of blood pressure [388]. In the second study, the authors utilized data from 19,486 European Americans and 6,287 African Americans. To better characterize the clustering of metabolic abnormalities commonly associated with MetS, they evaluated six phenotypic domains, which encompassed 19 quantitative traits, and analysis was performed using a multivariate association approach. The authors identified 19 significantly loci of which three were pleiotropic (associated with multiple phenotype domains) and located in or near apolipoprotein C1 (*APOC1*), phospholipase C gamma, and *BRAP* genes [389]. The third study utilized a subset-based meta-analysis approach in a population of 15,148 African Americans from the Population Architecture using Genomics Epidemiology study. The investigators identified 1 glucose and 4 lipoid loci associated with a diagnosis of MetS, of which three were replicated in a population of 5,172 Hispanics and one novel, pleiotophic loci was specific to African Americans. The authors also identified evidence for pleitrophy for *APOE, TOMM40, TCF7L2,* and *CETP* [390].

A multivariate GWA approach was also performed using data from the STAMPEED consortium (seven studies comprising 22,161 individuals of European ancestry). In this study, the authors compared every combination of pairwise comparisons between MetS components and, for each combination, individuals exceeding National Cholesterol Education Program defined thresholds for both traits were considered affected. Overall, a total of 29 common variants were associated with MetS or a pair of traits, with all but two

of the bivariate associations including alterations in lipid metabolism. However, the effect of the top 16 SNPs was relatively small and explained ~9% of the variance in triglycerides, 5.8% of HDL, 3.6% baseline glucose, and 1.4% of systolic blood pressure. The authors concluded that only a small portion of the covariation in traits could be explained by the common SNPs, and that the effects of genetic variants on lipid levels were more pronounced than for other traits [391].

Kong et al utilized GWA to investigate gender-specific loci for MetS and its components in a population of 9,932 Korean females. The authors defined MetS cases based on the criteria established by the International Diabetes Foundations, which resulted in 2,276 cases and 1,692 controls. GWA identified 14 loci showing moderate association for MetS in females but not in males. The authors also identified female-specific loci for fasting glucose concentrations and HDL cholesterol. This study provided evidence for sex-specific genetic architecture associated with MetS [392].

Although GWA studies evaluating MetS have had conflicting results, several consistent and key findings indicate: (i) that both common and rare variants contribute to MetS [393], (ii) different populations have both shared and unique loci, and (iii) a large number of variants are related to lipid metabolism, a result further replicated in GWA meta-analysis [394]. This information has been invaluable to understanding how genetics is contributing to MetS; however, despite decades of research, the fundamental genetic basis of this syndrome is still unknown which is partly due to the complexity of the phenotype, discrepancies in phenotype definition, unaccounted for environmental influences, and missing heritability. Unraveling the genes underlying thousands of loci on GWA remains one of the principal challenges in complex trait genetics.

Animal Models in Complex Disease Genetics: Naturally occurring animal models of disease can provide valuable insight into the genetic basis of complex traits in both humans and animals. Due to selective breeding, a small number of variants with a large effect size are likely contributing to a significant portion of the genetic variance of complex traits. For example, in horses four loci explain 83% of the genetic variation in height [395], a highly heritable trait with published h2 estimates up to 0.89 [396]. This is in contrast to humans where hundreds of genes with small effect control this trait [397]. Further, unlike

humans, animal management regimens are often standardized within environments. For example, horses on the same farm are typically fed the same hay, given a fixed daily amount of feed, graze the same pastures, and receive a similar level of exercise. Studies including both cases and controls from individuals housed at the same farm will effectively help to reduce bias secondary to environmental confounders. Similarly, animal models enable the manipulation of a single environmental variable or group of variables in order to parse out specific risk factors of complex diseases [398]. Finally, animal models provide a unique opportunity to collect trait-relevant multi-omics datasets on a large number of individuals within a single or limited number of visits such as genomic (blood), transcriptomic (tissue), metabolomic (serum), and microbiome (feces) data. Interrogation of a trait at multiple levels of the genome provides a powerful tool to identify potential causative changes and therapeutic targets [375,399].

Therefore, the decreased genetic diversity, controlled environment, and ability to collect large scale trait-relevant multi-omics datasets allows for layered genomics study design which cannot be recapitulated in humans. Identification of these variants within animal populations can be used to translate back to humans and provide a better understanding of the pathophysiology in both species, making comparative genomics the most efficient way of dissecting the genetic basis of complex traits [400]. The similarities between metabolic syndrome in humans and horses make horses an excellent naturally-occurring model of MetS and an ideal candidate for further exploring the genetic contribution to metabolic syndrome across species.

Evidence for a Genetic Contribution to EMS: One of the first studies evaluating the genetic basis of EMS was published in 2002 by Carter et al [13]. In this study, the authors estimated the heritability of pasture-associated laminitis using pedigree data from a single herd of 160 pure and crossbred Welsh and Dartmoor ponies. For this study, the authors grouped ponies based on whether they had a previous diagnosis of laminitis, were clinically laminitic, or non-laminitic. A total of 34% of ponies had a diagnosis of laminitis, of which there was an 8-fold lower prevalence in mature stallions versus females. The authors concluded that the observed prevalence of laminitis was consistent with the action of dominantly expressed gene(s) but with reduced penetrance due to sex-mediated factors,

age of onset and epigenetics [13]. Further, in a large, cross-bred population of horses evaluating the EMS phenotype, it was identified that a large percentage of the individual variation could not be explained by known explanatory variables [20]. The authors utilized a multi-level, multi-variate regression modeling to quantify the relative importance of environmental and individual factors. After adjusting for age, breed, gender, obesity, and season they determined that only 9.6-36.3% of the variation at the individual level was explained by these factors. The authors hypothesized that individual genetic differences were contributing to this unexplained metabolic trait variation [20].

A more recent study utilized GWA and haplotype analysis to propose a potential candidate allele in the 3' untranslated region (UTR) of the Family with Sequence Similarity 174 member A gene (*FAM174A*) for EMS and endocrinopathic laminitis in Arabian horses [401]. In this study, a significant association on chromosome 14 was identified in a genome-wide analysis for laminitis status in 64 Arabian horses. Genotypes of a single nucleotide polymorphism (SNP), BIEC2-263522 (T>C), correlated with both laminitis status and baseline insulin concentrations. In a second population of Arabian horses (n=50), BIEC2-263522 correlated with elevated body condition score (BCS) and modified insulin-glucose ratios (MIRG). *FAM174A* was identified as a candidate gene and sequencing resulted in the identification of a polymorphic guanine homopolymer region in the 3'UTR. The 11G allele was found to be in linkage disequilibrium with the original marker SNP and correlated with elevated insulin levels in their first population and BCS and MIRG in their second population. In addition, the 11G allele was found to be present in Welsh ponies and Tennessee Walking horses, but absent in Draft, Standardbred, and Thoroughbred horses [401]. In a population of Polish Arabian horses, the genotype frequency of the BIEC2-263522 SNP was found to be 51.6% for the heterozygous genotype and 16.8% for the homozygous variant [402]. However, the estimated SNP effect was not provided, significant thresholds were not corrected for multiple testing, and EMS phenotype data were unavailable for both the cross-breed and Polish Arabian populations preventing the validation of genotype to phenotype correlations in these cohorts.

Thus, there is evidence that EMS has a genetic component, but the identification of specific loci or genetic risk factors for EMS are in its infancy. It is imperative to continue to

investigate the specific genes and genetic risk factors for EMS in order to gain a better understanding of the root causes of these metabolic abnormalities. More importantly, knowledge of the genetics contributing to EMS will allow for the development of genetic tests which would enable veterinarians to evaluate a patient's risk for developing EMS by assessing the number of genetic and environmental risk factors for each individual. Understanding of a horse's risk of EMS will ultimately allow veterinarians to know which horses need frequent monitoring and would benefit from early environmental modification, as well as those that may not fully respond to environmental management alone (i.e., horses with a large number of risk variants or a modest number of variants with a moderate to large effect). In addition, this information should be used to guide owners in responsible breeding decisions (i.e., not breeding two horses who both have a high genetic risk for EMS).

Hypothesis and Objectives

Identification of the underlying genetic risk factors and key gene-by-environment interactions will improve our understanding of EMS pathophysiology and allow for early detection of high-risk individuals and intervention prior to the onset of laminitis. *We hypothesize that major genetic risk factors leading to EMS and laminitis susceptibility are shared across breeds, and that differences in the severity and secondary features of the EMS phenotype between breeds, or between individuals within a breed, are the result of modifying genetic risk alleles with variable frequencies between breeds.*

Objective 1: Estimation of the genetic contribution to metabolic traits. Heritability will be estimated using approximately 1,800,000 (1.8M) SNP genotype data from 264 Welsh ponies and 286 Morgan horses using a restricted maximum likelihood statistic with the inclusion of genetic relationship matrix corrected for linkage disequilibrium. The heritability of nine EMS biochemical measurements will be estimated, and for each trait the confounders of age, sex and season will be included based on the Akaike information criteria.

Objective 2: Identification of regions of the genome harboring EMS risk alleles. Genome-wide association analyses will be performed in a cohort of 264 Welsh ponies and 286 Morgan horses using approximately 1.8M SNP genotypes to identify loci harboring

EMS risk alleles. Within breed GWA will allow for identification of significant quantitative trait loci (QTL) within the Welsh ponies or Morgan horses. GWAS meta-analysis using the within breed GWA summary data from the Welsh ponies and Morgan horses will be performed to increase power to identify shared alleles of low to moderate effect. QTL will be given a high priority if they are shared between breeds or across EMS traits.

Objective 3: Identification of candidate genes and putative functional alleles contributing to EMS. Whole genome sequencing (WGS) in 19 Welsh ponies and 18 Morgan horses will be used to impute SNP genotype data from 264 Welsh ponies and 286 Morgan horses to WGS. QTL identified as high priority in Objective 2 will be fine mapped using linear regression analysis including the fixed effects of age and sex and the random effect of farm. Variant annotation software will be used to interrogate the fine mapped regions for putative functional alleles and a literature search will be performed to identify biological candidate genes to prioritize alleles for further follow-up.

Significance

Identification of EMS risk alleles and gene pathways has the potential to substantially expand our knowledge of EMS pathophysiology and will allow us to better predict disease, thus improving our ability to detect individuals who would benefit from management changes prior to the development of clinical signs, and prior to the development of laminitis. The focused approach, as outlined in this proposal, for using genomic data from WP will help us to identify unique loci responsible for metabolic differences between breeds (within breed approach), as well as increase our power for identifying major risk alleles shared across breeds (across breeds approach with Morgan horses).

Furthermore, based on the striking similarities between EMS and MetS, the results of this study could be used to further validate or identify additional genetic risk variants in other species with metabolic derangements, making EMS a naturally occurring polygenic animal model for MetS. In addition, unlike humans with MetS, horses with EMS are not hyperglycemic, indicating a sustained, compensated hyperinsulinemic state [17]. Identification of gene pathways could provide insights into novel treatment of uncompensated insulin resistance in humans.

Our *long-term goal* is to construct a genetic and environmental risk model to facilitate identification, management changes and early intervention in horses at high risk for developing EMS.

Chapter 2: Heritability of metabolic traits associated with equine metabolic syndrome in Welsh ponies and Morgan horses

Summary: Equine metabolic syndrome (EMS) is a complex clinical disorder with both environmental and genetic factors contributing to EMS phenotypes. Estimates of heritability determine the proportion of variation in a trait that is attributable to genetics. The objective of this chapter was to provide heritability estimates for nine metabolic traits associated with EMS in two high-risk breeds. High-density single nucleotide polymorphism (SNP) genotype data was used to estimate the heritability (h^2_{SNP}) of nine metabolic traits relevant to EMS in a cohort of 264 Welsh ponies and 286 Morgan horses. Traits included measurements of insulin, glucose, non-esterified fatty acids (NEFA), triglycerides, leptin, adiponectin, ACTH, and glucose (GLU-OST) and insulin (INS-OST) post oral sugar challenge. In Welsh ponies, seven of the nine traits had statistically significant h^2_{SNP} estimates that were considered moderately to highly heritable (h^2_{SNP} > 0.20) including: triglycerides (0.313; SE=0.146), glucose (0.408; SE=0.135), NEFA (0.434; SE=0.136), INS-OST (0.440; SE=0.148), adiponectin (0.488; SE=0.143), leptin (0.554; SE=0.132) and insulin (0.808; SE=0.108). In Morgans, six of the nine traits had statistically significant h^2_{SNP} estimates that were also determined to be moderately to highly heritable including: INS_OST (0.359; SE=0.185), leptin (0.486; SE=0.177), GLU-OST (0.566 SE=0.175), insulin (0.592; SE=0.195), NEFA (0.684; SE=0.164), and adiponectin (0.913; SE=0.181). This chapter provides the first concrete evidence of a genetic contribution to key phenotypes associated with EMS. Eight of these nine traits had moderate to high h^2_{SNP} estimates in this cohort. These data demonstrate that continued research for identification of the genetic risk factors for EMS phenotypes within and across breeds is warranted.

Note: This chapter was previously published in the *Equine Veterinary Journal* as follows: Norton, E.M.; Schultz, N.E.; Rendahl, A.K.; Geor, R.J.; Mickelson, J.R.; McCue, M.E. Heritability of metabolic traits associated with equine metabolic syndrome in Welsh ponies and Morgan horses. *Equine Vet J* 2019; 51(4): 475-480.

Introduction

Equine metabolic syndrome (EMS) is a term used to describe an interrelated group of metabolic disturbances including hyperinsulinemia, insulin resistance, dyslipidemia, and adiposity that often leads to an endocrinopathic laminitis. Both genetic and environmental factors likely play key roles in manifestation of EMS phenotypes. Support for this hypothesis comes from the accumulated evidence of a genetic component for human metabolic syndrome (MetS) [403], a clustering of metabolic traits with several similarities to EMS [7]. In horses, a dominant pattern of inheritance for pasture-associated laminitis has been suggested in a small pedigree of Welsh and Dartmoor ponies from a single farm [13]. More recently, *FAM174A* was identified as a candidate gene for endocrinopathic laminitis and insulin dysregulation by a genome wide analysis of DNA markers in a population of Arabian horses [401]. Further, acknowledged breed predispositions (including ponies, Morgan, Arabian, Peruvian Paso, Andalusian, and Tennessee Walking horses), as well as breed-specific differences in metabolic profiles [38,51], supports the hypothesis that unique genetic alleles within or across high risk breeds contribute to their susceptibility to EMS [404].

Determining which components of a disease are heritable increases our basic understanding of the condition and is a necessary step towards identifying the specific risk factors and developing genetic tests to detect high-risk individuals before they develop clinical disease. Heritability is defined as the proportion of variation in a trait that can be explained by an individual's genetics. Estimates of heritability are typically a measure of narrow sense heritability (h2), which is the ratio of the additive genetic effect (i.e., the sum of the effects that all contributing genetic variants have on the phenotype) to the total phenotypic variation in a trait. Historically, h2 has been estimated using pedigree data; however, the recent availability of high-density single nucleotide polymorphism (SNP) genotype data enables h2 to be estimated from a large population of unrelated individuals, resulting in "SNP-based heritability" (h^2_{SNP}), eliminating the reliance on accurate pedigree information and decreasing ascertainment bias [354,405].

The objective of this study was to provide h^2_{SNP} estimates of measures of insulin, glucose, adipokines and fat metabolism associated with the EMS phenotype and ACTH, accounting

for known covariates such as age, sex and season, in both Morgan horses and Welsh ponies. We hypothesized that the h^2_{SNP} estimates per trait would vary, and that traits would be mildly to moderately heritable.

Materials and Methods

Samples: History, signalment, environmental management, and EMS phenotype data were collected on 264 Welsh ponies and 286 Morgans as part of a large cross-breed study evaluating the EMS phenotype. The Welsh pony cohort was obtained from 28 farms throughout the United States and included 193 females and 71 males with a mean age of 11.5 years (age range of 2 to 33 years). The Morgan cohort was obtained from 31 farms throughout the United States and Canada, and included 184 females and 102 males, with a mean age of 12.3 years (age range of 2-29 years). Biochemical measurements collected on all individuals included insulin (INS) and glucose (GLU) after an 8 hour fast and 75 minutes post oral sugar challenge (OST) using 0.15ml/kg light Karo Syrup as previously described [406]. Additional samples at baseline included: triglycerides (TG), non-esterified fatty acids (NEFA), adiponectin, leptin, and ACTH. Season at the time of sampling was recorded as follows: winter (December-February; n=21Welsh ponies and 54 Morgans), spring (March-May; n=85 Welsh ponies and 135 Morgans), summer (July-August; n=132 Welsh ponies and 59 Morgans) or fall (September-November; n=26 Welsh ponies and 38 Morgans). See **Appendix A** for assay details and additional phenotype data. Horses with a history or phenotypic appearance of pars pituitary intermedia dysfunction were excluded from the study. This protocol was approved by the University of Minnesota Institutional Animal Care and Use Committee.

Genotype Data: Genomic DNA was isolated from whole blood or hair roots as per manufacturer recommendations (Puregene Blood Core Kit, Qiagen). 286 Morgans were genotyped on the Illumina EquineSNP50 (50K) BeadChip (54,602 SNPs), 220 Welsh ponies were genotyped on the Axiom Equine MCEc670 (670K) array (670,795 SNPs), and 44 Welsh ponies and 43 Morgans were genotyped on the MCEc2M (2M) array (2,011,826 SNPs). Haplotype phasing and genotype imputation was performed for horses genotyped on the lower density 50K or 670K arrays by using the high density MCEc2M array data and Beagle software [407,408], to yield a total of 1,923,776 SNPs in the Welsh ponies and

1,931,327 SNPs in the Morgans for quality control and pruning. Using data from Morgans genotyped on both the 50K and 2M arrays imputation concordance was determined to be 99.2%. We have previously shown an imputation concordance of 99.1% from 670K to 2M in the Welsh ponies [409].

Quality control was performed using the PLINK software package [410]. All horses passed quality control, including evaluation for discordant sex information and SNP genotyping rate (>95%). Individual SNPs that had a genotyping rate of <90%, a minor allele frequency (MAF) of <1.0%, or were outside of Hardy-Weinberg equilibrium, were removed. After genotype pruning and removal of the X-chromosome data, 1,158,831 and 1,428,337 SNPs remained for the Morgan and Welsh pony analyses, respectively.

Heritability Estimates: Biochemical measurements were tested for normality using a normal probability plot and Shapiro test. Insulin, INS-OST, TG and ACTH were log-transformed and NEFA, adiponectin, and leptin were square root transformed to achieve normality. Glucose and GLU-OST were not adjusted. All traits were treated as quantitative response variables.

Covariates for inclusion in the h^2_{SNP} estimates were selected using linear regression models with biochemical trait as the response variable and covariates as predictors. Age, sex and season were evaluated as covariates using model selection. Analysis was performed for all possible combinations of covariates, and model selection was based on the lowest value for the Akaike information criteria (AIC) (**Appendix A: Supplementary Table A1**).

Heritability estimates were obtained via a mixed linear model analysis performed with Genome-wide Complex Trait Analysis (GCTA) software [357]. Briefly, GCTA calculates h^2_{SNP} by fitting all SNPs simultaneously into the model using a restricted maximum likelihood estimation (REML), including a genetic relationship matrix (GRM). The GRM accounts for relatedness within the cohort by calculating pairwise comparison of relatedness based on the number of alleles two individuals inherit from a shared ancestor. For this analysis, the GRM was created using the software program Linkage Disequilibrium Adjusted Kinship (LDAK) [411], which adjusts the estimate based on the linkage disequilibrium (LD) between SNPs (i.e. an adjustment for lack of SNP independence due to SNPs being inherited together) to generate a weighted GRM (wGRM)

[411]. Heritability estimates were also calculated with GCTA and a standard GRM, as well as using LDAK's REML algorithm with the wGRM (**Appendix A: Supplementary Methods**). A Holm-Bonferroni correction was used to adjust for multiple comparisons to increase statistical power while minimizing the familywise Type 1 error rate (**Appendix A: Supplementary Table A2**).

Random subsetting of the data was performed to determine if individuals with close relationships unknown to the researchers (cryptically related individuals) were inflating the h^2_{SNP} estimate. Ten percent of the population was randomly removed from the analysis using the software program R's random number generator [412] and h^2_{SNP} estimates were calculated using GCTA. The average h^2_{SNP} and SE from 100 replicates per trait for each breed was calculated.

Results

Heritability Estimates: The covariate combinations of age, sex and season determined to be optimal for h^2_{SNP} estimates per trait in each breed are presented in **Table 2.1A**. In the Welsh ponies, seven of the nine biochemical traits had h^2_{SNP} estimates with p-values that exceeded the Holm-Bonferroni corrected cut-off (**Table 2.1B**). From lowest to most highly heritable these were: triglycerides (0.313), glucose (0.408), NEFA (0.434), INS-OST (0.44), adiponectin (0.488), leptin (0.554), and insulin (0.808). GLU-OST (0.226) and ACTH (0.305) did not meet the Holm-Bonferroni threshold for significance.

Six of the nine EMS traits in the Morgans had h^2_{SNP} estimates with p-values that exceeded the Holm-Bonferroni cutoff (**Table 2.1B**). From lowest to most highly heritable these were INS-OST (0.359), leptin (0.486), GLU-OST (0.566), insulin (0.592), NEFA (0.684), and adiponectin (0.913). Glucose (0.208), TG (0.273), and ACTH (0.408) had p-values that did not meet the threshold for Holm-Bonferroni corrected significance.

Random Subsetting of the Heritability Estimates: In the Welsh ponies, the differences between the mean h^2_{SNP} estimates obtained by randomly subsetting the data 100 times and the h^2_{SNP} estimates obtained from the entire cohort were very small (0.009-0.034; **Table 2.2**). There was a larger range (0.002-0.075) between the entire cohort and mean h^2_{SNP}

estimates in the Morgans (**Table 2.2**). However, the differences between the h^2_{SNP} values in the majority of the traits in both breeds were less than 0.025.

Discussion:

A thorough understanding of the heritability of a condition enables veterinarians to advise their clients on how likely it is that the trait will be passed from parent to offspring. In this study, we calculated h^2_{SNP} estimates for eight biochemical measurements reflective of EMS and ACTH in a population of 264 Welsh ponies and 286 Morgans by estimating additive genetic variance from high-density SNP genotype data (aka "SNP chip heritability"). This allowed us to overcome many of the biases in pedigree h2 estimates by including populations with a large proportion of unrelated individuals sampled across multiple farms throughout the United States and Canada.

Eight of nine traits had h^2_{SNP} estimates that were significant and moderately (0.21-0.40) to highly (>0.40) heritable in one or both breeds [338]. Leptin, INS-OST, TG, and ACTH had similar estimates across both breeds. However, glucose, GLU-OST, insulin, NEFA and adiponectin had differences in heritability of greater than 30% across breeds. Given that heritability is an estimate of the genetic variation in a trait, and is population specific, it is not surprising that we identified breed variation in h^2_{SNP} estimates for five of the nine traits. This can be explained by several factors. First, if more risk alleles contribute to a specific trait in one breed, it will have a higher heritability estimate. Second, if a trait or region of the genome is highly selected for, specific alleles may become fixed in the population (low to zero genetic variability in that region) leading to lower h^2_{SNP} estimate. Third, although a genetic variant may influence a trait across both breeds, if the variant is rare in one breed it will not contribute to the overall estimate of h^2_{SNP}.

We reported h^2_{SNP} estimates calculated in GCTA with the inclusion of the wGRM. Both GCTA and LDAK implement REML for estimating h^2_{SNP}, with differences between the algorithms based on assumptions of the effects of LD and MAF on h^2_{SNP} [405,413]. Without knowing the causal variants underlying these biochemical traits, it is impossible to know which method is most appropriate for these data. Therefore, we compared all three methods: GCTA with a standard GRM, GCTA with the wGRM, and LDAK with the wGRM (**Appendix A**) and found very little difference between estimates (**Appendix A:**

Supplementary Table A3), which gives us further confidence in the accuracy of our results.

Previous epidemiological studies have identified age, sex, and environmental management as risk factors for EMS [13,18,414], while season has been associated with variation in ACTH and glucose and insulin dynamics [61,67]. However, the suggested risk factors are not consistent across studies and do not correlate with all of the biochemical traits measured in this study, nor has a consensus on the effect of seasonal variation been established [48,66]. Some of these differences may be due to study design, populations sampled, or the represented geographic regions. We used AIC values to determine which covariates best fit a linear regression model for each biochemical trait in both breeds. Overall, season was included in h^2_{SNP} estimates for eight of the nine traits in both breeds, but inclusion of age and sex was more variable between breeds (**Table 2.1A**). The differences seen for age may reflect the tendency for Welsh ponies to express clinical disease at a younger age, and because of the range of ages in our cohort, where we had a larger number of younger Welsh ponies (46 ponies between 2 and 4 years old) than younger Morgans (7 horses between 2 and 4 years old). Notably, most h^2_{SNP} estimates based on model selection were similar to those that included age and sex, or age, sex and season (**Appendix A: Supplementary Table A4**). Estimates with the largest difference included leptin in the Welsh ponies and INS-OST in the Morgan; however, these estimates still overlapped within the range of the standard errors. Given that h^2_{SNP} estimates are population specific, we chose to report the estimates which best represented our data and tended to be the more conservative estimates for h^2_{SNP}.

Although several environmental risk factors have been identified for EMS, currently these factors only account for a portion of the environmental variation in the EMS phenotype [20]. We specifically chose farms in which at least one animal with clinical signs consistent with EMS and one normal control could be sampled, and where most horses on the same farm had the same management regardless of EMS status; thus, controlling for both known and unknown shared environmental factors (Supplemental Methods). Estimating heritability across varying environments (farms) avoids bias in the estimates caused by a single environment. When farm was included as a covariate in our analysis, the h^2_{SNP}

estimates were lower and often non-significant (**Appendix A: Supplementary Table A4**). This is likely due to non-independent sampling of horses from farms and excessive parsing of data (large number of farms with relatively few numbers of individuals). However, there still may be a small degree of inflation in our h^2_{SNP} estimates due to unaccounted for environmental variation.

Overestimation of h2 can occur due to unaccounted for population substructure, or genetically distinct subsets of individuals within a population. We tested for this in our cohort and found some evidence of population substructure (**Supplementary Methods** and **Supplementary Table A5**). To further explore this possibility, we subsetted the data by randomly removing 10% of the population, and repeated this process 100 times, to determine if a few individuals were artificially inflating the h^2_{SNP} estimates. The trait which had a larger difference (difference of 0.075) was ACTH in the Morgan. However, the mean h^2_{SNP} and SE for the remaining traits were similar to those obtained from the full dataset, indicating that population substructure or cryptic relatedness had minimal influence on the h^2_{SNP} estimates (**Table 2.2**).

Across both breeds, h^2_{SNP} for five of the 18 regression models (i.e., nine traits in each breed) did not meet a Holm-Bonferroni correction for statistical significance; however, h^2_{SNP} for 17 of the 18 regression models had uncorrected p-values of <0.05. Bonferroni corrections are conservative adjustments and can increase the probability of false negatives. Notably, all estimates were significant when including sex and age as a covariate (**Appendix A: Supplementary Table A4**). Therefore, it is likely that glucose (Morgans), TG (Morgans), GLU-OST (Welsh ponies) and ACTH (Welsh ponies) are within the appropriate range of h^2_{SNP} estimates for these breeds, and that increasing our population size would improve our power to estimate h^2_{SNP} for these traits. As we have previously shown, increasing the population size would also reduce the relatively large SE seen with all our estimates, but would not affect the overall h^2_{SNP} estimate [415]. The one trait with an unadjusted p-value >0.05 was ACTH in the Morgan, which was also the trait that had the largest difference between the original and mean subset value.

In humans, several published reports have estimated the h2 of traits associated with MetS, including insulin (0.09-0.51), glucose (0.10-0.33), GLU-OST (0.16-0.17), TG (0.11-0.60),

high-density lipoproteins (0.328-0.63), cholesterol (0.44-0.62), leptin (0.55), adiponectin (0.551), and proxy for insulin dysregulation (0.38-0.50) (**Appendix A: Supplementary Table A6**) [341,346,348,352,360,416,417]. Direct comparison between our estimates and studies in humans must be taken with caution due to differences in underlying physiology, populations, study design, methods for estimating h2, and measurement methods for the biochemical parameters. However, the large range in estimates for the same trait across ethnic groups may be analogous to the differences in our estimates between breeds. Further, the similarity in h2 estimates for several traits across species lends confidence to our estimates.

Most of our h^2_{SNP} estimates were <0.60, which is consistent with published heritability estimates of other complex traits in horses including: recurrent exertional rhabdomyolysis (0.34-0.49) [415], osteochondrosis (0.10-0.46) [418,419], and racing performance (0.19-0.61) [420]. Notably, three h^2_{SNP} estimates in our study were >0.60: insulin (0.80) in the Welsh ponies, and NEFA (0.68) and adiponectin (0.91) in the Morgans. These higher estimates are likely due to a small number of variants with a large effect size on each trait, which occurs commonly in animals due to selective breeding. For example, in horses four variants explain 83% of the genetic variation in height [395], a highly heritable trait with published h2 estimates up to 0.89 [396]. This is in contrast to humans where hundreds of genes with small effect control this trait [397]. However, another possibility could be inflation of these three estimates due to selection within subpopulations of our cohort.

Although several of the EMS traits likely have variants of large effect, in combination EMS is the result of dozens of variants of small, moderate, and large effect size. Given the complex, interrelated nature of the endocrine system, it is expected that several of these variants will also contribute to components outside of EMS. Therefore, attempting to eliminate EMS through breeding is not feasible, nor recommended, as it would have undesired consequences including decreasing genetic diversity within breeds. However, the availability of genetic tests would enable veterinarians to evaluate a patient's risk for developing EMS by assessing the number of genetic and environmental risk factors for each individual. Understanding of a horse's risk of EMS will ultimately allow veterinarians to know which horses need frequent monitoring and would benefit from early

environmental modification, as well as those that may not fully respond to environmental management alone (i.e. horses with a large number of risk variants or a modest number of variants with a moderate to large effect). In addition, this information should be used to guide owners in responsible breeding decisions (i.e., not breeding two horses who both have a high genetic risk for EMS).

In conclusion, through analysis of high-density SNP genotype data we determined that eight measured biochemical traits associated with EMS were moderately to highly heritable in both Morgan and Welsh ponies. Differences in h^2_{SNP} estimates in several traits between these two breeds is likely due to differences in risk alleles or the frequency of risk allele that are contributing to previously identified breed variability in metabolic traits. The results of this study provide the first concrete evidence of the genetic contribution to these eight phenotypes and that continued research for identification of the genetic risk factors for EMS is warranted.

A.

	Age	Sex	Season
Welsh Ponies			
Glucose			**X**
Glucose-OST			**X**
Insulin	**X**	**X**	
Insulin-OST	**X**	**X**	**X**
NEFA	**X**		**X**
TG	**X**	**X**	**X**
Leptin		**X**	**X**
Adiponectin		**X**	**X**
ACTH	**X**		**X**

	Age	Sex	Season
Morgan Horses			
Glucose		**X**	**X**
Glucose-OST		**X**	**X**
Insulin	**X**		**X**
Insulin-OST	**X**		**X**
NEFA		**X**	
TG		**X**	**X**
Leptin		**X**	**X**
Adiponectin	**X**		**X**
ACTH	**X**		**X**

B.

	Glucose	GLU-OST	Insulin	INS-OST	NEFA	TG	Leptin	ADIPON	ACTH
Welsh ponies									
h^2_{SNP} estimate	0.408	0.226	0.808	0.440	0.434	0.313	0.554	0.488	0.305
SE	0.135	0.142	0.108	0.148	0.136	0.146	0.132	0.143	0.154
P-Value	**<0.001**	0.05	**<0.001**	**<0.001**	**<0.001**	**0.02**	**<0.001**	**<0.001**	0.05
Morgan horses									
h^2_{SNP} estimate	0.208	0.565	0.592	0.359	0.684	0.273	0.486	0.913	0.408
SE	0.172	0.175	0.195	0.185	0.164	0.176	0.177	0.181	0.215
P-Value	0.05	**<0.001**	**<0.001**	**<0.05**	**<0.001**	0.05	**<0.001**	**<0.001**	0.06

Table 2.1: Covariate selection and heritability (h^2_{SNP}) estimates. Part A: Summary tables of the covariates selected for each trait based on model analysis for the Welsh ponies and Morgans. Part B: Heritability estimates using GCTA with the wGRM for nine biochemical traits with the selected covariates for both breeds. P-values are adjusted by a Holm-Bonferroni correction (**Appendix A: Supplementary Table A2**), bolded values were <0.05 after correction, bolded p-values listed as <0.05 are those which were less than 0.05 prior to rounding, p-values in red are those which the unadjusted p-value was >0.05. Abbreviations: SE: standard error, GLU-OST: glucose post oral sugar test, INS-OST: insulin post oral sugar test, NEFA: non-esterified fatty acids, TG: triglycerides, ADIPON: adiponectin.

	Glucose	GLU-OST	Insulin	INS-OST	NEFA	TG	Leptin	ADIPON	ACTH
Welsh ponies									
Mean h^2_{SNP}	0.397	0.200	0.817	0.419	0.468	0.302	0.579	0.500	0.281
Mean SE	0.147	0.150	0.115	0.16	0.155	0.157	0.148	0.153	0.164
Diff h^2_{SNP}	0.011	0.026	0.009	0.021	0.034	0.011	0.025	0.012	0.018
Morgan horses									
Mean h^2_{SNP}	0.199	0.612	0.630	0.385	0.748	0.294	0.507	0.911	0.333
Mean SE	0.192	0.193	0.217	0.208	2.10	0.197	0.196	0.204	0.236
Diff h^2_{SNP}	0.009	0.047	0.038	0.026	0.002	0.021	0.021	0.002	0.075

Table 2.2: Repeated subsetting of heritability (h^2_{SNP}) estimates. Summary table from random subsetting of the data including: mean heritability estimates (mean h^2_{SNP}), mean standard error (SE), and the difference between the mean h^2_{SNP} value and the estimate with the entire cohort (diff h^2_{SNP}), for nine biochemical traits in both Morgans and Welsh ponies using GCTA and the wGRM. Covariates included in the model were based on Akaike information criterion values. Abbreviations: GLU-OST: glucose post oral sugar test, INS-OST: insulin post oral sugar test, NEFA: non-esterified fatty acids, TG: triglycerides, ADIPON: adiponectin.

Chapter 3: Genome-wide association analyses of EMS phenotypes in Welsh ponies and Morgan horses.

Summary: Equine metabolic syndrome (EMS) is thought to be complex trait, yet the genetic risk alleles contributing to EMS have yet to be discovered. Successful genome-wide association analyses (GWA) identify regions of the genome harboring genetic risk alleles. The objectives of this chapter were to perform within breed GWA to identify significant contributing loci in two high risk breeds followed by a meta-analysis to identify shared and unique loci between both breeds. GWA was performed for eleven EMS traits and identified 130 associated genomic regions in a cohort of 264 Welsh ponies and 142 associated regions in a cohort of 286 Morgans. The boundaries of GWA regions were then defined based on either a fixed-size or on the breakdown of linkage disequilibrium (LD). Approximately 60% of the fixed-size boundaries were found to be larger than the LD boundaries. GWA meta-analysis demonstrated that 65 of the 272 GWA regions were shared across breeds. GWA regions were subsequently prioritized if they were: shared between breeds (high priority), shared across traits (high priority), identified in a single GWA cohort (medium priority), or shared across traits but no SNPs exceeded the threshold for genome-wide significance (low priority). Prioritization resulted in 56 high priority, 26 medium priority, and 7 low priority regions, for a total of 1,853 candidate genes in the Welsh ponies; and 39 high priority, 8 medium priority, and 9 low priority regions, for a total of 1,167 candidate genes in the Morgan horses. These data clearly support the hypothesis that EMS is a polygenic trait with breed-specific risk alleles as well as those shared across breeds.

Introduction

Equine metabolic syndrome (EMS) is best described as a clustering of metabolic derangements, which often lead to a pasture-associated or endocrinopathic laminitis. The term EMS was coined based on similarities with metabolic syndrome in humans (MetS) including: insulin dysregulation, derangements in fat metabolism, regional adiposity and alterations in inflammatory mediators. Both syndromes also share known risk factors including age, sex, diet, exercise, and seasonal variation [13,66,67,414,421-423].

Although a dominant mode of inheritance for laminitis status was proposed for a small group of ponies [13], breed differences in EMS susceptibility, metabolic profiles, and clinical severity have led to the more widely applicable, alternative hypothesis that EMS is a complex disease, with both environmental and genetic risk factors contributing to disease severity. As a complex trait, it is likely that EMS is the result of a combination of genes with variable modes of inheritance, penetrance and effect size [424]. Recently, our laboratory provided evidence for this hypothesis through estimation of narrow sense heritability in a cohort of Morgan horses and Welsh ponies, where eight of nine metabolic measurements were estimated to have low, moderate or high heritability (**chapter 2**). Further, several heritability estimates varied across the two breeds, which provided further evidence for breed related differences and was consistent with heritability estimates across ethnic groups for MetS.

Although heritability estimates provide valuable insight on the genetic contribution to a trait, they do not provide information on the (i) number of contributing genes, (ii) specific genes involved, or (iii) where these genes are located within the genome. Identification of the coding and non-coding variants contributing to a complex trait are important for understanding its complete pathophysiology and to gain a better understanding of how genes interact or are influenced by the environment. Further, this information is necessary for the development of genetic tests which would allow veterinarians to assess a patient's risk for developing EMS before they show clinical signs, identify horses that need frequent monitoring or early environmental modifications, and provide responsible breeding recommendations.

Genome wide association analyses (GWA) use single nucleotide polymorphism (SNP) genotype data to identify regions of the genome that are associated with a trait. GWA has been used and validated across multiple species for both simple and complex traits to narrow down the genome to specific regions of interest harboring the risk alleles and can provide valuable information about the genetic architecture of a trait. For example, GWA for MetS have led to identification of quantitative trait loci harboring candidate genes in several metabolic pathways, including alleles influencing lipoprotein particle size and glucose, insulin and lipid homeostasis [425,426]. Further, these studies have identified different risk alleles amongst ethnic groups [389].

We hypothesized that major genetic risk factors leading to EMS are shared across breeds, and that differences in the severity and secondary features of the EMS phenotype between breeds, or between individuals within a breed, are the result of modifying genetic risk alleles with variable frequencies between breeds. The first objective of this chapter was to perform within breed GWA to identify significant contributing loci in Welsh ponies and Morgan horses, two breeds known to be high risk for EMS. The second objective was to use meta-analysis to identify shared and unique loci between both breeds.

Materials and Methods

Samples: Horses used in this study were a part of a large, across breeds study evaluating the EMS phenotype [20]. From this dataset, 264 Welsh ponies (194 females and 70 males with a mean age of 11.7 years) and 287 Morgan horses (184 females and 102 males with a mean age of 12.3 years) were included in this analysis. Samples were collected from 31 and 28 farms throughout the United States and Canada for the Morgan horses and Welsh ponies, respectively.

Phenotype data collected on all horses included: signalment, medical history, laminitis status, environmental management (feed, supplements, turnout and exercise regimen), and morphometric measurements (body condition score (BCS), wither height, and neck and girth circumference). After an eight hour fast, an oral sugar test (OST) was performed using 0.15mg/kg Karo lite corn syrup as previously described [406]. Biochemical measurements at baseline included insulin, glucose, non-esterified fatty acids (NEFA),

triglycerides (TG), adiponectin, leptin and ACTH. Biochemical measurements 75 minutes after the OST included insulin (INS-OST) and glucose (GLU-OST).

For inclusion in the study, each farm had to have at least one control and one horse with clinical signs consistent with EMS (including horses with regional adiposity, hyperinsulinemia or an exaggerated response to the OST, elevations in TG, and decreased levels of adiponectin at the time of sampling) under the same management. Horses with a history or phenotypic appearance of pars pituitary intermedia dysfunction (PPID) were excluded from the study. Previously laminitic was defined as an individual who had been diagnosed with pasture-associated or endocrinopathic laminitis by a veterinarian, had radiographic evidence of laminitis, or had signs indicative of chronic laminitis observed by the researchers at the time of sampling. Horses in which laminitis could have been caused by another inciting factor (history of illness, grain-overload, corticosteroid administration or PPID), or who had clinically-evident, acute laminitis at the time of sampling, were excluded from the study.

Genotype Data: DNA was isolated from whole blood or hair roots using the Puregene Blood Core Kit, (Qiagen) per manufacturer's instructions. Genome-wide single nucleotide polymorphism (SNP) genotyping was performed on all horses. Horses were genotyped either on the Illumina EquineSNP50 BeadChip (268 Morgan horses), Axiom Equine MCEc670 array (220 Welsh ponies), or Axiom Equine MCEc2M array (44 Welsh ponies and 43 Morgan horses), containing 54,602 SNPs, 670,795 SNPs, and 2,011,826 SNPs across the equine genome including the 31 autosomes and X chromosome, respectively.

Haplotype phasing and genotype imputation of up to the two million SNPs present on the Axiom Equine MCEc2M array was performed on horses genotyped on the two lower density arrays using Beagle software [407]. Based on published recommendations, a cross breed population of 496 horses genotyped on the MCEc2M array, including the Welsh ponies and Morgan horses described above, were used as the reference population [408]. Imputation concordance was calculated by comparing data from individuals who were genotyped on both the low and high-density arrays. Briefly, for the 44 Welsh ponies and 43 Morgan horses genotyped on the MCEc2M array, genotypes from the MCEc2M array were masked down to those found on the Illumina EquineSNP50 BeadChip (Morgan

horses) or the MCEc670 array (Welsh ponies). Imputation was performed and the imputed genotypes were compared to the masked genotypes to determine genotype concordance. Concordance was 99.2% in the Morgan horses and 99.1% in the Welsh ponies. SNPs that did not have 100% concordance were removed from the data, yielding a total of 1,931,327 SNPs in the Welsh ponies and 1,932,766 SNPs in the Morgan horses.

Quality control on the imputed data was performed using the Plink software package [410]. All horses passed quality control, including evaluation for discordant sex information and SNP genotyping rate (>95%). Individual SNPs that had a genotyping rate of <90%, a minor allele frequency (MAF) of <1.0% or were outside of Hardy-Weinberg equilibrium (p-values <1.0e-05), were removed. After genotype pruning, 1,428,337 and 1,158,831 SNPs remained for subsequent analyses in the Welsh ponies and Morgan horses, respectively. Of these, a total 688,471 SNPs were shared between both breeds. Base pair locations for all SNPs were remapped from EquCab2 to EquCab3 using the script from Beeson et al [427].

Welsh Pony Population Structure: The Welsh Pony and Cob Society (http://wpcs.uk.com) registers Welsh ponies into six sections based on pedigree, withers height and conformation as follows: section A (sire and dam must both be section A, and the pony can be up to 50 inches for withers height), section B (either sire and dam are both section B or one parent can be a section A, and the pony can be up to 58 inches for withers height), section C (at least one parent must be C or D and the pony can be up to 54 inches for withers height), section D (at least one parent must be C or D and the pony must be over 54 inches for withers height), section H (either the sire or dam is a registered Welsh pony, and there are no height restrictions) and section P (either sire or dam is at least 50% Welsh pony with no height restrictions). Our cohort included 74 section As, 146 section Bs, 3 section Cs, 15 section Ds, 19 section Hs, 7 sections Ps, and 10 unknown/unregistered Welsh ponies (**Appendix B: Supplementary Table B1**). [428]. Principal components analysis (PCA) revealed population stratification in the Welsh pony cohort based on clustering of the registered sections (**Figure 3.1)**.To account for this population substructure, and avoid over-fitting the model, three separate GWA were performed using the full cohort (n=264), sections A, B, C and D (n=238) and sections A and B (n=220).

Genome-Wide Association Analyses (GWA): Eleven traits significantly associated with EMS including insulin, glucose, adiponectin, leptin, NEFA, TG, ACTH, insulin-OST, glucose-OST, and measures of obesity (neck circumference to whither height ratio (NH), and girth circumference to whither height ratio (GH)[20] were treated as quantitative response variables in the GWA analyses. Laminitis status was treated as a binary response variable. All quantitative traits were tested for normality using a normal probability plot and Shapiro test and adjusted for normality as appropriate. Adiponectin, leptin, and NEFA were square root adjusted and insulin, INS-OST and triglycerides and ACTH were log transformed. Glucose, GLU-OST and NH and GH ratios were normally distributed and did not need to be adjusted.

Traits measurements were adjusted to account for known confounding covariates using the residuals from a linear mixed effects model in the R software program Linear and Nonlinear Mixed Effects Models (nlme) with sex and age included as fixed effects and farm as a random effect [429]. For each trait, within breed GWA were performed from the imputed SNP genotype data using a custom code for an improved mixed linear regression analysis [20]. This algorithm utilizes a three-step process, which combines a Bayesian Sparse Linear Mixed Model (BSLMM) [430] available in the software program Genome-wide Efficient Mixed Model Association (GEMMA) [379] and a linear mixed model implemented in FaST-LMM [431]. In step one, the genome is divided by chromosome and SNPs are placed into 500kb bins. Based on results from BSLMM, the SNP with the highest model frequency and the two adjacent SNPs were chosen to represent the corresponding bin. In step two, a likelihood ratio test was performed to determine if inclusion of the top ranked bin as a random effect will improve the null model (model with sex and age as fixed effects and farm as a random effect). If the model was improved, the alternative model became the null model and the next highest-ranked bin was tested. If the model was not improved, the bin was discarded, and the next highest-ranked bin was tested against the null hypothesis. After all bins were evaluated, SNPs which improved the model were utilized to create the select SNP genetic relationship matrix (GRM). In step three, all imputed SNPs were tested for an effect on the trait using FastLMM with the inclusion of the select SNP GRM in place of the standard all-SNP GRM. For the GWA, the tested SNP,

and SNPs within 1Mb of the tested SNP, were removed from the select SNP GRM to avoid proximal contamination.

The number of iterations for the Markov chain Monte Carlo (MCMC) implemented in BSLMM has not been previously assessed [20,430], and our initial results with the default of 550 thousand (k) iterations with 10k burn-in iterations provided inconsistent results across seeds. Therefore, we took appropriate steps to determine the number of iterations for the MCMC to converge and provide consistent results across seeds. First, to assess the concordance of SNPs identified by BSLMM, we performed this step using 10 million (M) iterations with 100K burn-in iterations, which was repeated across ten different seeds. SNPs with a beta value greater than zero (i.e. the posterior mean for SNPs which were estimated to have a large effect on the outcome variable) were extracted from the dataset for each seed. For this subset of SNPs, the intersect between seeds was determined, and correlations between gamma values (proportion of iterations that the SNP was estimated to have a large effect) were calculated. For 10M iterations, the total number of SNPs with a beta value greater than zero ranged from 486,937 to 497,207 SNPs. Approximately 50% of the SNPs were shared between two seeds, ~13% were shared between four seeds, and ~3% were shared across all ten seeds. Pearson's correlation coefficient between gamma values were minimal at <0.01. Thus, this process was repeated using 20M iterations (200K burn-in iterations) and then increasing in 10M and 100k increments up to 100M iterations (1M burn-in iterations). Although SNP concordance improved as the number of iterations increased, the gain plateaued after 50M iterations where all SNPs had a beta value greater than zero. In addition, the Pearson's correlation coefficient for gamma values was still poor at 0.20 at 100M iteration. Computational time was extensive at 30 and 60 days to complete the 50M and 100M iterations, respectively, utilizing six processors per node and running seeds in parallel [432].

Previous studies have averaged the values of MCMC estimates across repeated chains [433]. For this analysis the goal was to maximize sensitivity; therefore, using data from the 10M iterations, the max gamma value across all ten seeds was chosen to represent each SNP in which beta was greater than zero. These values were then used to choose the most influential SNP per 500kb bins (step 1). To assess the repeatability of these results, this

process was repeated using 10 different seeds at 10M iterations and 20M iterations. Although differences were present, most hits were shared across all three results (**Appendix B: Supplementary Table B2**). Thus, to maximize computational efficiency and sensitivity, we used the max gamma value across 10 seeds obtained from 10M iterations (100K burn-in) and prioritized regions of interest (see *Prioritization of GWA Regions*). However, validation of this technique and calculation of the sensitivity and specificity will require identification of the genetic variants within these regions.

The threshold for genome wide significance was based on the effective number of independent tests (SNPs not in linkage disequilibrium [LD]) as calculated by the Genetic Type 1 Error Calculator (GEC) [434]. In the Welsh ponies, this value was 841,750 SNPs, resulting in a Bonferroni-corrected threshold for genome wide significance of 5.98×10^{-8}. For the Morgan horses, the effective number of independent tests was 657,030 SNPs, resulting in a Bonferroni corrected threshold for genome wide significance of 7.61×10^{-8}. The suggestive threshold for both breeds was set at 1.00e-05 [435,436].

Meta-Analysis: A GWA meta-analysis was performed with the software program METASOFT [437] using the Morgan horse and Welsh pony GWA summary data from the 688,471 SNPs that were shared between breeds. Briefly, the METASOFT algorithm uses a random effects model which adjusts for heterogeneity between studies by allowing the effect size of the alternative allele to vary between populations. Unlike other random effects models, where both the null and alternative models assume heterogeneity, METASOFT uses a likelihood ratio test that assumes heterogeneity only under the alternative model [437]. The effective number of shared SNPs was 306,023 in the Morgan horses and 307,349 in the Welsh ponies as calculated by GEC. For a more conservative p-value, the threshold for genome wide significance was determined using the effective number of SNPs for the Welsh ponies (0.05/307,349) and set at 1.63×10^{-7}. The suggestive threshold was set at 1.00×10^{-5} [435,436]. To be considered a region of interest identified on meta-analysis (MA-ROI), at least one SNP needed to exceed the threshold for genome-wide significance.

Prioritization of GWA Regions and Identification of Positional Candidate Genes: All GWA regions where SNPs exceeded the suggestive threshold for significance were

reviewed. To be considered within a single region, consecutive SNPs on the same chromosome had to be within 500kb of each other [438,439]. Regions of interest had to contain a minimum of five SNPs exceeding the suggestive threshold, with at least one SNP exceeding the threshold for genome wide significance.

Fixed-Size Regions: The boundaries of the fixed region were defined as 500kb 5' of the base pair position of the minimum SNP within the region and 500kb 3' of the base pair position of the maximum SNP [438-443]. A region was identified as shared if it was within the boundaries of another region and prioritized as described below.

LD-Regions: To define the boundaries of the LD region, the software program Plink was utilized to calculate the pairwise LD measures for all SNPs within the region [410]. Window size was set at 1Mb and pairwise calculation for LD with the test SNP was performed for all SNPs within the window. The threshold for SNPs within LD was set at greater than 0.3 [441]. A custom code was used to identify regions where LD for all SNPs dropped below 0.3 and spanned at least 100kb both 5' and 3' to the widest peak of LD within the window, which was used to define the boundaries of the ROI. If LD did not drop for at least 100kb on either side of the LD peak, window size was increased by 1Mb until the ROI could be defined. An LD-region was identified as shared if it was within the boundaries of another LD-region and prioritized as described below.

Prioritization*:* Regions were prioritized based on whether they were identified as shared between breeds on meta-analysis, an ROI, or shared across traits within a single GWA cohort (for example, a region shared between insulin and adiponectin in the Morgan horses). The prioritized regions were categorized as high, medium or low priority (**Figure 3.2**) as follows:

- High priority: Region was identified as an MA-ROI or it was shared across traits with at least one region being considered an ROI.
- Medium priority: Region was identified as an ROI in at least one GWA cohort.
- Low priority: Region was shared across traits, but no regions met the criteria to be considered an ROI.

- If a region met the criteria for more than one category (for example a region identified as a MA-ROI and was also shared across traits but not an ROI) then the region was assigned the higher priority level.

Identification of Positional Candidate Genes*:* Positional candidate genes were identified using the Bioconductor/ R software package biomaRt [444] with EquCab3 as the reference genome [445]. Boundaries were based on the fixed and LD-regions as described above. Positional candidate genes were defined as all protein coding genes, pseudogenes, and RNA genes within the region.

Results

GWA Results for Welsh Ponies: Principal components analysis (PCA) revealed population stratification in the Welsh pony cohort based on clustering of the registered sections (**Figure 3.1)**. GWA across all twelve traits for the entire Welsh pony cohort (n=264), resulted in 130 regions where at least one SNP exceeded the suggestive threshold. Of these regions, 33 were identified as ROI (minimum of five SNPs exceeding the suggestive threshold with at least one SNP exceeding the threshold for genome wide significance). Specifically, GWA identified 1 ROI for insulin post oral sugar test, 5 ROI for baseline insulin, 1 ROI for adiponectin, 2 ROI for leptin, 2 ROI for ACTH, 9 for NH, 8 ROI for GH, and 5 ROI for laminitis status. ROI were not identified for INS-OST, glucose, NEFA, or TG (**Table 3.1**).

GWA of the adjusted cohort including only section A, B, C and D Welsh ponies (n=238), resulted in a total of 139 regions where at least one SNP exceeded the suggestive threshold. Of these regions, 23 were identified as ROI as follows: 2 ROI for baseline insulin, 2 ROI for NEFA, 1 ROI for adiponectin, 1 ROI for leptin, 5 ROI for NH, 9 ROI for GH, and 3 ROI for laminitis status. ROI were not identified for INS- or GLU-OST, glucose, triglycerides, or ACTH levels (**Table 3.1**).

Additional GWA analysis that included only the section A and B Welsh ponies (n=220) resulted in a total of 82 regions where at least one SNP exceeded the suggestive threshold. Of these regions, 13 were identified as ROI as follows: 1 ROI for INS-OST, 1 ROI for GLU-OST, 2 ROI for baseline insulin, 3 ROI for glucose, 1 ROI for NEFA, 1 ROI for

triglycerides, 1 ROI for ACTH, 1 ROI for NH, and 2 ROI for laminitis status (**Table 3.1**). ROI were not identified for adiponectin, leptin, or GH (**Table 3.1**).

Across all 12 traits, 38 regions were shared with two of the Welsh pony GWA cohorts and 5 regions were shared with all three of them. Fifteen of the 43 shared regions contained at least one GWA where the region met the criteria to be considered an ROI (**Tables 3.2 and 3.3**). The 43 shared regions represented 18.91%, 26.57%, and 30.49% of the total regions identified in the full cohort, the section A, B, C and D Welsh ponies, and the section A and B Welsh ponies, respectively. Eight regions had an ROI identified in the full cohort (24.24% of the total ROI for this cohort), 6 regions had an ROI identified in the section A, B, C and D ponies (26.09% of the total ROI for this cohort), and 6 regions had an ROI identified in the section A and B ponies (46.15% of the total ROI identified in this cohort). For example, analysis of ACTH identified five shared regions. The region on equine chromosome (ECA) 5 was shared across all three cohorts but was only identified as an ROI in the GWA of the full cohort (**Table 3.3 and Figure 3.3**).

GWA Results for Morgan Horses: GWA across all twelve traits for the Morgan horses, identified 142 regions where at least one SNP exceeded the suggestive threshold. Of these regions, 37 ROI were identified including, 1 ROI for INS-OST, 3 ROI for GLU-OST, 1 ROI for baseline insulin, 2 ROI for glucose, 4 ROI for NEFA, 4 ROI for adiponectin, 3 ROI for leptin, 3 ROI for ACTH, 5 for NH, 4 for GH, and 7 for laminitis status. ROI were not identified for plasma triglyceride levels (**Table 3.1**).

Shared Regions Across Welsh Ponies and Morgan Horses: Identification of the shared regions between the Morgan horses and at least one Welsh pony cohort from the boundaries of the fixed region obtained from the GWA results identified 1 shared region for laminitis status (all ponies), 1 shared region for ACTH (Morgan horses with section A, B, C, and D ponies), and 1 shared region for insulin-OST (for Morgan horses with both the section A, B, C, and D and section A and B ponies; **Figure 3.4**). The boundaries defined by the LD-region, identified the above shared regions as well as an additional shared region for GH on ECA 22 between the Morgan horses and the full Welsh pony cohort.

Meta-analysis identified all four shared regions, as well as an additional 56 regions and 5 unique regions (regions not identified in either breed as significant on GWA), for a total of

65 shared regions of interest (MA-ROI). MA-ROI included 2 for INS-OST, 4 for GLU-OST, 3 for insulin, 2 for glucose, 4 for NEFA, 7 for adiponectin, 5 for leptin, 15 for NH, 8 for GH, and 12 for laminitis status. Unique regions were found for INS-OST (1 MA-ROI), GLU-OST (1 MA-ROI), and NH (3 MA-ROI). No MA-ROI were identified for plasma triglyceride levels (**Table 3.4**).

Of the 56 regions identified on meta-analysis that were only significant in one breed in the breed-specific GWA, 30 (22 ROI) were called in at least one Welsh pony cohort and 26 (20 ROI) were called in the Morgan horses. Twenty-eight of the MA-ROI contained less than 5 SNPs of which 11 were single SNP regions. Comparison of the results using a fixed effects model identified 32 of the 65 MA-ROI and the traditional random effect model identified 2 of the 65 MA-ROI (**Table 3.4**).

Prioritization of GWA Results and Identification of Positional Candidate Genes Based on Fixed-Size Regions in Welsh Ponies: For the full Welsh pony cohort, 78 of the 130 regions were eliminated from further prioritization, 35 were categorized as high priority, 12 were categorized as medium priority and 5 were categorized as low priority (**Tables 3.5 and 3.6**). For the section A, B, C and D Welsh ponies, 94 of the 139 regions were eliminated from further prioritization, 19 were categorized as high priority, 19 were categorized as medium priority and 8 were categorized as low priority (**Tables 3.5 and 3.7**). For the section A and B Welsh ponies, 57 of the 82 regions identified on GWA were eliminated from further prioritization, 9 were categorized as high priority, 10 were categorized as medium priority and 6 were categorized as low priority (**Tables 3.5 and 3.8**).

Combining the results from all three Welsh pony cohorts resulted in 114 regions and 1,898 positional candidate genes with 46 high priority regions containing 890 positional candidate genes, 34 medium priority regions containing 719 positional candidate genes, and 35 low priority regions containing 289 positional candidate genes. Accounting for the 19 shared regions resulted in 91 unique regions and 1,511 positional candidate genes (**Tables 3.5 and 3.9**).

Prioritization of GWA Results and Identification of Positional Candidate Genes Based on Fixed-size Regions in Morgan Horses: For the Morgan horses, 88 of the 142 regions were eliminated from further prioritization (**Table 3.10**). This resulted in 54 regions being

prioritized and 1,104 positional candidate genes with 38 high priority regions containing 801 positional candidate genes, 8 medium priority regions containing 139 positional candidate genes, and 8 low priority regions containing 164 positional candidate genes. Accounting for the 10 shared regions resulted in 44 unique regions and 963 positional candidate genes (**Tables 3.5, 3.10 and 3.11**).

Prioritization of GWA Results and Identification of Positional Candidate Genes Based on LD-defined Regions in Welsh Ponies: The boundaries for the regions identified by LD for the 130 regions identified on GWA for the full Welsh pony cohort, the 139 regions identified on GWA for the Section A, B, C and D Welsh ponies, and the 82 regions identified on GWA for the Section A and B Welsh ponies are presented in **Tables 3.12, 3.13, and 3.14,** respectively. Across Welsh pony cohorts, the LD boundaries identified 5 additional regions shared across traits (ECA1 for adiponectin and INS-OST, ECA5 for insulin and leptin, ECA6 for leptin and GH, ECA9 for INS-OST and NEFA, and ECA18 for insulin and GH) but did not identify six regions as shared across traits that were identified with the fixed boundaries (ECA4 for leptin and GH, ECA10 for NH and GH, ECA14 for leptin and laminitis status, ECA19 for ACTH and laminitis status, ECA 28 for insulin and INS-OST, and ECA28 for adiponectin and leptin). This resulted in 89 regions being prioritized with 56 high priority regions containing 1,567 positional candidate genes, 26 medium priority regions containing 620 positional candidate genes, and 7 low priority regions containing 30 positional candidate genes for a total of 2,217 positional candidate genes. Accounting for the 18 shared regions resulted in 16 unique regions and 1,853 positional candidate genes (**Table 3.15**).

Prioritization of GWA Results and Identification of Positional Candidate Genes Based on LD-defined Regions in Morgan Horses: Using the boundaries of the LD-ROI for the Morgan horse GWA results identified three additional regions shared across traits (ECA 21 for triglycerides and adiponectin, ECA 6 for adiponectin and INS-OST, and ECA 19 for NH and laminitis status) but did not identify two regions as shared across traits that were identified with the fixed boundaries (ECA 20 for adiponectin and insulin and ECA 24 for insulin and NEFA). This resulted in 39 high priority regions containing 1,142 positional candidate genes, 8 medium priority regions containing 155 positional candidate genes, and

9 low priority regions containing 176 positional candidate genes for a total of 1,473 positional candidate genes. Accounting for the 12 shared regions resulted in 1,167 positional candidate genes (**Tables 3.16 and 3.17**).

Discussion

In this study, we used high density SNP genotype data and GWA in two high risk breeds to identify hundreds of regions of the genome contributing to 11 EMS traits. Both fixed (500kb) and linkage disequilibrium-based approaches were used to identify the boundaries of genomic regions of interest and positional candidate genes within these regions. Within breed prioritization of the LD-defined regions resulted in 56 high priority, 26 medium priority, and 7 low priority regions, for a total of 1,853 candidate genes in the Welsh ponies; and 39 high priority, 8 medium priority, and 9 low priority regions, for a total of 1,167 candidate genes in the Morgan horses. Meta-analysis demonstrated that 65 of these regions were shared across breeds. These data support the hypothesis that EMS is a polygenic trait with both across breed and breed-specific genetic variants.

Age and sex were included in our model as fixed effects based on epidemiological studies which identified both as risk factors for EMS [18,70]. Season [67,421], diet [13,414], exercise [132,414], and endocrine-disrupting chemicals [74] have also been identified as environmental risk factors for EMS, but a large percentage of environmental variation has yet to be explained [20]. Further, several studies have produced conflicting findings as to the effect of season on EMS biochemical measurements [51,66], as well as the long-term effect of high non-structural carbohydrate diets on insulin sensitivity [47,48]. Therefore, we included farm as a random variable to account for both known and unknown environmental risk factors, as well as non-independent sampling of our data (each farm was required to have one control and one horse with EMS to be included in the study).

Selective breeding for traits such as conformation or athletic performance can lead to population stratification within breeds [446], and not accounting for this population stratification can lead to spurious associations on GWA [428]. For this data, principal components analysis revealed population stratification in the Welsh pony cohort based on clustering of the registered sections (**Figure 3.1)**. This was not unexpected as the Welsh pony sections are distinct subpopulations based on pedigree and conformation.

Mixed linear models are a common way to account for population stratification and relatedness in GWA [378,379]. However, the Welsh ponies presented a unique challenge since, although the GRM would account for genetic similarities between Welsh pony sections, conformational traits such as height are considered complex traits and therefore the GRM would not account for all the phenotypic variation between sections. On the other hand, including both the GRM and section as a covariate would lead to over-fitting of the model by accounting for relatedness both as a random effect (GRM) and fixed effect (section). Further, accounting for population stratification by limiting the GWA to specific sections of Welsh ponies would reduce power to identify low frequency variants and prevent the identification of variants that are fixed within a section. For example, GWA for the full cohort identified an ROI on ECA6 for baseline insulin, which was not identified on the GWA for the section A and B ponies (**Table 3.3**). The allele frequency for a missense mutation in exon 1 of the high mobility AT hook gene (*HMGA2*) was found to be fixed in the section A ponies, 74% in the section B ponies, 3% in the section D ponies, and 64% in the section H and P ponies (see **chapter 5**). Thus, the frequency for the minor allele in the section A and B ponies (n=220) was 15%; whereas, the minor allele frequency for the full cohort (n=264), was 22% which was high enough to be detected on that GWA.

Therefore, to account for population stratification within our Welsh pony cohort while maximizing sensitivity to identify genetic variations contributing to EMS both within and across sections, we chose to perform the GWA using the full data set and then subset the data to the section A, B, C and D ponies, and the section A and B ponies. Ideally, we would have also included the section C and D ponies as a separate GWA cohort but were under-powered due to the low number of ponies from these sections represented in our population. Interestingly, less than a third of the GWA regions in the Welsh ponies were shared across cohorts (**Table 3.3**), which provides support for our approach. However, it should be acknowledged that this could also reflect spurious associations.

Comparison of the regions identified by the fixed and LD based approaches found a total of four shared regions between breeds. This could indicate that breed differences account for more of the risk alleles for EMS than previously thought, or that additional regions were shared but not identified in one breed on GWA, which can occur for several reasons.

First, if the allele frequency of the variant is low in one breed, then it will not be detected on that GWA. Second, the effect of the variant on a trait can vary between breeds. The within breed population sizes were powered to detect variants of moderate to high effect but would not find variants of low effect [447,448]. Third, variations in recombination of the ancestral chromosome can lead to differences in marker alleles between populations [449]. Depending on the markers represented on the genotyping array, the variant may be identified in one breed but not the other. Increasing the power of the study by performing across-breed GWA could identify more shared regions between breeds. However, combining data can lead to the inclusion of additional population substructure and unknown confounding variables into the model [450]. Further, subpopulations within the Welsh pony cohort prevented the feasibility of an across-breeds GWA for our data.

Meta-analysis uses GWA summary statistics to effectively combine GWA studies, increasing the number of individuals within the study and improving the power to find unique associations, variants of low effect, and additional shared regions across populations [451,452]. Both fixed and random effects models have been used for GWA meta-analyses. Fixed effects models assume the true effect of each risk allele is the same across populations; whereas, the random effects model assumes the effect size of the risk allele will vary across populations, explicitly modeling the between-study heterogeneity often encountered in these studies [452,453]. Surprisingly, random effects models were shown to be less powerful than those of the fixed effect models [437,454]. Han *et. al.* proposed that this was due to the assumption of heterogeneity under both the null and alternative hypothesis using the traditional random effects model [437]. The authors argued that under the null hypothesis the variant would have no effect in either population and thus would be a violation of this assumption; they proposed using a likelihood ratio test that assumes heterogeneity only under the alternative model which was implemented in the software program METSOFT [437].

We chose to perform the meta-analysis using this algorithm and identified 65 shared regions, of which 5 were unique (not identified in either breed specific GWA). The fixed effects model identified 32 of these regions and the traditional random effect model identified 2 of these regions, which is consistent what with Han *et. al.* found for

METASOFT [437]. Neither the fixed or traditional random effects models identified unique regions for meta-analysis. However, given the limited number of published studies using this algorithm, it is possible that the additional shared regions identified with METASOFT represent false positives. Identification of the causal variants in these regions for both breeds will enable studies to validate these results. Nonetheless, these data show the power of meta-analysis to identify additional and unique shared regions across breeds.

For this analysis, we defined an ROI as a GWA region in which a minimum of five SNPs exceeded the suggestive threshold and at least one SNP exceeded the threshold for genome wide significance. ROI accounted for 14-26% of the regions identified on GWA across all cohorts. Of the remaining regions, 25-49% were single SNP regions, 23-41% were regions with less than five SNPs, and 7-16% were regions with greater than five SNPs but none that exceeded the threshold for genome-wide significance.

To reduce false positives, regions were prioritized and those not assigned a priority were removed. Regions shared across breeds (MA-ROI) were given high priority, as these regions were not breed specific and likely to be found in other high-risk breeds. Regions shared across traits with at least one ROI were also assigned high priority. Many components and downstream effects of the endocrine system are highly interrelated; therefore, a variant affecting multiple traits would be expected to have a larger biological effect then a variant affecting a single trait. An ROI identified in one GWA cohort was assigned medium priority as these regions were likely breed or section (Welsh pony) specific and, based on the power of our study, variants of moderate to high effect. Finally, regions that were not ROI but shared across traits were assigned low priority. Because these regions were identified across multiple GWA it is possible that these regions are less likely to be false positives and/or that these regions contain variants of low effect.

Our prioritization removed 61% of the 130 regions for the full Welsh pony cohort, 66% of the 139 regions in the section A, B, C and D Welsh ponies, 63% of the 82 regions in the section A and B Welsh ponies, and 58% of the 142 regions in the Morgan horses. Of the 310 removed regions, 152 (49%) were single SNP regions, 118 (38%) were regions with less than five SNPs, and 40 (13%) were regions with greater than or equals to five SNPs but no SNPs which exceeded the threshold for genome wide significance. Given (i) the

large percentage of single or low SNP regions that were removed, (ii) the high-density genotype data used in these analyses, and (iii) the use of the max gamma value for BSLMM (improving sensitivity at the cost of specificity), it is likely most of these regions were false positives. However, we utilized Bonferroni corrected p-values which tend to be more conservative corrections [455]; therefore, some of the removed regions may harbor genetic variants associated with EMS but represent variants with very low effect or poorly annotated regions of the genome (relative decreased number of SNPs is that region). Increasing the number of individuals, or represented Welsh pony sections, would improve the power of the study to determine which of these regions were true or false positives.

Markers present on the genotyping arrays are common variants within the population and are used on GWA to tag the causal variant if they are in LD [456]. In other words, identification of causal variants and positional candidate genes is directly related to the region of the genome in LD with the SNP markers identified as significantly associated with the trait on GWA. In order to identify candidate genes, we first used a fixed boundary of 500kb 5' of the SNP identified on GWA with the lowest base pair position and 3' of the SNP with the highest base pair position. 500kb was chosen based on the average distance for LD to breakdown to ≤ 0.25 in Thoroughbred horses [440,441,443]. Although LD decay varies between horse breeds [442], using the more conservative Thoroughbred estimate gave a higher likelihood that we would capture all variants within LD ($r^2>0.3$) of the marker SNPs in our cohorts. From the fixed boundaries, 1,511 and 963 positional candidate genes were identified in the Welsh ponies and Morgan horses, respectively.

Estimates of LD decay are based on the average r^2 across chromosomal segments and do not represent specific regions of the genome [440,442]. Newer variants or variants within regions under selection will have longer LD blocks whereas older/ancestral variants will have shorter LD blocks due to longer periods of recombination. Therefore, using a fixed region has the potential to exclude causal variants or to include candidate genes that are not in LD with the marker SNPs. To more precisely call positional candidate genes for GWA regions, we calculated LD using the squared correlation coefficients between SNPs. SNPs within LD were defined as an $r^2 >0.3$ [441]. Boundaries were identified based on gaps of LD, i.e. were all SNPs dropped below 0.3 for a span of 100kb 5' (defined the start

of the LD block) and 3' (defined the end of LD block) to the widest peak of LD. Initially, a 50kb distance for the LD gap was used based on breed specific LD decay of an $r^2 < 0.3$ [442]. However, this distance was intractable as large peaks of LD were often identified on either side of the LD gaps. The distance was increased to 100kb and consistently identified where regional LD of marker SNPs declined below threshold, with ~70% LD gaps being >100kb.

Across all Welsh pony cohorts, 70% of the boundaries identified by LD were smaller than those identified by the fixed region, with an average difference of 645.4kb (range of 11.4kb to 1.7Mb); whereas, in the Morgan horses, 57% of the LD boundaries were smaller than that of the fixed regions, with an average difference of 566.6 kb (range of 51.5kb to 2.2Mb). The large percentage of fixed boundaries likely overestimating the region size is not surprising given that the fixed regions were based on data from Thoroughbreds, which have one of the highest inbreeding coefficients and LD amongst horse breeds [438,442]. Ponies and Morgan horses were identified to have LD similar to Quarter Horses [442], a breed with a high level of genetic diversity. For the remaining regions, the LD boundaries were an average of 1.9Mb longer (range of 22.8kb to 9.3Mb) in the Welsh ponies and 1.4Mb longer (range of 12.6kb to 8.2Mb) in the Morgan horses then defined by the standard region and likely represent regions under selection.

Further, assessment of LD provided additional information about the regional genetic architecture. For example, ECA4 in the Morgan horses for GH had several SNPs that had a second peak of LD after the central peak (**Figure 3.5**). Assessment of the reference genome identified an inversion corresponding to the location of the second peak. In addition, LD identified regions where SNPs within 500kb of another SNP were not in LD, indicating two separate regions. Conversely, LD also identified regions where SNPs which were identified as separate regions were found to be within LD of each other, indicating a single region.

In conclusion, the results of these data provide strong evidence that EMS is a complex, polygenic syndrome with dozens of risk alleles contributing to the phenotype. Prioritization of the hundreds of regions identified on the GWA of 12 individual traits let to the identification of thousands of positional candidate genes. Further work to narrow

down the candidate gene pool could include: (i) identification of biological candidate genes based on known gene function and evidence from other species; (ii) use of whole genome sequencing and linear regression analysis to fine map regions; (iii) use of intermediate phenotypes such as metabolomics or transcriptomics to identify shared regions; or (iv) network analysis for identification of genes within similar, relevant pathways. Nonetheless, this data was an important first step in the identification of the genetic risk alleles associated with EMS.

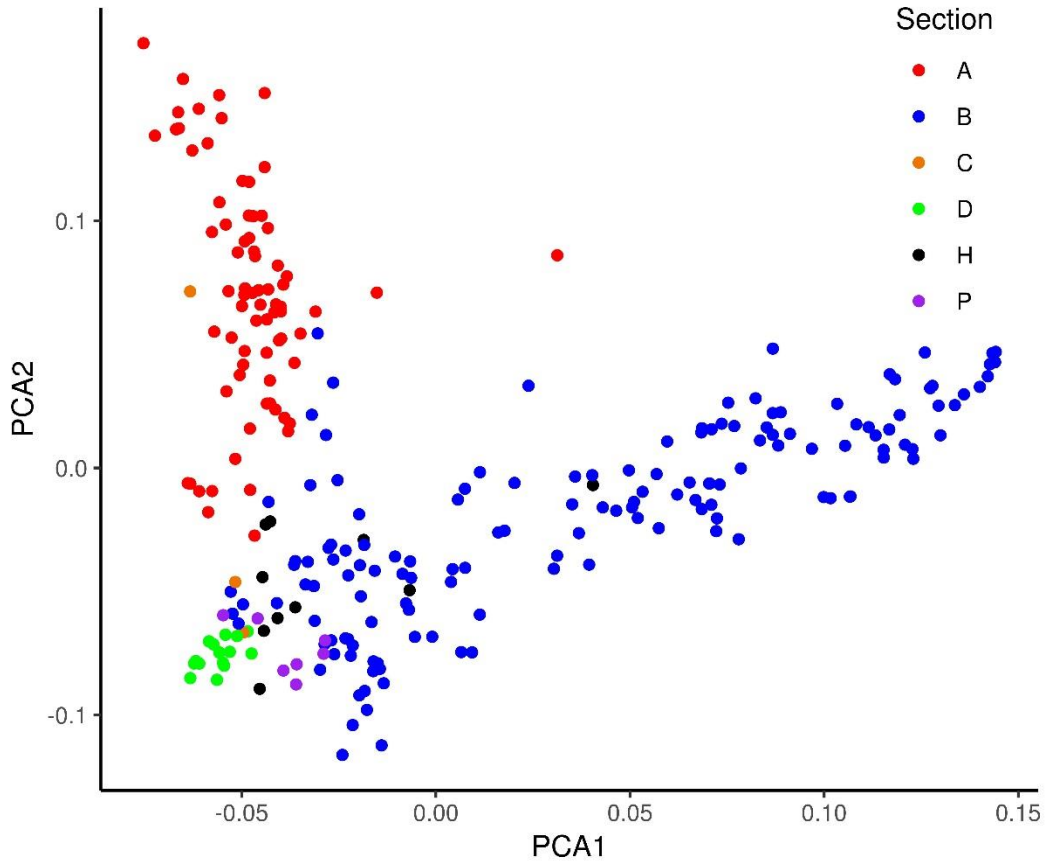


Figure 3.1: Principle components analysis for the Welsh ponies. The first two principal components are plotted on the x-axis (PCA1) and y-axis (PCA2). Distinct clustering is evident across the Welsh pony sections A, B, and D, indicating population substructure. The two section C ponies cluster with the section A (both ponies had a full section A parent) and the section H and P ponies are intermixed with the section B and D ponies (see appendix 1 for pedigree information).

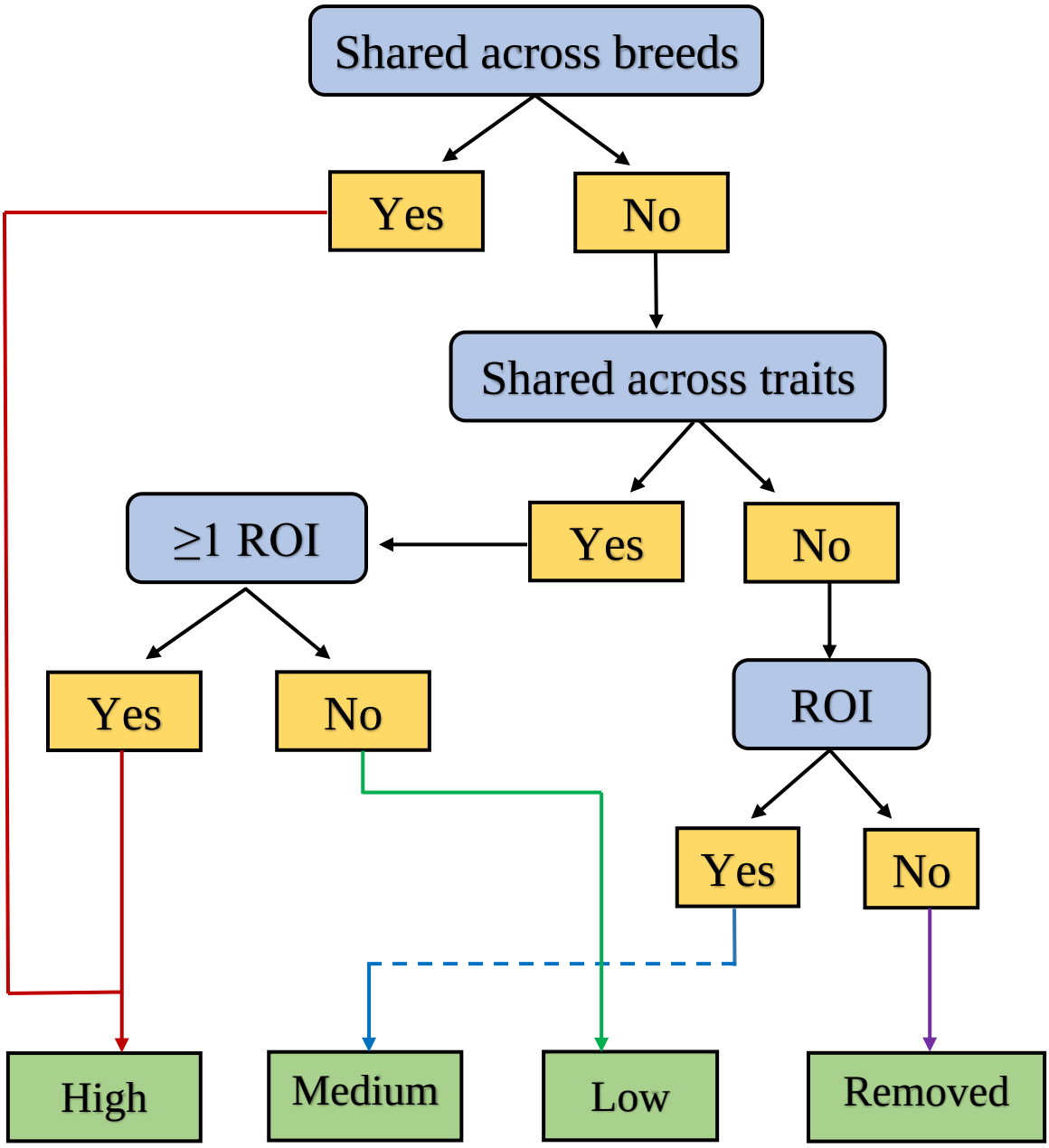


Figure 3.2: **Flow chart of the prioritization of the regions identified on genome-wide association analyses (GWA).** Regions were prioritized as high priority if they were identified as shared across breeds (MA-ROI) or it was shared across traits with at least one region being considered a region of interest (ROI). Regions were prioritized as medium priority if they were identified as an ROI in at least one GWA cohort. Regions were identified as low priority if they were identified as shared across traits, but no regions met the criteria to be considered an ROI. Regions which were not shared and were not considered an ROI were removed from further analysis.

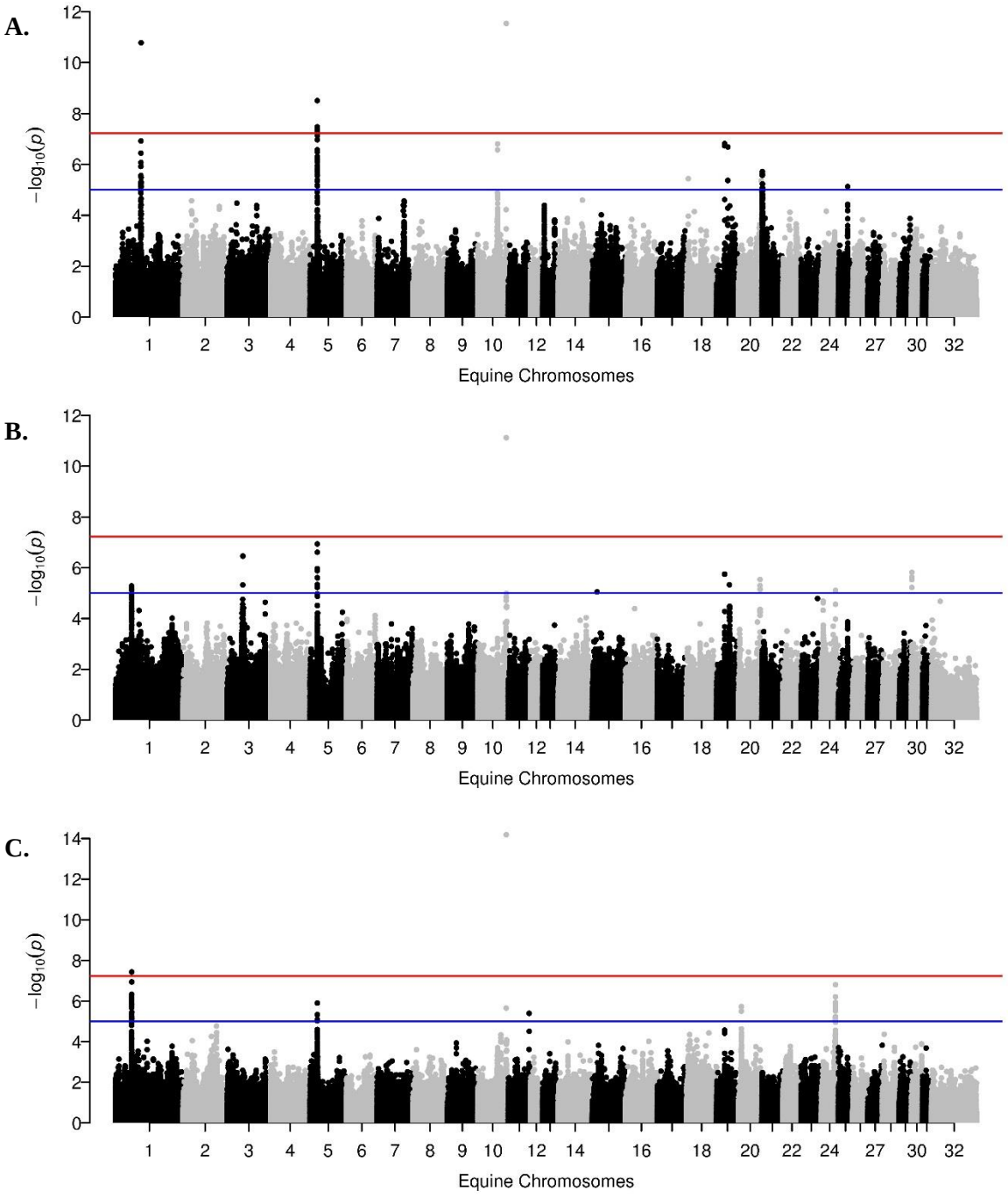


Figure 3.3: Manhattan plots of the genome wide association results for ACTH in (A) full Welsh pony cohort, (B) the section A, B, C and D Welsh ponies, and (c) the section A and B Welsh ponies. The equine chromosomes (ECA) are plotted on the x-axis and the -log of the p-value is plotted on the y-axis. The blue line indicates the suggestive threshold (1.0e-05) and the red line represents the genome-wide significant threshold (5.9e-08). In all three GWA, the same region on ECA5 exceeds the suggestive threshold but is only identified as an ROI in the full cohort (A).

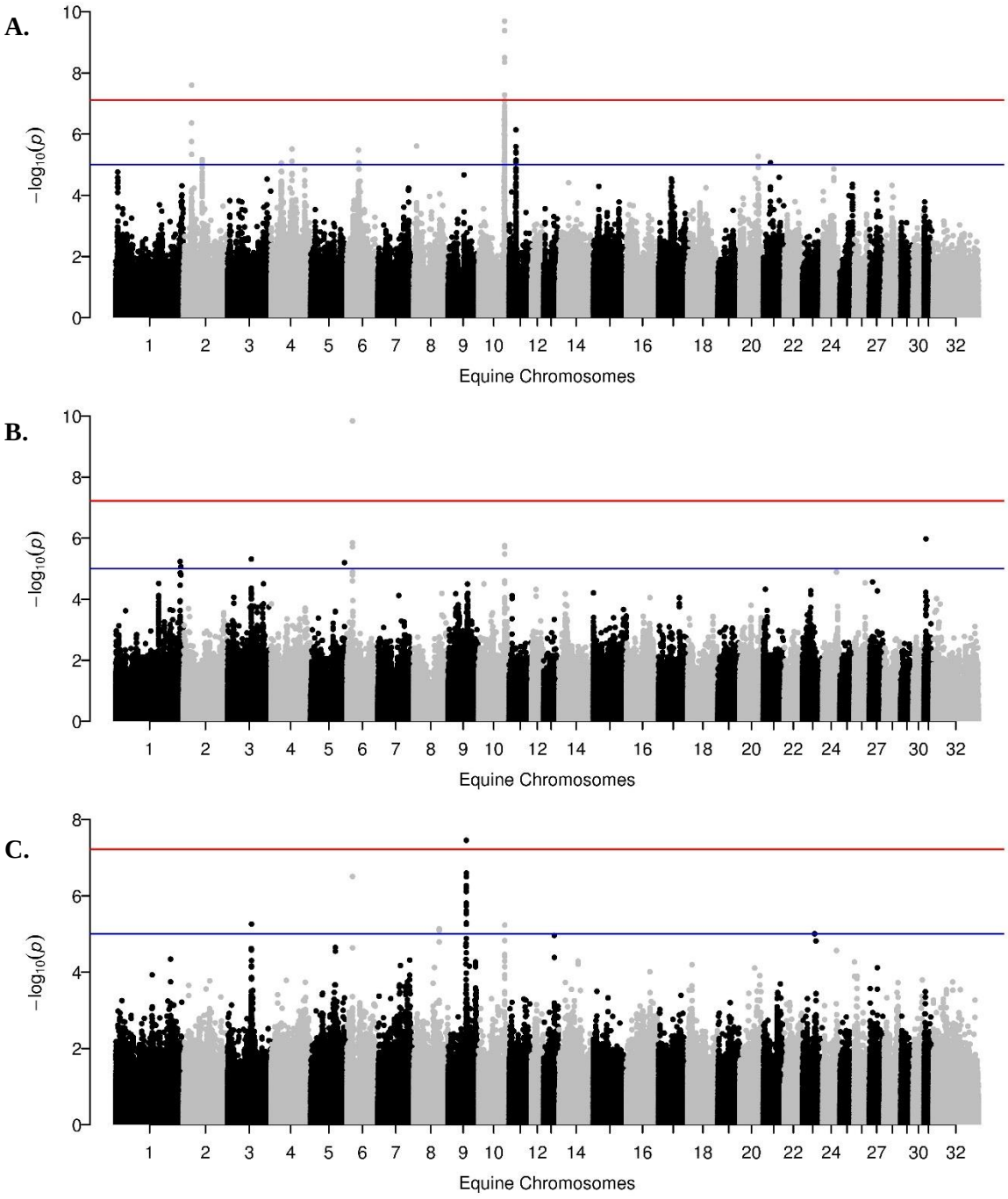


Figure 3.4: Manhattan plots of the genome wide association results for insulin concentration post oral sugar test in (A) Morgan horses, (B) the section A, B, C and D Welsh ponies, and (c) the section A and B Welsh ponies. The equine chromosomes (ECA) are plotted on the x-axis and the -log of the p-value is plotted on the y-axis. The blue line indicates the suggestive threshold (1.0e-05) and the red line represents the genome-wide significant threshold (7.61e-08 in the Morgans and 5.98e-08 the Welsh ponies). In all three GWA, the same region on ECA10 exceeds the suggestive threshold but is only identified as an ROI in the Morgan horses (A). However, GWA meta-analysis identified this region as shared across both breeds.

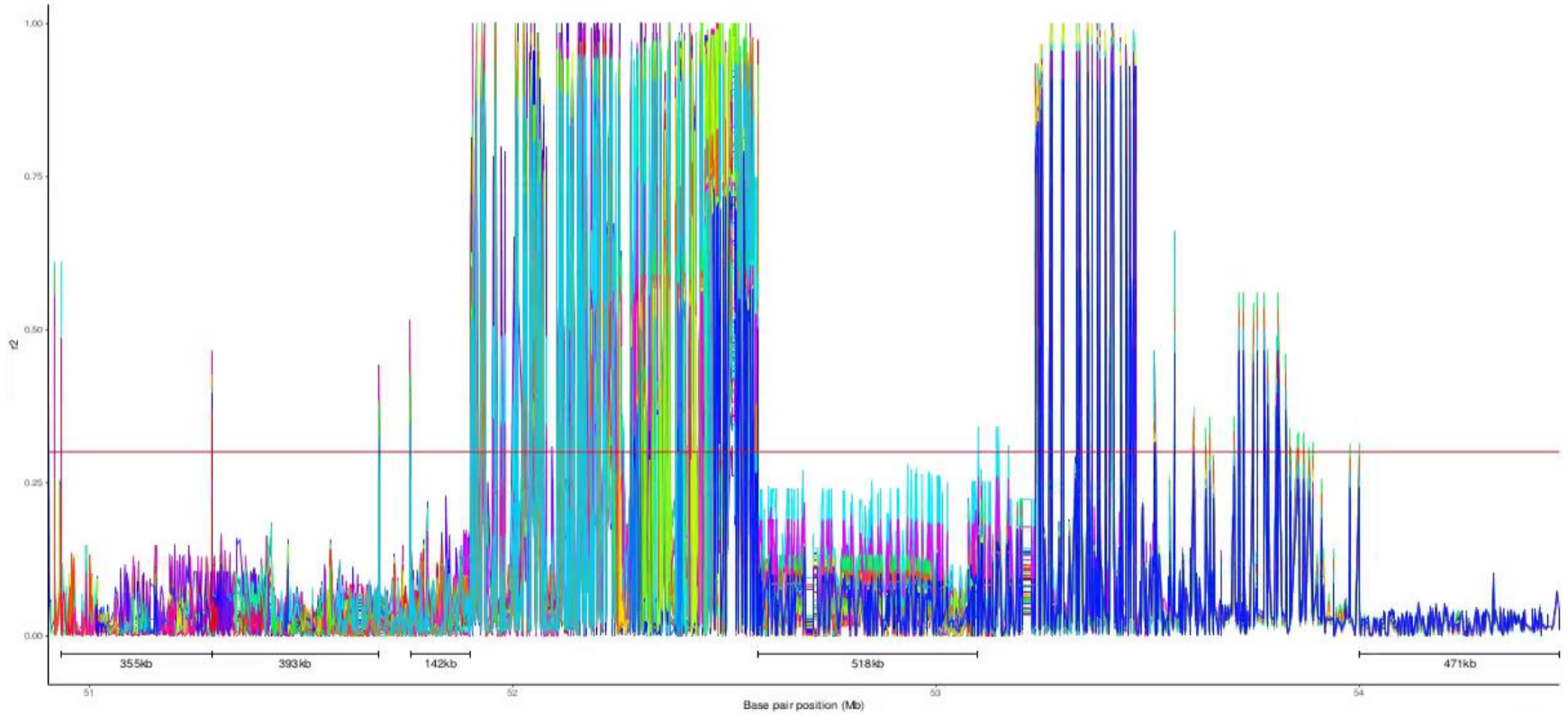


Figure 3.5: Linkage disequilibrium (LD) for neck-to-height-ratio (NH) on equine chromosome 4 (ECA4) in the Morgan horses. Base pair (bp) positions are on the x-axis and values for the pairwise comparisons of LD (r2) are on the y-axis. Red horizontal line indicates the threshold for LD at an r2 of 0.3. Individual colors represent the LD for each SNP identified on genome wide association analysis. The length in bp of the regions where the LD dropped below 0.3 for a minimum of 100kg for all SNPs are labeled parallel to the x-axis. Two peaks were identified with the first between bp 51900767-52580849 and the second peak from bp 53099275 to 54002853. Evaluation of the reference region identified an inversion at the position of the second peak.

	All Welsh ponies		Section ABCD		Section AB		Morgan Horses	
Trait	**Regions**	**ROI**	**Regions**	**ROI**	**Regions**	**ROI**	**Regions**	**ROI**
Insulin	15	5	10	2	10	2	12	1
Insulin-OST	7	0	7	0	6	1	10	1
Glucose	8	0	7	0	8	3	6	2
Glucose-OST	5	1	4	0	3	1	6	3
NEFA	10	0	10	2	11	1	13	4
TG	3	0	5	0	6	1	4	0
Adiponectin	2	1	6	1	4	0	17	4
Leptin	7	2	16	1	3	0	8	3
ACTH	10	2	10	0	6	1	18	3
NH	22	9	22	5	6	1	16	5
GH	23	8	30	9	9	0	14	4
LAM	18	5	12	3	10	2	18	7
Total	**130**	**33**	**139**	**23**	**82**	**13**	**142**	**37**

Table 3.1: Summed regions for each of the 12 EMS traits from the Welsh pony and Morgan horse genome-wide association analyses (GWA). Data includes the full Welsh pony cohort (n=264), individuals identified by pedigree as section A, B, C or D (n=238), individuals identified by pedigree as section A or B (n=220), and the Morgan horses (n=286). The column listed as ROI (region of interest) indicates that total number of regions which met the criteria to be considered an ROI (minimum of five SNPs with at least one SNP exceeding the threshold for genome wide significance). Abbreviations: OST: oral sugar test, NH: neck-to-height ratio, GH: girth-to-height ratio, LAM: laminitis status.

Trait	All WP & Section ABCD	All WP & Section AB	Section ABCD & Section AB	All Three	At least one ROI
Insulin	1	1	1	0	3
Insulin-OST	0	1	3	0	0
Glucose	1	1	0	0	1
Glucose-OST	1	0	2	0	1
NEFA	1	0	5	1	2
TG	1	0	1	0	0
Adiponectin	0	0	1	0	0
Leptin	1	0	0	1	1
ACTH	1	0	2	2	2
NH	4	0	0	1	2
GH	5	0	0	0	2
LAM	2	2	0	0	1
Total	**18**	**5**	**15**	**5**	**15**

Table 3.2: Summary table of the shared regions across two or three cohorts for each of the 12 EMS traits from the Welsh pony (WP) genome-wide association analyses (GWA) including the full cohort (n=264), individuals identified by pedigree as section A, B, C or D (n=238), individuals identified by pedigree as section A or B (n=220). The column listed as at least one ROI (region of interest) indicates that total number of shared regions where at least one region met the criteria to be considered an ROI (minimum of five SNPs exceeding the suggestive threshold with at least one SNP exceeding the threshold for genome wide significance).

Table 3.3: Specific shared regions from the Welsh pony genome-wide association analyses								
			All WP		Section ABCD WP		Section AB WP	
Trait	**Chr**	**Total_GWAS**	**Min_SNP**	**Max_SNP**	**Min_SNP**	**Max_SNP**	**Min_SNP**	**Max_SNP**
Insulin	6	2	**81074650**	**81566120**	**81421330**	**82660343**	-	-
	9	2	-	-	**58976739**	**59099678**	**58477773**	**60003081**
	15	2	**5887873**	**6278651**	-	-	5899834	NA
Insulin-OST	3	2	-	-	65320573	NA	65980441	NA
	6	2	-	-	15393073	15402993	15393073	NA
	8	2	69942950	69982846	-	-	69942980	69982846
	10	2	-	-	72158447	72240841	72238960	NA
Glucose	8	2	81424426	81518794	81284977	81428684	-	-
	29	2	21472582	21475253	-	-	**21440455**	**22135257**
Glucose-OST	5	2	-	-	66212381	66719700	66618266	66719700
	23	2	10907371	10951165	10942382	10951165	-	-
	28	2	-	-	33915296	NA	**33387547**	**33915296**
NEFA	6	2	76161874	NA	76161874	NA	-	-
	7	2	-	-	7268673	7382898	7268673	NA
	7	2	-	-	8181330	8243021	7268673	NA
	9	2	-	-	**47030376**	**48595497**	**47219472**	**48722431**
	14	2	-	-	33871722	NA	33829080	33974280
	22	3	18575108	NA	18575108	NA	18575108	18594384
	28	2	-	-	**33731242**	**34441427**	33819949	33831231
TG	12	2	32054230	32083040	32072315	32083040	-	-
	20	2	-	-	55609506	55705820	55239314	NA
Adiponectin	22	2	-	-	36975989	37058774	37058774	NA
Leptin	7	2	**65773875**	**65782930**	65773875	65782930	-	-
	10	3	865540	883471	856640	883471	871456	NA
ACTH	1	2	-	-	44050526	44285580	**43943376**	**44773532**
	5	3	**19628265**	**20107907**	19859591	20010745	19859591	20010745

Table 3.3: Specific shared regions from the Welsh pony genome-wide association analyses (cont.)								
			All WP		Section ABCD WP		Section AB WP	
Trait	**Chr**	**Total_GWAS**	**Min_SNP**	**Max_SNP**	**Min_SNP**	**Max_SNP**	**Min_SNP**	**Max_SNP**
ACTH (cont)	10	3	78846710	NA	78845710	NA	78703637	78846710
	19	2	21867680	21871015	21867680	21871015	-	-
	24	2	-	-	39069140	NA	38145005	39231598
NH	4	3	67875816	68337160	**68879163**	**69478180**	**67379332**	**69246252**
	4	2	77152103	NA	76199121	77653150	-	-
	4	2	78075875	78460889	76199121	77653150	-	-
	8	2	61139637	61236848	61139637	61177365	-	-
	21	2	**20193411**	**21497651**	20193411	21059497	-	-
GH	1	2	119770589	NA	119519666	119549672	-	-
	4	2	68337160	NA	69000484	69423480	-	-
	4	2	**84181768**	**85381459**	**83940435**	**85259515**	-	-
	20	2	**29233068**	**29537740**	29233068	29252036	-	-
	21	2	21387986	21398724	20919577	20922494	-	-
LAM	1	2	49077969	NA	49077969	NA	-	-
	2	2	**35906741**	**36414648**	36104151	36414648	-	-
	15	2	49986709	50013578	-	-	49986709	50013578
	16	2	65111190	NA	-	-	64888181	64938437

Table 3.3: Specific shared regions from the Welsh pony (WP) genome-wide association analyses (GWA) including the full cohort (n=264), individuals identified by pedigree as section A, B, C or D (n=238) and individuals identified by pedigree as section A or B (n=220). Bolded values are regions which met the criteria for being considered a region of interest (ROI). Values indicated as NA for the maximum SNP are those in which a single SNP exceeded suggestive or genome-wide significant threshold; values indicated as a dash (-) for both minimum and maximum SNP are those in which GWA did not identify the region as shared in that cohort. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.4: Meta-analysis results for 11 metabolic traits								
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**FE**	**Trad_RE**
Insulin	5	**A**	44104129	45081679	39	4	-	-
	15	**A**	5887873	6225014	21	2	-	-
	24	**A**	28804043	29076914	3	2	-	-
Insulin-OST	10	**B C E**	71620835	72425049	19	4	X	-
	28		38307699	38344594	2	1	-	-
Glucose	4	**B F**	18053357	18550035	20	1	-	-
	8	**D**	9289661	9312611	2	1	-	-
Glucose-OST	3	**B**	55982921	56558742	57	39	X	-
	4	**B**	27802674	28514796	18	4	X	-
	15		79697363	79717603	3	3	-	-
	28	**A**	34861664	34868420	2	2	-	-
NEFA	1	**B**	183532379	184178932	21	15	X	-
	17	**B**	13355958	14014858	23	1	X	-
	24	**D**	20975408	NA	1	1	-	-
	30	**B**	20148173	20205201	10	3	X	-
Adiponectin	2	**B**	16725632	17531903	25	19	X	-
	4	**B**	37105938	37523046	6	2	X	-
	6	**B**	31582345	31708194	17	1	X	-
	6	**B**	67097628	68036518	16	1	-	-
	18	**D**	41399862	41533081	9	1	-	-
	18	**A**	60138400	60241267	10	2	-	-
	20	**D**	3447045	3609674	10	4	X	X
Leptin	7	**A**	65731012	65804974	6	3	X	-
	10	**C**	871456	NA	1	1	-	-
	19	**B**	48839140	49627683	44	22	X	-
	24	**B**	28551544	28744981	17	6	-	-
ACTH	1	**A**	69730886	70257187	4	1	-	-

Table 3.4: Meta-analysis results for 11 metabolic traits (cont.)								
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**FE**	**Trad_RE**
ACTH (cont.)	1	**B**	82755708	82879246	10	1	-	-
	3	**C D E**	41684754	NA	1	1	-	-
	3	**B**	101236287	101618645	42	24	X	-
	5	**B**	28822515	29342972	12	3	X	-
	10	**C**	78846710	NA	1	1	-	-
NH	1		88009187	NA	1	1	-	-
	3		58464229	NA	1	1	-	-
	4	**B**	51903203	53474757	64	40	-	-
	6	**B**	63614756	63814984	20	10	-	-
	9		22745020	NA	1	1	-	-
	9	**A**	33549797	34165892	31	1	-	-
	11	**A F**	18987272	19176693	10	8	-	-
	14	**A**	63778931	63876998	7	2	X	-
	19	**D**	1134701	1139669	2	2	-	-
	19	**B**	32230245	33643392	55	2	-	-
	20	**A**	39797561	40162785	7	4	X	-
	20	**A**	59659997	60403627	11	4	X	X
	21	**A**	20193411	21256032	18	11	X	-
	24	**A**	33852631	34812035	36	23	X	-
GH	1	**A B E**	121484057	121775873	47	19	-	-
	1	**A**	131512239	131621826	3	3	X	-
	4	**A**	84181768	85275183	29	11	X	-
	11	**A F**	18987272	19176693	10	9	X	-
	17	**B**	32120145	32544617	23	4	X	-
	19	**A F**	28934939	NA	1	1	-	-
	20	**A**	63560971	63691145	10	6	-	-
	22	**C**	40135963	40167502	4	4	X	-

Table 3.4: Meta-analysis results for 11 metabolic traits (cont.)								
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**FE**	**Trad_RE**
LAM	1	**C**	49077969	NA	1	1	-	-
	2	**A**	36104151	36108219	6	6	-	-
	4	**B F**	17765473	18991639	11	3	-	-
	12	**B**	29378128	30296509	19	11	X	-
	14	**B**	88430222	89591967	20	5	-	-
	18	**B**	31679672	33134556	51	26	X	-
	19	**B F**	28057756	28417335	5	2	-	-
	19	**A**	57605404	58429206	36	20	X	-
	22	**B C E**	3565315	4307679	62	38	X	-
	23	**B**	12226548	12763020	35	24	X	-
	28	**A**	9446507	9643240	13	5	X	-

Table 3.4: Meta-analysis results for 11 metabolic traits. To be considered an MA-ROI, at least one SNP had to exceed the threshold for genome-wide significance (1.6e-07). Provided is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP, as well as the number of SNPs per region which exceeded the suggestive (Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. Summary column: region was identified as an ROI in the Welsh pony (A) or Morgan horse (B) GWA, region was identified in the Welsh pony (C) or Morgan horse (D) GWA for that trait but did not meet the criteria for an ROI, region was also identified as a shared region when analyzing the results of the standard ROI or LD-ROI (E), region was shared across two traits in the metanalysis (F) which is also represented by the corresponding highlighted chromosomes (Chr). Regions which were statistically significant using a fixed effects models (FE) or the traditional random effects model (Trad_RE) are indicated by an X in the respective columns.

	All Welsh ponies			Section ABCD			Section AB			Welsh Pony Final			Morgan Horses		
Trait	**High**	**Med**	**Low**	**High**	**Med**	**Low**	**High**	**Med**	**Low**	**High**	**Med**	**Low**	**High**	**Med**	**Low**
Insulin	3	3	1	0	2	0	2	1	0	4	4	1	1	1	1
Insulin-OST	2	0	0	1	0	1	2	1	0	3	1	0	2	0	1
Glucose	0	1	0	0	0	0	0	3	0	0	3	0	2	0	0
Glucose-OST	0	1	0	1	0	0	1	0	0	1	1	0	2	1	0
NEFA	0	0	0	2	0	0	1	1	0	2	0	0	4	1	0
TG	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0
Adiponectin	1	0	0	0	1	1	0	0	0	1	1	0	6	0	0
Leptin	3	1	0	2	1	2	1	0	1	3	2	3	3	1	0
ACTH	3	1	0	4	2	0	1	1	0	6	2	0	4	0	2
NH	8	1	2	3	3	2	0	0	2	10	4	5	5	2	1
GH	8	2	2	3	8	2	0	0	2	9	9	5	2	2	2
LAM	7	2	0	3	2	0	1	2	1	8	6	1	7	0	1
Total	**35**	**12**	**5**	**19**	**19**	**8**	**9**	**10**	**6**	**47**	**34**	**16**	**38**	**8**	**8**
Cohort Total	**52**			**46**			**25**			**114**			**54**		

Table 3.5: Summary table of prioritization of the fixed-sized regions in the Welsh ponies and Morgan horses. Data includes the full Welsh pony cohort (n=264), the section A, B, C and D (n=238), the section A and B Welsh ponies (n=220), the combined Welsh pony data (Welsh Pony Final), and the Morgan horses (n=286). Regions were categorized as high priority (regions found on metanalysis or region was shared with another trait and at least one region was considered an ROI), medium priority (region was an ROI in at least on GWA cohort), or low priority (region was shared across traits but region was not an ROI).

Table 3.6: Prioritization of the GWA results of the full Welsh pony cohort (n=264) based on fixed-sized regions									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Insulin	1		46119989	NA	1	0	45619989	46619989	3
	5	**H**	40632818	41895313	86	4	40132818	42395313	76
	6	**A**	82238815	82729921	11	0	81738815	83229921	17
	8		75410291	75771110	6	0	74910291	76271110	14
	9		83981022	84014912	6	0	83481022	84514912	49
	13		14234078	14849603	7	0	13734078	15349603	9
	13		37700109	37723843	3	1	37200109	38223843	1
	14		31226680	31583686	4	0	30726680	32083686	14
	15	**A H**	5748638	6140956	28	5	5248638	6640956	2
	15		54081224	54559632	5	1	53581224	55059632	19
	19		9606463	9637331	4	3	9106463	10137331	3
	23		46084858	46952228	6	3	45584858	47452228	6
	24	**H**	28580621	29056428	15	7	28080621	29556428	6
	34		38174280	NA	1	0	37674280	38674280	17
	28	**F**	38543945	NA	1	0	38043945	39043945	23
Insulin-OST	1		119102659	119140428	6	0	118602659	119640428	13
	8	**A**	73418239	73458142	4	0	72918239	73958142	5
	9		27879884	NA	1	0	27379884	28379884	13
	9		28468074	NA	1	0	27968074	28968074	15
	20		59181583	59182258	2	0	58681583	59682258	5
	28	**F H**	39385975	39462810	4	2	38885975	39962810	44
	31		8856537	8855069	2	0	8356537	9355069	18
Glucose	3		32093888	32098312	4	0	31593888	32598312	21
	8	**C**	84913969	85008392	2	2	84413969	85508392	3
	15		83778178	NA	1	1	83278178	84278178	14
	16		86563618	86743699	2	0	86063618	87243699	6
	22		42270349	42320092	2	0	41770349	42820092	10

Table 3.6: Prioritization of the GWA results of the full Welsh pony cohort (n=264) based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Glucose (cont.)	28		14969841	15045427	2	1	14469841	15545427	6
	28		33576312	NA	1	0	33076312	34076312	10
	29	**A**	22555245	22557916	2	0	22055245	23057916	15
Glucose-OST	1		185361795	NA	1	0	184861795	185861795	9
	4		8502301	8522723	4	0	8002301	9022723	10
	4		40150197	40352671	11	4	39650197	40852671	12
	7		76975314	NA	1	0	76475314	77475314	41
	23	**C**	10182647	10226427	4	2	9682647	10726427	5
NEFA	4		14831152	14840371	4	0	14331152	15340371	24
	6		68206430	68512033	2	0	67706430	69012033	36
	6	**C**	77102911	NA	1	0	76602911	77602911	6
	8		11128642	11158885	2	0	10628642	11658885	12
	8		69737476	NA	1	0	69237476	70237476	11
	19		1055718	NA	1	1	555718	1555718	11
	20		8830210	NA	1	0	8330210	9330210	9
	20		26078001	NA	1	0	25578001	26578001	31
	22	**C**	19009107	NA	1	0	18509107	19509107	13
	31		13902942	NA	1	0	13402942	14402942	6
TG	1		153409995	153700953	2	0	152909995	154200953	6
	4		93870436	NA	1	0	93370436	94370436	16
	12	**C**	35927778	35956541	4	1	35427778	36456541	45
Adiponectin	17		61546409	61552964	3	0	61046409	62052964	9
	18	**H**	60290699	60393507	10	5	59790699	60893507	12
Leptin	1		72370796	73160541	31	1	71870796	73660541	12
	5		43015591	43412260	24	0	42515591	43912260	56
	7	**A H**	67955613	67964668	5	4	67455613	68464668	4
	10	**C H**	866333	884264	3	1	366333	1384264	14

Table 3.6: Prioritization of the GWA results of the full Welsh pony cohort (n=264) based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Leptin (cont.)	21	**H**	22944751	23022779	2	1	22444751	23522779	1
	26		11291558	NA	1	0	10791558	11791558	36
	28		36456338	36459615	3	0	35956338	36959615	21
ACTH	1	**A H**	70266479	70832972	20	1	69766479	71332972	25
	5	**A**	16869826	17349383	30	4	16369826	17849383	3
	10	**H**	55658306	56077011	2	1	55158306	56577011	27
	10	**C H**	80023665	NA	1	0	79523665	80523665	6
	18		4529063	NA	1	0	4029063	5029063	18
	19	**C**	24243287	24246621	2	0	23743287	24746621	20
	19		33315383	33342063	4	0	32815383	33842063	11
	20		63684506	NA	1	0	63184506	64184506	93
	21		264658	2467359	12	0	-235342	2967359	25
	25		26250218	NA	1	0	25750218	26750218	22
NH	1		91537471	91969415	11	0	91037471	92469415	14
	4		62017772	62060721	2	2	61517772	62560721	6
	4	**C F G**	68114618	68576476	2	0	67614618	69076476	6
	4	**C**	77390519	NA	1	0	76890519	77890519	14
	4	**C**	78314683	78699729	3	1	77814683	79199729	22
	4		79698145	80390074	14	7	79198145	80890074	8
	6	**F**	1019810	1033178	3	0	519810	1533178	7
	7	**E**	93233594	93628623	10	6	92733594	94128623	9
	8	**C**	64510733	64609130	2	1	64010733	65109130	6
	8		88125499	88327659	2	1	87625499	88827659	13
	9	**H**	33913440	35808721	39	9	33413440	36308721	55
	11	**E H**	19050799	19240093	11	8	18550799	19740093	5
	12		7654801	7676262	2	0	7154801	8176262	18
	12		15601877	NA	1	0	15101877	16101877	7

Table 3.6: Prioritization of the GWA results of the full Welsh pony cohort (n=264) based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
NH (cont.)	14	**H**	63736228	63834285	7	3	63236228	64334285	25
	20	**H**	40661395	41066022	10	4	40161395	41566022	3
	20	**E H**	60832063	61575820	11	2	60332063	62075820	23
	21	**A E H**	20812917	22117426	22	13	20312917	22617426	16
	24		21429112	21604747	3	0	20929112	22104747	29
	24	**H**	33796794	35472785	62	33	33296794	35972785	11
	27		14461113	14463955	2	0	13961113	14963955	7
	29		33232105	33233161	2	0	32732105	33733161	20
GH	1	**C**	120905261	NA	1	0	120405261	121405261	25
	1	**H**	132203667	133711337	30	14	131703667	134211337	12
	1		150735268	NA	1	0	150235268	151235268	8
	4		67153317	67163513	2	0	66653317	67663513	5
	4	**C F G**	68576476	NA	1	0	68076476	69076476	49
	4	**A H**	84285316	85497218	63	33	83785316	85997218	9
	6	**F**	1019810	1154034	4	3	519810	1654034	6
	7	**E**	93233594	93580126	7	4	92733594	94080126	13
	9		55626969	55685330	2	0	55126969	56185330	17
	10		3673095	3673552	2	0	3173095	4173552	3
	10		32529022	32559811	4	0	32029022	33059811	59
	11	**E H**	18827291	19240093	14	9	18327291	19740093	29
	12		25641997	NA	1	0	25141997	26141997	24
	15		15062753	15656836	12	4	14562753	16156836	4
	19	**H**	31283482	31445588	9	1	30783482	31945588	53
	20	**A**	30141925	30449510	12	4	29641925	30949510	6
	20	**H**	64731849	64861251	12	7	64231849	65361251	7
	21	**C E**	22007711	NA	2	0	21507711	22507711	5
	22	**H**	41033715	41065262	4	4	40533715	41565262	6

Table 3.6: Prioritization of the GWA results of the full Welsh pony cohort (n=264) based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
GH (cont.)	24		18072215	18172937	4	0	17572215	18672937	17
	25		19485041	NA	1	1	18985041	19985041	7
	31		17898824	17912707	4	0	17398824	18412707	4
	31		18700158	18819670	4	0	18200158	19319670	22
LAM	1	**C H**	49441032	NA	1	1	48941032	49941032	27
	2	**A H**	36123836	36633565	14	11	35623836	37133565	3
	5		79658109	NA	1	0	79158109	80158109	7
	7		97437120	97439429	2	0	96937120	97939429	10
	10		64224504	64281425	2	0	63724504	64781425	16
	13		27143211	27035221	5	0	26643211	27535221	8
	14		52578019	52579053	2	0	52078019	53079053	3
	15	**C H**	50978261	51005138	3	1	50478261	51505138	7
	15		64654206	64769743	12	0	64154206	65269743	9
	16		35123235	35595543	3	0	34623235	36095543	8
	16	**C**	66694166	NA	1	0	66194166	67194166	24
	19		39125743	39626653	11	1	38625743	40126653	24
	19	**H**	59885237	61849890	78	35	59385237	62349890	4
	20	**E**	62018962	62085163	4	3	61518962	62585163	10
	22	**B**	3551367	NA	1	0	3051367	4051367	37
	25		32816803	32852556	2	0	32316803	33352556	5
	28	**H**	10461982	10666731	19	7	9961982	11166731	8
	31		10611327	10509324	5	3	10111327	11009324	3

Table 3.6: Prioritization of the GWA results of the full Welsh pony cohort (n=264) based on fixed-sized regions. To be considered an ROI, at least five SNP had to exceed the suggestive threshold (1.0e-05) with one SNP exceeding the threshold for genome-wide significance (5.98e-08). Provided in the table is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP in the region, the min (Min_Region) and maximum (Max_Region) boundaries of the region based on a fixed value of 500Kb 5' of the Min_SNP and 500Kb 3' of the Max_SNP, as well as the number of SNPs per region which exceeded the suggestive (Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. Total_Genes

includes all protein-coding, pseudogenes, and RNA genes identified for the region based on EquCab3. A black box in the summary column indicates the region did not meet the criteria to be considered an ROI and was not shared with another GWA cohort or trait. Letters in the summary column represent: (A) region was shared with another Welsh pony cohort and at least one region was considered an ROI, (B) region was shared with the Morgans and at least one region was considered an ROI, (C) region was shared with another Welsh pony cohort but neither regions met the criteria for an ROI, (D) region was shared with Morgan but neither regions met the criteria for an ROI, (E) region was shared with another trait in this cohort and at least one region was considered an ROI, (F) region was shared with another trait in this cohort but no regions met the criteria for an ROI, (G) region was shared across multiple traits in another cohort, (H) region was identified as shared across breeds on metanalysis and was considered an MA-ROI. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.7: Prioritization of the GWA results of the section A, B, C and D Welsh ponies based on fixed-sized regions									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Insulin	1		88317885	NA	1	0	87817885	88817885	2
	4		16025941	16029966	4	0	15525941	16529966	17
	6	**A**	82585066	83826234	32	2	82085066	84326234	26
	9	**A**	61021946	61144885	13	3	60521946	61644885	7
	12		859517	NA	1	0	359517	1359517	14
	12		5811391	5812573	3	0	5311391	6312573	3
	17		11193396	11343933	4	0	10693396	11843933	9
	18		79126354	79216656	3	1	78626354	79716656	10
	21		34649027	NA	1	0	34149027	35149027	3
	24		38174280	NA	1	0	37674280	38674280	17
Insulin-OST	1		176823704	NA	1	0	176323704	177323704	18
	1		181205641	NA	1	0	180705641	181705641	3
	3	**C**	67119398	NA	1	0	66619398	67619398	14
	5		89878763	NA	1	0	89378763	90378763	9
	6	**C F G**	15257536	15267456	3	1	14757536	15767456	12
	10	**B C H**	73334761	73417042	3	0	72834761	73917042	6
	31		5253579	NA	1	0	4753579	5753579	7
Glucose	2		88732913	88775982	2	0	88232913	89275982	6
	4		57433023	57463516	2	1	56933023	57963516	20
	8	**C**	84774486	84918226	5	0	84274486	85418226	3
	8		92368897	NA	1	0	91868897	92868897	12
	15		71112709	NA	1	0	70612709	71612709	10
	19		8738837	8781089	2	0	8238837	9281089	6
	29		4401376	4855454	4	1	3901376	5355454	7
Glucose-OST	5	**C**	63041759	63549216	12	0	62541759	64049216	16
	23	**C**	10217644	10226427	2	0	9717644	10726427	4
	28		14969841	NA	1	0	14469841	15469841	6

Table 3.7: Prioritization of the GWA results of the section A, B, C and D Welsh ponies based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Max_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Glucose-OST (cont)	28	**A E G H**	34960948	NA	1	1	34460948	35460948	23
NEFA	6	**C**	77102911	NA	1	0	76602911	77602911	6
	7	**C**	7744001	7858148	2	0	7244001	8358148	9
	7	**C**	8657121	8718823	3	0	8157121	9218823	3
	9	**A**	48687570	50301924	74	5	48187570	50801924	32
	13		3866723	NA	1	0	3366723	4366723	14
	14	**C**	33187338	NA	1	0	32687338	33687338	7
	22	**C**	19009107	NA	1	0	18509107	19509107	13
	28	**A E G**	34777499	35488520	25	6	34277499	35988520	43
	31		9275456	8325401	9	0	8775456	8825401	0
	32		21391267	21497776	2	0	20891267	21997776	0
TG	7		28031826	28039745	2	0	27531826	28539745	22
	12	**C**	35945816	35956541	2	0	35445816	36456541	45
	17		17532266	NA	2	0	17032266	18032266	4
	17		33912651	NA	1	0	33412651	34412651	5
	20	**C**	56719186	56815453	4	1	56219186	57315453	2
Adiponectin	1		175782149	177072407	32	1	175282149	177572407	23
	7		75100837	NA	1	0	74600837	75600837	54
	20		8415408	9139191	5	0	7915408	9639191	19
	22	**C**	37875269	37957795	3	0	37375269	38457795	20
	25		9125953	NA	1	0	8625953	9625953	6
	28	**F**	41052952	NA	1	0	40552952	41552952	18
Leptin	2		87434404	NA	1	0	86934404	87934404	13
	4	**F**	48014169	48031048	6	0	47514169	48531048	7
	6		2348093	2376386	6	1	1848093	2876386	3
	6		21686436	22141052	9	0	21186436	22641052	10
	7	**A H**	67955613	67964668	2	0	67455613	68464668	4

Table 3.7: Prioritization of the GWA results of the section A, B, C and D Welsh ponies based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Max_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Leptin (cont.)	8		87412707	NA	1	0	86912707	87912707	6
	10	**C H**	857433	884264	4	0	357433	1384264	14
	10		83363991	NA	1	0	82863991	83863991	15
	10		84395615	NA	1	0	83895615	84895615	8
	11		33592865	NA	1	0	33092865	34092865	40
	12		25937470	25914980	2	0	25437470	26414980	28
	13		7690179	NA	1	0	7190179	8190179	30
	15		24814175	24816607	2	0	24314175	25316607	3
	16		42665270	NA	1	0	42165270	43165270	14
	26		11291558	11425566	4	1	10791558	11925566	2
	28	**F**	40504716	NA	1	0	40004716	41004716	23
ACTH	1	**A F**	44391917	44627074	13	0	43891917	45127074	2
	3	**D**	44073772	44105888	8	0	43573772	44605888	6
	5	**A**	17101043	17252354	9	0	16601043	17752354	19
	10	**C H**	78845710	NA	1	1	78345710	79345710	10
	15		13711487	NA	1	0	13211487	14211487	11
	19	**C**	24243287	24246621	2	0	23743287	24746621	18
	19	**E**	37642432	NA	1	0	37142432	38142432	8
	20	**E**	60431850	NA	3	0	59931850	60931850	0
	24	**C G**	39497717	NA	1	0	38997717	39997717	14
	30		1302176	1304866	4	0	802176	1804866	13
NH	1	**F**	44398249	NA	1	0	43898249	44898249	2
	3		109783963	NA	1	0	109283963	110283963	11
	4	**A E G**	69118549	69714717	17	14	68618549	70214717	7
	4		72715285	73055056	2	0	72215285	73555056	5
	4	**C G**	76437287	77891737	17	0	75937287	78391737	14
	4	**E**	83194842	85546563	54	38	82694842	86046563	67

Table 3.7: Prioritization of the GWA results of the section A, B, C and D Welsh ponies based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Max_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
NH (cont.)	8	**C**	64510733	64548459	2	0	64010733	65048459	9
	14		57777287	NA	1	0	57277287	58277287	6
	15		44638315	NA	1	3	44138315	45138315	1
	15		73207370	73338081	5	3	72707370	73838081	13
	16		19756213	20195426	10	0	19256213	20695426	12
	16		26012920	NA	1	0	25512920	26512920	8
	17		36798348	36856368	9	9	36298348	37356368	5
	18		68917978	NA	1	1	68417978	69417978	5
	20	**F**	30160893	NA	1	1	29660893	30660893	42
	21		6664894	6835706	5	2	6164894	7335706	11
	21	**A F G H**	20812917	21679286	4	0	20312917	22179286	19
	24		8561717	11119679	17	0	8061717	11619679	43
	24		30419181	30419482	2	0	29919181	30919482	2
	26		13669849	NA	1	0	13169849	14169849	7
	29		12571950	12620905	5	0	12071950	13120905	11
	30		30980395	NA	1	0	30480395	31480395	26
GH	1		73428660	73434597	2	0	72928660	73934597	2
	1	**C**	120654637	120684527	5	0	120154637	121184527	22
	4	**F**	46793329	47851529	10	0	46293329	48351529	10
	4		52253860	NA	1	0	51753860	52753860	9
	4		61808230	62195564	3	0	61308230	62695564	17
	4	**A E G**	69236860	69660028	5	0	68736860	70160028	6
	4		74331018	76480368	41	7	73831018	76980368	25
	4		79807650	80390074	12	8	79307650	80890074	22
	4	**A E H**	84044345	85375688	34	18	83544345	85875688	56
	5		20984165	NA	1	0	20484165	21484165	8
	9		75259988	75263736	2	1	74759988	75763736	10

Table 3.7: Prioritization of the GWA results of the section A, B, C and D Welsh ponies based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Max_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
GH (cont.)	10		51267938	NA	1	0	50767938	51767938	9
	11		15497051	16248131	36	26	14997051	16748131	40
	13		1235274	531380	29	19	735274	1031380	8
	14		5739617	6265883	11	0	5239617	6765883	20
	15		85423316	NA	1	0	84923316	85923316	27
	16		27103995	28231661	41	11	26603995	28731661	17
	16		88165628	88202104	5	0	87665628	88702104	10
	17		167021	NA	1	0	-332979	667021	12
	17		57101997	NA	1	0	56601997	57601997	8
	18		70918093	NA	1	0	70418093	71418093	4
	18		75058371	76075236	26	6	74558371	76575236	31
	18		80391110	81050756	27	1	79891110	81550756	25
	20	**A F**	30141925	30160893	3	1	29641925	30660893	42
	20	**E**	60935600	61788330	8	1	60435600	62288330	3
	21		18238312	18263289	2	0	17738312	18763289	15
	21	**C F G**	21539433	21542349	2	0	21039433	22042349	9
	22		43609456	NA	1	0	43109456	44109456	6
	24		22090203	22552582	9	0	21590203	23052582	32
	25		25896326	25897963	2	0	25396326	26397963	16
LAM	1	**C H**	49441032	NA	1	0	48941032	49941032	22
	2		29737934	29777141	17	7	29237934	30277141	33
	2	**A H**	36322824	36633565	8	0	35822824	37133565	27
	11		37530491	37555597	2	0	37030491	38055597	33
	16		9221468	NA	1	0	8721468	9721468	10
	17		46013130	46020667	3	0	45513130	46520667	10
	18		26676637	NA	1	0	26176637	27176637	2
	19	**E**	37272294	37328619	15	3	36772294	37828619	10

Table 3.7: Prioritization of the GWA results of the section A, B, C and D Welsh ponies based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Max_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
LAM (cont.)	20		48609221	NA	1	0	48109221	49109221	15
	23		4297964	4341498	5	0	3797964	4841498	8
	27		4129106	NA	1	0	3629106	4629106	10
	30		12607858	13128439	15	9	12107858	13628439	15

Table 3.7: Prioritization of the GWA results of the section A, B, C and D Welsh ponies (n=238) based on fixed-sized regions. To be considered an ROI, at least five SNP had to exceed the suggestive threshold (1.0e-05) with one SNP exceeding the threshold for genome-wide significance (5.98e-08). Provided in the table is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP in the region, the min (Min_Region) and maximum (Max_Region) boundaries of the region based on a fixed value of 500Kb 5' of the Min_SNP and 500Kb 3' of the Max_SNP, as well as the number of SNPs per region which exceeded the suggestive (Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. Total_Genes includes all protein-coding, pseudogenes, and RNA genes identified for the region based on EquCab3. A black box in the summary column indicates the region did not meet the criteria to be considered an ROI and was not shared with another GWA cohort or trait. Letters in the summary column represent: (A) region was shared with another Welsh pony cohort and at least one region was considered an ROI, (B) region was shared with the Morgans and at least one region was considered an ROI, (C) region was shared with another Welsh pony cohort but neither regions met the criteria for an ROI, (D) region was shared with Morgan but neither regions met the criteria for an ROI, (E) region was shared with another trait in this cohort and at least one region was considered an ROI, (F) region was shared with another trait in this cohort but no regions met the criteria for an ROI, (G) region was shared across multiple traits in another cohort, (H) region was identified as shared across breeds on metanalysis and was considered an MA-ROI. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.8: Prioritization of the GWA results of the section A and B Welsh ponies based on fixed-sized regions									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Insulin	4		4672416	NA	1	0	4172416	5172416	4
	8	**E**	71911322	73187367	71	48	71411322	73687367	30
	9	**A**	60513208	62048722	21	6	60013208	62548722	14
	10		72007185	72019609	2	0	71507185	72519609	8
	11		20890216	NA	1	0	20390216	21390216	53
	14		44171754	NA	1	0	43671754	44671754	9
	15	**A H**	5760603	NA	1	0	5260603	6260603	2
	16		85072895	85226499	2	0	84572895	85726499	23
	18		27406051	27434219	3	0	26906051	27934219	12
	21		37151683	NA	1	0	36651683	37651683	4
Insulin-OST	3	**C**	67827963	NA	1	0	67327963	68327963	15
	6	**C**	15257536	NA	1	0	14757536	15757536	12
	8	**E**	73418276	73458142	3	0	72918276	73958142	5
	9		51694853	52360209	26	1	51194853	52860209	7
	10	**B C H**	73415161	NA	1	0	72915161	73915161	6
	23		33075103	NA	2	0	32575103	33575103	8
Glucose	2		78104573	NA	1	0	77604573	78604573	13
	4		91598735	91632300	3	0	91098735	92132300	6
	5		60653615	61066511	3	1	60153615	61566511	5
	8		89457249	90171577	19	3	88957249	90671577	5
	14		6606837	6628311	4	3	6106837	7128311	13
	17		79020897	79997119	21	4	78520897	80497119	33
	29	**A**	22523122	23217070	6	5	22023122	23717070	22
	31		7659497	7654406	3	0	7159497	8154406	7
Glucose-OST	5	**C**	63447777	63549216	3	0	62947777	64049216	8
	16		86361940	NA	1	1	85861940	86861940	6
	28	**A E H**	34434081	34960948	5	1	33934081	35460948	32

Table 3.8: Prioritization of the GWA results of the section A and B Welsh ponies based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
NEFA	5		17682899	NA	1	0	17182899	18182899	12
	6		74667806	74721945	4	0	74167806	75221945	61
	7		5590146	NA	1	0	5090146	6090146	42
	7	**C**	7744001	NA	1	0	7244001	8244001	8
	7		90384141	90387298	2	0	89884141	90887298	9
	9	**A**	48876850	50428786	66	10	48376850	50928786	29
	14	**C**	33144705	33289979	13	0	32644705	33789979	8
	18		21325941	22264264	2	0	20825941	22764264	10
	20		31639261	NA	1	0	31139261	32139261	20
	22	**C**	19009107	19028315	2	0	18509107	19528315	14
	28	**A E G**	34865969	34877252	7	0	34365969	35377252	16
TG	1		47645272	NA	1	0	47145272	48145272	1
	2		98328483	NA	1	0	97828483	98828483	3
	4		88686448	NA	1	0	88186448	89186448	12
	7		26533379	26635921	7	1	26033379	27135921	26
	9		73409149	73438018	4	0	72909149	73938018	11
	20	**C**	56347955	NA	1	0	55847955	56847955	4
Adiponectin	8		5894342	NA	1	0	5394342	6394342	20
	18		39196722	NA	1	0	38696722	39696722	9
	20		26633993	NA	1	0	26133993	27133993	41
	22	**C**	37957795	NA	1	1	37457795	38457795	19
Leptin	10	**C H**	872249	NA	1	0	372249	1372249	14
	14	**F**	60295756	NA	1	0	59795756	60795756	6
	17		5633648	NA	1	1	5133648	6133648	6
ACTH	1	**A**	44284734	45133993	30	1	43784734	45633993	5
	5		17101043	17252354	4	0	16601043	17752354	19
	10	**C H**	79880592	80023665	2	1	79380592	80523665	27

Table 3.8: Prioritization of the GWA results of the section A and B Welsh ponies based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
ACTH (cont.)	11		58160240	NA	1	0	57660240	58660240	29
	20		8331002	8355327	7	0	7831002	8855327	16
	24	**C G**	38516095	39660384	23	0	38016095	40160384	34
NH	1		9347701	NA	1	0	8847701	9847701	13
	3		69624972	NA	1	0	69124972	70124972	6
	4	**A G**	67618110	69482711	14	3	67118110	69982711	16
	10	**F**	10827320	NA	1	0	10327320	11327320	47
	14		73473354	NA	1	0	72973354	73973354	11
	21	**F**	23990259	24995726	2	0	23490259	25495726	19
GH	1		166271712	NA	1	0	165771712	166771712	7
	10	**F**	11229405	NA	1	0	10729405	11729405	37
	10		70502635	70536766	2	1	70002635	71036766	13
	12		20064456	NA	1	0	19564456	20564456	19
	17		27064422	NA	1	0	26564422	27564422	8
	21	**F**	23776930	23991948	5	0	23276930	24491948	14
	24		49764166	NA	1	0	49264166	50264166	0
	22		23930066	NA	1	0	23430066	24430066	26
	25		15030393	NA	1	0	14530393	15530393	13
LAM	3		77977500	NA	1	0	77477500	78477500	21
	8		45552432	NA	1	0	45052432	46052432	4
	10		15374259	15988198	15	2	14874259	16488198	73
	13		24882636	25740597	8	2	24382636	26240597	32
	14	**F**	58930834	59667233	4	0	58430834	60167233	15
	15	**C H**	50978261	51005138	3	1	50478261	51505138	3
	16	**C**	66471008	66521264	2	0	65971008	67021264	9
	18		15365144	NA	1	0	14865144	15865144	4
	19		54249861	54263396	4	0	53749861	54763396	5

Table 3.8: Prioritization of the GWA results of the section A and B Welsh ponies based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
LAM (cont.)	20		43136147	43150142	3	0	42636147	43650142	39

Table 3.8: Prioritization of the GWA results of the section A and B Welsh ponies (n=220) based on fixed-sized regions. To be considered an ROI, at least five SNP had to exceed the suggestive threshold (1.0e-05) with one SNP exceeding the threshold for genome-wide significance (5.98e-08). Provided in the table is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP in the region, the min (Min_ROI) and maximum (Max_ROI) boundaries of the region based on a fixed value of 500Kb 5' of the Min_SNP and 500Kb 3' of the Max_SNP, as well as the number of SNPs per region which exceeded the suggestive (Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. The total number of genes includes all protein-coding genes, pseudogenes, and RNA genes identified for region based on EquCab3. A black box in the summary column indicates the region did not meet the criteria to be considered an ROI and was not shared with another GWA cohort or trait. Letters in the summary column represent: (A) region was shared with another Welsh pony cohort GWA and at least one region was considered an ROI, (B) region was shared with the Morgan GWA and at least one region was considered an ROI, (C) region was shared with another Welsh pony cohort GWA but no regions met the criteria for an ROI, (D) region was shared with Morgan GWA but no regions met the criteria for an ROI, (E) region was shared with another trait in this GWA cohort and at least one region was considered an ROI, (F) region was shared with another trait in this GWA cohort but no regions met the criteria for an ROI, (G) region was shared across multiple traits in another GWA cohort, (H) region was identified as shared across breeds on metanalysis and was considered a MA-ROI. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.9: High Priority Regions Welsh ponies based on fixed-sized regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	5	40132818	42395313	67	0	9	76
	8	71411322	73687367	16	0	14	30
	15	5248638	6640956	0	0	2	2
	24	28080621	29556428	0	0	6	6
Insulin-OST	8	72918276	73958142	1	0	4	5
	10	72834761	73917042	5	1	0	6
	28	38885975	39962810	42	0	2	44
Glucose-OST	28	33934081	35460948	26	0	6	32
NEFA	9	48187570	50801924	19	0	13	32
	28	34277499	35988520	34	0	9	43
Adiponectin	18	59790699	60893507	8	0	4	12
Leptin	7	67455613	68464668	4	0	0	4
	10	366333	1384264	2	0	12	14
	21	22444751	23522779	2	0	5	7
ACTH	1	69766479	71332972	10	0	11	21
	3	43573772	44605888	2	1	3	6
	10	55158306	56577011	2	0	1	3
	10	79380592	80523665	20	0	7	27
	19	37142432	38142432	7	0	1	8
	20	59931850	60931850	0	0	0	0
NH	4	68618549	70214717	5	0	2	7
	4	82694842	86046563	47	1	19	67
	7	92733594	94128623	2	0	5	7
	9	33413440	36308721	6	0	7	13
	11	18550799	19740093	51	1	3	55
	14	63236228	64334285	2	0	5	7
	20	40161395	41566022	14	0	11	25
	20	60332063	62075820	1	0	2	3

Table 3.9: High Priority Regions Welsh ponies based on fixed-sized regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
NH (cont.)	21	20312917	22617426	4	0	19	23
	24	33296794	35972785	26	0	3	29
GH	1	131703667	134211337	17	0	8	25
	4	68736860	70160028	5	0	1	6
	4	83544345	85997218	39	1	17	57
	7	92733594	94080126	2	0	4	6
	11	18327291	19740093	55	1	3	59
	19	30783482	31945588	4	0	0	4
	20	60435600	62288330	1	0	2	3
	20	64231849	65361251	3	0	3	6
	22	40533715	41565262	1	0	4	5
LAM	1	48941032	49941032	6	0	16	22
	2	35623836	37133565	20	0	7	27
	15	50478261	51505138	2	0	1	3
	19	36772294	37828619	10	0	0	10
	19	59385237	62349890	17	1	6	24
	20	61518962	62585163	2	0	2	4
	22	3051367	4051367	5	0	5	10
	28	9961982	11166731	4	0	1	5
Total				**618**	**7**	**265**	**890**

Medium Priority Regions Welsh ponies based on fixed-sized regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	6	82085066	84326234	11	0	15	26
	9	60013208	62548722	4	0	10	14
	15	53581224	55059632	7	0	12	19
	23	45584858	47452228	1	0	5	6
Insulin-OST	9	51194853	52860209	4	0	3	7

Table 3.9: Medium Priority Regions Welsh ponies based on fixed-sized regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Glucose	8	88957249	90671577	2	0	3	5
Glucose (cont.)	17	78520897	80497119	21	0	12	33
	29	22023122	23717070	16	0	6	22
Glucose-OST	4	39650197	40852671	6	0	6	12
TG	7	26033379	27135921	24	0	2	26
Adiponectin	1	175282149	177572407	12	0	11	23
Leptin	1	71870796	73660541	9	0	3	12
	6	1848093	2876386	2	0	1	3
ACTH	1	43784734	45633993	2	0	3	5
	5	16369826	17849383	19	0	6	25
NH	4	79198145	80890074	14	0	8	22
	15	72707370	73838081	7	0	6	13
	17	36298348	37356368	1	0	4	5
	21	6164894	7335706	4	0	7	11
GH	4	73831018	76980368	13	0	12	25
	4	79307650	80890074	14	0	8	22
	11	14997051	16748131	34	0	6	40
	13	735274	1031380	6	0	2	8
	15	14562753	16156836	15	0	9	24
	16	26603995	28731661	14	0	3	17
	18	74558371	76575236	26	0	5	31
	18	79891110	81550756	13	0	12	25
	20	29641925	30949510	45	0	8	53
LAM	2	29237934	30277141	30	0	3	33
	10	14874259	16488198	64	0	9	73
	13	24382636	26240597	29	0	3	32
	19	38625743	40126653	23	0	1	24
	30	12107858	13628439	7	0	8	15

Table 3.9: Medium Priority Regions Welsh ponies based on fixed-sized regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
LAM (cont.)	31	10111327	11009324	4	0	4	8
Total				**503**	**0**	**216**	**719**

Low Priority Regions Welsh ponies based on fixed-sized regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	28	38043945	39043945	18	0	5	23
Adiponectin	28	40552952	41552952	14	0	4	18
Leptin	4	47514169	48531048	3	0	4	7
	14	59795756	60795756	5	0	1	6
	28	40004716	41004716	17	0	6	23
NH	1	43898249	44898249	1	0	1	2
	6	519810	1533178	6	0	2	8
	10	10327320	11327320	42	2	3	47
	20	29660893	30660893	36	0	6	42
	21	23490259	25495726	16	0	3	19
GH	4	46293329	48351529	3	0	7	10
	6	519810	1654034	6	0	3	9
	10	10729405	11729405	34	0	3	37
	21	21039433	22042349	3	0	6	9
	21	23276930	24491948	12	0	2	14
LAM	14	58430834	60167233	9	0	6	15
Total				**225**	**2**	**62**	**289**

Table 3.9: Final prioritization of the GWA results of the Welsh pony cohorts based on fixed-sized regions. Regions were categorized as high priority (regions found on metanalysis or was shared with another trait and considered an ROI), medium priority (region was identified as an ROI in at least one GWA), or low priority (region was shared with across traits but region was not an ROI). Final region boundaries of the region were defined as 500Kb 5' of the lowest SNP (Min_ROI) and 500Kb 3' of the highest (Max_ROI) SNP across relevant GWA data. The total number of genes includes all protein-coding genes, pseudogenes, and RNA genes identified for region based on EquCab3. Shared regions across prioritized traits are indicated by highlighted chromosomes.

Table 3.10: Prioritization of the GWA results of Morgan horses based on fixed-sized regions									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Insulin	2		117366086	117410894	5	1	116866086	117910894	9
	3		115316619	115326166	4	0	114816619	115826166	12
	4		97370223	NA	1	0	96870223	97870223	30
	5		88722709	NA	2	0	88222709	89222709	6
	8		36946690	NA	1	0	36446690	37446690	13
	8		62414695	62422169	3	0	61914695	62922169	5
	10		54997568	55022644	3	0	54497568	55522644	1
	18		38197723	NA	1	0	37697723	38697723	7
	19		20841248	NA	1	0	20341248	21341248	17
	20	**F**	4635861	4702640	7	0	4135861	5202640	16
	24	**F H**	21134897	NA	1	0	20634897	21634897	14
	26		39653507	NA	1	0	39153507	40153507	22
Insulin-OST	2		22468309	22541921	4	1	21968309	23041921	16
	2		51548258	51661415	7	0	51048258	52161415	33
	4	**E**	28373202	NA	1	0	27873202	28873202	2
	4		57780431	57786154	2	0	57280431	58286154	15
	6		32931767	33694226	2	0	32431767	34194226	30
	8		10116471	NA	1	0	9616471	10616471	17
	10	**B H**	71996093	73613162	50	5	71496093	74113162	17
	11	**F**	18848207	19009809	7	0	18348207	19509809	47
	20		51914168	NA	1	0	51414168	52414168	22
	21		20781491	NA	1	0	20281491	21281491	12
Glucose	4	**E H**	17981325	18477651	33	11	17481325	18977651	9
	8	**H**	11530408	12159746	5	1	11030408	12659746	18
	16		42711571	NA	1	0	42211571	43211571	15
	28		36615983	NA	1	0	36115983	37115983	34
	29		9494870	NA	1	0	8994870	9994870	14
	31		21504871	NA	1	0	21004871	22004871	9

Table 3.10: Prioritization of the GWA results of Morgan horses based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Glucose-OST	2		62607747	NA	1	0	62107747	63107747	11
	3	**H**	56674808	58220254	85	53	56174808	58720254	19
	4	**E H**	27505119	28710128	39	4	27005119	29210128	8
	14		28998387	29000329	2	0	28498387	29500329	24
	25		18872032	NA	1	0	18372032	19372032	23
	26		22407530	23379414	23	2	21907530	23879414	6
NEFA	1		166669064	166888483	3	0	166169064	167388483	8
	1	**H**	185892360	186617146	25	15	185392360	187117146	37
	2		106012533	106052266	6	1	105512533	106552266	16
	7		86986401	87004808	3	0	86486401	87504808	4
	9		76549280	76571642	3	0	76049280	77071642	13
	15		66056425	NA	1	0	65556425	66556425	12
	17	**H**	13427110	14189583	14	1	12927110	14689583	6
	18		7685942	9565563	44	0	7185942	10065563	18
	19		48235446	NA	1	0	47735446	48735446	21
	24	**F H**	20381260	20888104	2	1	19881260	21388104	32
	24		45325106	45675218	5	0	44825106	46175218	28
	30		6239856	6258423	5	0	5739856	6758423	9
	30	**H**	20974703	21044590	11	4	20474703	21544590	3
TG	1		126407798	127401777	6	0	125907798	127901777	32
	10		65383517	NA	1	0	64883517	65883517	6
	20		52368013	52589211	4	1	51868013	53089211	9
	21		49201984	49202284	2	0	48701984	49702284	2
Adiponectin	1		129650721	129653375	2	0	129150721	130153375	18
	1		138037003	NA	1	0	137537003	138537003	8
	2	**H**	16747148	17739125	38	27	16247148	18239125	50
	4	**H**	36557672	38544490	54	4	36057672	39044490	35
	6	**H**	32601529	32727370	19	1	32101529	33227370	15

Table 3.10: Prioritization of the GWA results of Morgan horses based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Adiponectin (cont)	6	**H**	67997807	69847785	68	6	67497807	70347785	82
	7		21524454	21986901	14	0	21024454	22486901	26
	7		32963159	32963459	2	0	32463159	33463459	37
	8		3347264	3419299	6	0	2847264	3919299	26
	15		21830373	21834175	2	0	21330373	22334175	4
	15		66865469	66893151	4	0	66365469	67393151	8
	18	**F H**	41448414	NA	1	1	40948414	41948414	10
	18		49705278	49893633	7	0	49205278	50393633	28
	19		25833383	25859655	2	0	25333383	26359655	14
	20	**F H**	3734902	3954772	12	0	3234902	4454772	16
	20		1882774	NA	1	0	1382774	2382774	11
	21		49478363	NA	1	0	48978363	49978363	1
Leptin	1		130957068	131062691	3	0	130457068	131562691	11
	4	**E**	52373692	52614368	22	0	51873692	53114368	14
	6		38446793	NA	1	0	37946793	38946793	21
	8		8682147	NA	1	0	8182147	9182147	37
	19	**H**	51360775	53132722	57	27	50860775	53632722	26
	21		16547954	16608200	3	0	16047954	17108200	8
	24	**H**	27275709	29038412	65	14	26775709	29538412	10
	25		27438558	27907420	14	2	26938558	28407420	31
ACTH	1	**E H**	83546191	83734040	17	4	83046191	84234040	23
	3	**D H**	43335201	44116411	13	0	42835201	44616411	10
	3	**H**	103056163	103438726	49	34	102556163	103938726	8
	5	**H**	25785666	27061038	32	10	25285666	27561038	26
	10		67992633	67997136	2	0	67492633	68497136	12
	10		70528773	NA	1	0	70028773	71028773	12
	11	**F**	18728679	18904099	4	0	18228679	19404099	44
	11		52897545	53669056	32	0	52397545	54169056	24

Table 3.10: Prioritization of the GWA results of Morgan horses based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
ACTH (cont.)	13		25806289	NA	1	0	25306289	26306289	15
	16		31200001	NA	1	0	30700001	31700001	13
	18	**F**	41392781	NA	1	0	40892781	41892781	10
	20		29056288	NA	1	0	28556288	29556288	45
	21		11112604	NA	1	0	10612604	11612604	11
	21		24436227	24439739	3	0	23936227	24939739	12
	25		13299542	NA	1	0	12799542	13799542	15
	25		14989527	NA	1	0	14489527	15489527	13
	31		16965044	17737242	4	0	16465044	18237242	30
NH	1	**F**	78493587	79782621	37	0	77993587	80282621	20
	1	**E**	82958480	83232130	10	0	82458480	83732130	15
	2		93824111	93833011	2	0	93324111	94333011	3
	4	**E H**	52076906	53659651	149	110	51576906	54159651	22
	5		59796357	60233277	10	0	59296357	60733277	7
	5		65804297	65824216	3	0	65304297	66324216	3
	6	**H**	64502443	65350057	44	12	64002443	65850057	17
	8		29756282	NA	1	0	29256282	30256282	7
	9		49062306	49078134	2	0	48562306	49578134	8
	14		74532493	NA	1	0	74032493	75032493	9
	18		2306238	NA	1	0	1806238	2806238	7
	19	**H**	1188889	1197320	3	2	688889	1697320	10
	19	**H**	34421059	36247260	102	23	33921059	36747260	48
	19		46479290	47156982	5	1	45979290	47656982	25
	21		4745903	NA	1	0	4245903	5245903	31
	24		42026470	42450741	9	1	41526470	42950741	23
GH	1	**F**	79175507	79234421	5	0	78675507	79734421	9
	1		109778420	109819993	4	0	109278420	110319993	13
	1	**H**	122383349	123036781	71	31	121883349	123536781	20

Table 3.10: Prioritization of the GWA results of Morgan horses based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
GH (cont.)	2		85183513	86093522	21	10	84683513	86593522	14
	4		3032922	NA	1	1	2532922	3532922	7
	6	**F**	3139850	3453652	2	0	2639850	3953652	10
	6		6272129	6335115	8	0	5772129	6835115	27
	6		15729023	16202020	8	0	15229023	16702020	16
	7		26684853	26701040	6	0	26184853	27201040	28
	8		63557829	NA	1	0	63057829	64057829	9
	17	**H**	32020513	33031579	39	2	31520513	33531579	6
	18		2423391	NA	1	0	1923391	2923391	9
	22		45719751	48733979	2	0	45219751	49233979	70
	29		19108245	19432974	7	2	18608245	19932974	15
LAM	2		66192812	NA	1	1	65692812	66692812	11
	3		3294278	NA	1	0	2794278	3794278	10
	4	**E H**	17509325	19295909	52	4	17009325	19795909	18
	6	**F**	3466933	NA	1	0	2966933	3966933	8
	6		79661858	NA	1	0	79161858	80161858	4
	8		59199626	60121756	24	0	58699626	60621756	21
	12	**H**	33127411	34414133	53	27	32627411	34914133	54
	14		66311023	66688404	15	0	65811023	67188404	7
	14	**H**	88975206	90135630	48	9	88475206	90635630	37
	16		64556111	NA	1	0	64056111	65056111	8
	16		74667638	NA	1	0	74167638	75167638	21
	18	**H**	31710749	33317633	65	33	31210749	33817633	18
	19	**H**	30133826	NA	51	3	29633826	30633826	6
	22	**B H**	3616445	4853827	75	45	3116445	5353827	21
	22		13852015	NA	1	0	13352015	14352015	11
	22		23806850	NA	1	0	23306850	24306850	25
	23	**H**	11116499	12515439	51	46	10616499	13015439	19

Table 3.10: Prioritization of the GWA results of Morgan horses based on fixed-sized regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
LAM (cont.)	31		6804894	NA	1	0	6304894	7304894	3

Table 3.10: Prioritization of the GWA results of Morgan horses based on fixed-sized regions (n=296). To be considered an ROI, at least five SNP had to exceed the suggestive threshold (1.0e-05) with one SNP exceeding the threshold for genome-wide significance (5.98e-08). Provided in the table is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP in the region, the min (Min_ROI) and maximum (Max_ROI) boundaries of the region based on a fixed value of 500Kb 5' of the Min_SNP and 500Kb 3' of the Max_SNP, as well as the number of SNPs per region which exceeded the suggestive (Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. The total number of genes includes all protein-coding genes, pseudogenes, and RNA genes identified for region based on EquCab3. A black box in the summary column indicates the region did not meet the criteria to be considered an ROI and was not shared with one or more Welsh pony cohorts or trait. Letters in the summary column represent: (B) region was shared with one or more Welsh pony GWA cohorts and at least one region was considered an ROI, (D) region was shared with one or more Welsh pony GWA cohort but no regions met the criteria for an ROI, (E) region was shared with another trait in the Morgan GWA and at least one region was considered an ROI, (F) region was shared with another trait in the Morgan GWA but no regions met the criteria for an ROI, (H) region was identified as shared across breeds on metanalysis and was considered a MA-ROI. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.11: High Priority Regions Morgan horses based on fixed-sized regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	24	20634897	21634897	11	0	3	14
Insulin-OST	4	27873202	28873202	2	0	0	2
	10	71496093	74113162	11	1	5	17
Glucose	4	17481325	18977651	6	0	3	9
	8	11030408	12659746	9	0	9	18
Glucose-OST	3	56174808	58720254	12	0	7	19
	4	27005119	29210128	5	0	3	8
NEFA	1	185392360	187117146	24	0	13	37
	17	12927110	14689583	4	0	2	6
	24	19881260	21388104	28	0	4	32
	30	20474703	21544590	1	0	2	3
Adiponectin	2	16247148	18239125	26	1	23	50
	4	36057672	39044490	28	0	7	35
	6	32101529	33227370	9	0	6	15
	6	67497807	70347785	77	0	5	82
	18	40948414	41948414	4	0	6	10
	20	3234902	4454772	12	0	4	16
Leptin	4	51873692	53114368	6	0	8	14
	19	50860775	53632722	13	0	13	26
	24	26775709	29538412	0	0	10	10
ACTH	1	83046191	84234040	17	1	5	23
	3	42835201	44616411	2	1	7	10
	3	102556163	103938726	3	0	5	8
	5	25285666	27561038	12	0	14	26
NH	1	82458480	83732130	10	0	5	15
	4	51576906	54159651	10	0	12	22
	6	64002443	65850057	9	0	8	17
	19	688889	1697320	5	1	4	10

Table 3.11: High Priority Regions Morgan horses based on fixed-sized regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
NH (cont.)	19	33921059	36747260	40	0	8	48
GH	1	121883349	123536781	19	0	1	20
	17	31520513	33531579	2	1	3	6
LAM	4	17009325	19795909	8	1	9	18
	12	32627411	34914133	31	0	23	54
	14	88475206	90635630	23	0	14	37
	18	31210749	33817633	11	0	7	18
	19	29633826	30633826	6	0	0	6
	22	3116445	5353827	12	0	9	21
	23	10616499	13015439	4	0	15	19
Total				**512**	**7**	**282**	**801**

Medium Priority Regions Morgan horses based on fixed-sized regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	2	116866086	117910894	6	0	3	9
Glucose-OST	26	21907530	23879414	0	0	6	6
NEFA	2	105512533	106552266	13	0	3	16
Leptin	25	26938558	28407420	29	1	1	31
NH	19	45979290	47656982	22	0	3	25
	24	41526470	42950741	15	0	8	23
GH	2	84683513	86593522	8	0	6	14
	29	18608245	19932974	13	0	2	15
Total				**106**	**1**	**32**	**139**

Low Priority Regions Morgan horses based on fixed-sized regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	20	4135861	5202640	9	0	7	16
Insulin-OST	11	18348207	19509809	44	1	2	47

Table 3.11: Low Priority Regions Morgan horses based on fixed-sized regions (cont)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
ACTH	11	18228679	19404099	42	1	1	44
	18	40892781	41892781	4	0	6	10
NH	1	77993587	80282621	8	0	12	20
GH	1	78675507	79734421	5	0	4	9
	6	2639850	3953652	4	0	6	10
LAM	6	2966933	3966933	4	0	4	8
Total				**120**	**2**	**42**	**164**

Table 3.11: Final prioritization of the GWA results for the Morgan horses based on fixed-sized regions. Regions were categorized as high priority (regions found on metanalysis or was shared with another trait and considered an ROI), medium priority (region was identified as an ROI in at least one GWA), or low priority (region was shared with across traits but region was not an ROI). Final region boundaries of the region were defined as 500Kb 5' of the lowest SNP (Min_ROI) and 500Kb 3' of the highest (Max_ROI) SNP across relevant GWA data. The total number of genes includes all protein-coding genes, pseudogenes, and RNA genes identified for region based on EquCab3. Shared regions across prioritized traits are indicated by highlighted chromosomes.

Table 3.12: Prioritization of the GWA results of the full Welsh pony cohort based on LD-defined regions									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SN**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Insulin	1		46119989	NA	1	0	46069989	46169989	0
	5	**E H**	40632818	41895313	86	4	35409104	44806458	306
	6	**A**	82238815	82729921	11	0	81685572	83066256	17
	8	**A**	75410291	75771110	6	0	74768024	76080554	13
	9		83981022	84014912	6	0	83924785	84136547	11
	13		14234078	14849603	7	0	12830836	16025943	19
	13		37700109	37723843	3	1	37642362	37732522	0
	14		31226680	31583686	4	0	31004830	31595632	5
	15	**A H**	5748638	6140956	28	5	5748377	6612684	1
	15		54081224	54559632	5	1	54076168	54634446	5
	19		9606463	9637331	4	3	9604680	9680011	0
	23		46084858	46952228	6	3	45940500	46233500	2
	24	**H**	28580621	29056428	15	7	28451012	29887250	6
	34		38174280	NA	1	0	38124280	38224280	0
	28		38543945	NA	1	0	38493945	38593945	4
Insulin-OST	1		119102659	119140428	6	0	117422338	119310838	17
	8	**C**	73418239	73458142	4	0	73223448	73648399	2
	9		27879884	NA	1	0	27829884	27929884	1
	9		28468074	NA	1	0	28418074	28518074	3
	20		59181583	59182258	2	0	59165490	59819879	5
	28	**H**	39385975	39462810	4	2	39322188	39488807	9
	31		8856537	8855069	2	0	8487501	9386141	16
Glucose	3		32093888	32098312	4	0	32038651	32100834	2
	8	**C**	84913969	85008392	2	2	84907297	85013865	0
	15	**H**	83778178	NA	1	1	83728178	83828178	2
	16		86563618	86743699	2	0	86419997	87063084	3
	22		42270349	42320092	2	0	42222066	42357143	2

Table 3.12: Prioritization of the GWA results of the full Welsh pony cohort based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SN**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Glucose (cont.)	28		14969841	15045427	2	1	14968044	15079576	0
	28		33576312	NA	1	0	33526312	33626312	1
	29	**A**	22555245	22557916	2	0	22420947	22764383	9
Glucose-OST	1		185361795	NA	1	0	185311795	185411795	2
	4		8502301	8522723	4	0	8380352	8750255	5
	4		40150197	40352671	11	4	40143954	40782593	9
	7		76975314	NA	1	0	76925314	77025314	5
	23	**C**	10182647	10226427	4	2	9741040	10226824	2
NEFA	4		14831152	14840371	4	0	14262252	15798579	40
	6		68206430	68512033	2	0	67662851	69169794	39
	6	**C**	77102911	NA	1	0	77052911	77152911	1
	8		11128642	11158885	2	0	10809267	11158996	5
	8		69737476	NA	1	0	69687476	69787476	1
	19	**H**	1055718	NA	1	1	1005718	1105718	2
	20		8830210	NA	1	0	8780210	8880210	1
	20		26078001	NA	1	0	26028001	26128001	15
	22	**C**	19009107	NA	1	0	18959107	19059107	1
	31		13902942	NA	1	0	13852942	13952942	0
TG	1		153409995	153700953	2	0	152034119	154351987	11
	4		93870436	NA	1	0	93820436	93920436	3
	12	**C**	35927778	35956541	4	1	35859835	36313678	13
Adiponectin	17		61546409	61552964	3	NA	61267646	61749127	7
	18	**H**	60290699	60393507	10	5	60060215	61349045	13
Leptin	1		72370796	73160541	31	1	71902092	78569116	57
	5	**E**	43015591	43412260	24	NA	39751797	50431769	239
	7	**A H**	67955613	67964668	5	4	67910114	68117086	1
	10	**C H**	866333	884264	3	1	692055	956048	4

Table 3.12: Prioritization of the GWA results of the full Welsh pony cohort based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SN**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Leptin (cont.)	21	**E H**	NA	NA	2	1	22940681	23516697	1
	26		11291558	NA	1	NA	11241558	11341558	1
	28		36456338	36459615	3	NA	36070224	36467537	15
ACTH	1	**H**	70266479	70832972	20	1	69558737	70960589	23
	5	**A**	16869826	17349383	30	4	16534115	18234765	26
	10	**H**	55658306	56077011	2	1	55060512	56255134	2
	10	**C H**	80023665	NA	1	0	79973665	80073665	6
	18		4529063	NA	1	0	4479063	4579063	1
	19	**C**	24243287	24246621	2	0	24226051	24312494	6
	19		33315383	33342063	4	0	33290960	33465193	4
	20		63684506	NA	1	0	63634506	63734506	1
	21		264658	2467359	12	0	-566031	5646555	188
	25		26250218	NA	1	0	26200218	26300218	5
NH	1		91537471	91969415	11	0	91257387	92299877	13
	4		62017772	62060721	2	2	61306607	62232293	14
	4	**A F G**	68114618	68576476	2	0	67588953	69039376	6
	4		77390519	NA	1	0	77340519	77440519	1
	4		78314683	78699729	3	1	77300209	79809543	26
	4	**A**	79698145	80390074	14	7	77298241	81186565	40
	6	**F**	1019810	1033178	3	0	903258	1451922	2
	7	**E**	93233594	93628623	10	6	93176991	93628686	1
	8	**C**	64510733	64609130	2	1	64277141	64856816	5
	8		88125499	88327659	2	1	87999710	88892517	5
	9	**H**	33913440	35808721	39	9	32632235	37587269	18
	11	**E H**	19050799	19240093	11	8	18342117	19876247	60
	12		7654801	7676262	2	0	7314424	7786773	1
	12		15601877	NA	1	0	15551877	15651877	1

Table 3.12: Prioritization of the GWA results of the full Welsh pony cohort based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SN**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
NH (cont.)	14	**H**	63736228	63834285	7	3	63702522	63847210	2
	20	**H**	40661395	41066022	10	4	40244007	41210876	14
	20	**H**	60832063	61575820	11	2	60723014	61735694	2
	21	**A E H**	20812917	22117426	22	13	19515280	23543447	33
	24		21429112	21604747	3	0	20723403	21637158	13
	24	**H**	33796794	35472785	62	33	31843480	36758215	57
	27		14461113	14463955	2	NA	14185003	14663824	6
	29		33232105	33233161	2	NA	32443212	34147999	11
GH	1	**C**	120905261	NA	1	0	120855261	120955261	3
	1	**H**	132203667	133711337	30	14	132184772	133716124	16
	1		150735268	NA	1	0	150685268	150785268	1
	4		67153317	67163513	2	0	67079978	67434976	4
	4	**F**	68576476	NA	1	0	68526476	68626476	1
	4	**A H**	84285316	85497218	63	33	81804323	85719241	69
	6	**F**	1019810	1154034	4	3	903258	1734708	5
	7	**E**	93233594	93580126	7	4	93191676	93628672	1
	9		55626969	55685330	2	0	55565458	56169206	9
	10		3673095	3673552	2	0	3671537	3801071	0
	10		32529022	32559811	4	0	32333094	33276546	2
	11	**E H**	18827291	19240093	14	9	18613895	19317536	26
	12		25641997	NA	1	NA	25591997	25691997	2
	15		15062753	15656836	12	4	13131438	16662645	55
	19	**H**	31283482	31445588	9	1	31204596	31799125	0
	20	**A**	30141925	30449510	12	4	29486630	30976763	62
	20	**H**	64731849	64861251	12	7	64722427	65336095	4
	21	**E H**	22007711	NA	2	0	21957711	22057711	1
	22	**H**	41033715	41065262	4	4	41032889	41066045	0

Table 3.12: Prioritization of the GWA results of the full Welsh pony cohort based on on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SN**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
GH (cont.)	24		18072215	18172937	4	0	18024756	18508703	2
	25	**H**	19485041	NA	1	1	19435041	19535041	4
	31		17898824	17912707	4	0	17847583	18148648	2
	31		18700158	18819670	4	0	18685559	18858145	1
LAM	1	**C H**	49441032	NA	1	1	49391032	49491032	1
	2	**A H**	36123836	36633565	14	11	35909634	36665473	13
	5		79658109	NA	1	0	79608109	79708109	1
	7		97437120	97439429	2	0	97415193	97521397	0
	10		64224504	64281425	2	0	64224315	64303766	1
	13		27143211	27035221	5	0	27081260	27274065	2
	14		52578019	52579053	2	0	52549348	52667915	2
	15	**C H**	50978261	51005138	3	1	50973563	51006110	0
	15		64654206	64769743	12	0	64265615	65947460	16
	16		35123235	35595543	3	0	35116321	35769589	6
	16	**C**	66694166	NA	1	0	66644166	66744166	1
	19		39125743	39626653	11	1	37990377	39825664	27
	19	**H**	59885237	61849890	78	35	57082025	62825378	59
	20		62018962	62085163	4	3	61971048	62085845	1
	22	**D**	3551367	NA	1	0	3501367	3601367	1
	25		32816803	32852556	2	0	32681105	32892597	7
	28	**H**	10461982	10666731	19	7	9990892	10844823	4
	31		10611327	10509324	5	3	10611124	10918134	3

Table 3.12: Prioritization of the GWA results of the full Welsh pony cohort (n=264) based on LD-defined regions. To be considered an ROI, at least five SNP had to exceed the suggestive threshold (1.0e-05) with one SNP exceeding the threshold for genome-wide significance (5.98e-08). Provided in the table is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP in the region, the min (Min_Region) and maximum (Max_Region) boundaries of the region based on LD, as well as the number of SNPs per region which exceeded the suggestive

(Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. Total_Genes includes all protein-coding genes, pseudogenes, and RNA genes identified for region based on EquCab3. A black box in the summary column indicates the region did not meet the criteria to be considered an ROI and was not shared with another GWA cohort or trait. Letters in the summary column represent: (A) region was shared with another Welsh pony cohort GWA and at least one region was considered an ROI, (B) region was shared with the Morgan GWA and at least one region was considered an ROI, (C) region was shared with another Welsh pony cohort GWA but no regions met the criteria for an ROI, (D) region was shared with Morgan GWA but no regions met the criteria for an ROI, (E) region was shared with another trait in this GWA cohort and at least one region was considered an ROI, (F) region was shared with another trait in this GWA cohort but no regions met the criteria for an ROI, (G) region was shared across multiple traits in another GWA cohort, (H) region was identified as shared across breeds on metanalysis and was considered a MA-ROI. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.13: Prioritization of the GWA results of section A, B, C and D Welsh ponies based on LD-defined regions									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Insulin	1		88317885	NA	1	0	88267885	88367885	0
	4		16025941	16029966	4	0	15938986	16180572	4
	6	**A**	82585066	83826234	32	2	80572788	83599194	37
	9	**A**	61021946	61144885	13	3	58372873	61241976	19
	12		859517	NA	1	0	809517	909517	2
	12		5811391	5812573	3	0	5254445	6105807	3
	17		11193396	11343933	4	0	10851587	11442936	3
	18	**E**	79126354	79216656	3	1	78720858	79634082	6
	21		34649027	NA	1	0	34599027	34699027	0
	24		38174280	NA	1	0	38124280	38224280	2
Insulin-OST	1	**E**	176823704	NA	1	0	176773704	176873704	1
	1		181205641	NA	1	0	181155641	181255641	0
	3		67119398	NA	1	0	67069398	67169398	2
	5		89878763	NA	1	0	89828763	89928763	2
	6	**C**	15257536	15267456	3	1	15328522	15802598	5
	10	**B C H**	73334761	73417042	3	0	71967783	72438937	3
	31		5253579	NA	1	0	5203579	5303579	2
Glucose	2		88732913	88775982	2	0	87700669	88704605	7
	4		57433023	57463516	2	1	56762597	57781825	21
	8	**C**	84774486	84918226	5	0	81206774	81830595	2
	8	**A**	92368897	NA	1	0	92318897	92418897	2
	15		71112709	NA	1	0	71062709	71162709	2
	19		8738837	8781089	2	0	6175910	6939325	4
	29		4401376	4855454	4	1	3141136	3956266	3
Glucose-OST	5	**C**	63041759	63549216	12	0	64707107	68397655	30
	23	**C**	10217644	10226427	2	0	10729234	11001562	0
	28		14969841	NA	1	0	14919841	15019841	0

Table 3.13: Prioritization of the GWA results of section A, B, C and D Welsh ponies based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Glucose-OST (cont)	28	**A E G H**	34960948	NA	1	1	34910948	35010948	4
NEFA	6	**C**	77102911	NA	1	0	77052911	77152911	1
	7	**C**	7744001	7858148	2	0	7084283	7235090	1
	7		8657121	8718823	3	0	7794692	8979276	3
	9	**A**	48687570	50301924	74	5	43402596	51140717	71
	13		3866723	NA	1	0	3816723	3916723	2
	14	**C**	33187338	NA	1	0	33137338	33237338	1
	22	**C**	19009107	NA	1	0	18959107	19059107	1
	28	**A E G**	34777499	35488520	25	6	32909542	35703535	76
	31		9275456	8325401	9	0	9225456	9325456	2
	32		21391267	21497776	2	0	20203883	21904415	0
TG	7		28031826	28039745	2	0	27130249	27583217	6
	12	**C**	35945816	35956541	2	0	31936287	32490177	17
	17		17532266	NA	2	0	17482266	17582266	2
	17		33912651	NA	1	0	33862651	33962651	1
	20		56719186	56815453	4	1	53870768	56991611	21
Adiponectin	1	**E**	175782149	177072407	32	1	171861236	178270042	49
	7		75100837	NA	1	0	75050837	75150837	7
	20		8415408	9139191	5	0	6054702	11308877	65
	22	**C**	37875269	37957795	3	0	36564923	37148074	9
	25		9125953	NA	1	0	9075953	9175953	1
	28		41052952	NA	1	0	41002952	41102952	1
Leptin	2		87434404	NA	1	0	87384404	87484404	1
	4	**F**	48014169	48031048	6	0	47052514	48193459	6
	6		2348093	2376386	6	1	488137	4012580	25
	6		21686436	22141052	9	0	21323668	22378607	6
	7	**A H**	67955613	67964668	2	0	65678376	65985348	2

Table 3.13: Prioritization of the GWA results of section A, B, C and D Welsh ponies based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Leptin (cont.)	8		87412707	NA	1	0	87362707	87462707	1
	10	**C H**	857433	884264	4	0	809627	1068890	5
	10		83363991	NA	1	0	83313991	83413991	1
	10		84395615	NA	1	0	84345615	84445615	2
	11		33592865	NA	1	0	33542865	33642865	7
	12		25937470	25914980	2	0	22339448	2294098	0
	13		7690179	NA	1	0	7640179	7740179	3
	15		24814175	24816607	2	0	23820813	24351119	3
	16		42665270	NA	1	0	42615270	42715270	1
	26		11291558	11425566	4	1	10695805	10988327	1
	28		40504716	NA	1	NA	40454716	40554716	0
ACTH	1	**A F**	44391917	44627074	13	0	43064008	44872306	7
	3	**D**	44073772	44105888	8	0	41547556	43484146	9
	5	**A**	17101043	17252354	9	0	18357723	21333724	36
	10	**C H**	78845710	NA	1	1	78795710	78895710	3
	15		13711487	NA	1	0	13661487	13761487	1
	19	**C**	24243287	24246621	2	0	21800444	22098695	7
	19		37642432	NA	1	0	37592432	37692432	2
	20	**E**	60431850	NA	3	0	60381850	60481850	0
	24	**C**	39497717	NA	1	0	39447717	39547717	0
	30		1302176	1304866	4	0	962895	1567967	8
NH	1	**F**	44398249	NA	1	0	44348249	44448249	0
	3		109783963	NA	1	0	109733963	109833963	1
	4	**A E G**	69118549	69714717	17	14	68425678	69636837	6
	4		72715285	73055056	2	0	72321830	72867956	2
	4	**C**	76437287	77891737	17	0	73229449	79629933	56
	4	**E**	83194842	85546563	54	38	83144842	83244842	1

Table 3.13: Prioritization of the GWA results of section A, B, C and D Welsh ponies based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
NH (cont.)	8	**C**	64510733	64548459	2	0	60856045	61415070	6
	14		57777287	NA	1	0	57727287	57827287	1
	15		44638315	NA	1	3	44588315	44688315	0
	15		73207370	73338081	5	3	73033562	73478127	3
	16		19756213	20195426	10	0	17609338	19833772	15
	16		26012920	NA	1	0	25962920	26062920	0
	17		36798348	36856368	9	9	36704887	37121452	0
	18		68917978	NA	1	1	68867978	68967978	1
	20	**F**	30160893	NA	1	1	30110893	30210893	1
	21		6664894	6835706	5	2	5280993	6396786	8
	21	**A F H**	20812917	21679286	4	0	20606675	21752563	9
	24		8561717	11119679	17	0	4971965	15097475	125
	24		30419181	30419482	2	0	30330777	30578714	0
	26		13669849	NA	1	0	13619849	13719849	0
	29		12571950	12620905	5	0	10813364	11907852	14
	30		30980395	NA	1	0	30930395	31030395	5
GH	1		73428660	73434597	2	0	72506297	72913313	1
	1	**C**	120654637	120684527	5	0	119265228	119651584	9
	4	**F**	46793329	47851529	10	0	47576773	47975728	4
	4		52253860	NA	1	0	52203860	52303860	1
	4		61808230	62195564	3	0	61460640	62542569	11
	4	**E**	69236860	69660028	5	0	68425678	69636837	6
	4		74331018	76480368	41	7	70026254	81648125	95
	4		79807650	80390074	12	8	80015281	81132522	9
	4	**A E H**	84044345	85375688	34	18	82570011	86366835	75
	5		20984165	NA	1	0	20934165	21034165	2
	9		75259988	75263736	2	1	73067974	73332751	3

Table 3.13: Prioritization of the GWA results of section A, B, C and D Welsh ponies based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
GH (cont.)	10		51267938	NA	1	0	51217938	51317938	1
	11		15497051	16248131	36	26	15414337	16451463	25
	13		1235274	531380	29	19	4097503	6272661	18
	14		5739617	6265883	11	0	6039980	7464225	18
	15		85423316	NA	1	0	85373316	85473316	2
	16		27103995	28231661	41	11	25105634	30681811	44
	16		88165628	88202104	5	0	85453703	87408801	25
	17		167021	NA	1	0	117021	217021	2
	17		57101997	NA	1	0	57051997	57151997	2
	18		70918093	NA	1	0	70868093	70968093	0
	18		75058371	76075236	26	6	74790214	76353283	23
	18	**E**	80391110	81050756	27	1	79527484	81467661	25
	20	**A F**	30141925	30160893	3	1	28634038	30114993	64
	20	**E**	60935600	61788330	8	1	59464566	61015217	3
	21		18238312	18263289	2	0	17104955	18679882	20
	21	**F**	21539433	21542349	2	0	20611963	21174919	6
	22		43609456	NA	1	0	43559456	43659456	0
	24		22090203	22552582	9	0	21281696	23226701	36
	25		25896326	25897963	2	0	23582383	26321736	50
LAM	1	**C H**	49441032	NA	1	0	49391032	49491032	1
	2		29737934	29777141	17	7	29447761	29803535	8
	2	**A H**	36322824	36633565	8	0	35880861	36496556	12
	11		37530491	37555597	2	0	36920316	37335219	15
	16		9221468	NA	1	0	9171468	9271468	1
	17		46013130	46020667	3	0	46036409	46209083	3
	18		26676637	NA	1	0	26626637	26726637	1
	19		37272294	37328619	15	3	34513667	34812456	4

Table 3.13: Prioritization of the GWA results of section A, B, C and D Welsh ponies based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
LAM (cont.)	20		48609221	NA	1	0	48559221	48659221	3
	23		4297964	4341498	5	0	3933380	4310086	5
	27		4129106	NA	1	0	4079106	4179106	1
	30		12607858	13128439	15	9	11660801	12736188	12

Table 3.13: Prioritization of the GWA results of section A, B, C and D Welsh ponies (n=238) based on LD-defined regions. To be considered an ROI, at least five SNP had to exceed the suggestive threshold (1.0e-05) with one SNP exceeding the threshold for genome-wide significance (5.98e-08). Provided in the table is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP in the region, the min (Min_Region) and maximum (Max_Region) boundaries of the region based on LD, as well as the number of SNPs per region which exceeded the suggestive (Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. Total_Genes includes all protein-coding genes, pseudogenes, and RNA genes based on EquCab3. Black box in the summary column indicates the region did not meet the criteria to be considered an ROI and was not shared with another GWA cohort or trait. Letters in the summary column represent: (A) region was shared with another Welsh pony cohort and at least one region was considered an ROI, (B) region was shared with the Morgans and at least one region was considered an ROI, (C) region was shared with another Welsh pony cohort but no regions met the criteria for an ROI, (D) region was shared with Morgan but no regions met the criteria for an ROI, (E) region was shared with another trait in this cohort and at least one region was considered an ROI, (F) region was shared with another trait in this cohort but no regions met the criteria for an ROI, (G) region was shared across multiple traits in another cohort, (H) region was identified as shared across breeds on metanalysis and was considered a MA-ROI. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.14: Prioritization of the GWA results of section A and B Welsh ponies based on LD-defined regions									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Insulin	4		4672416	NA	1	0	4622416	4722416	1
	8	**A E**	71911322	73187367	71	48	69350844	75906595	55
	9	**A**	60513208	62048722	21	6	60380309	63287617	19
	10		72007185	72019609	2	0	71920915	72410422	5
	11		20890216	NA	1	0	20840216	20940216	9
	14		44171754	NA	1	0	44121754	44221754	2
	15	**A H**	5760603	NA	1	0	5710603	5810603	2
	16		85072895	85226499	2	0	84940660	85543835	12
	18		27406051	27434219	3	0	27257163	27649509	10
	21		37151683	NA	1	0	37101683	37201683	2
Insulin-OST	3		67827963	NA	1	0	67777963	67877963	3
	6	**C**	15257536	NA	1	0	15207536	15307536	2
	8	**C E**	73418276	73458142	3	0	73173455	73699198	2
	9		51694853	52360209	26	1	51519922	52222979	2
	10	**B C H**	73415161	NA	1	0	73365161	73465161	2
	23		33075103	NA	2	0	33025103	33125103	1
Glucose	2		78104573	NA	1	0	78054573	78154573	0
	4		91598735	91632300	3	0	91430847	91770165	2
	5		60653615	61066511	3	1	58829349	61396408	8
	8	**A**	89457249	90171577	19	3	86176351	93189207	41
	14		6606837	6628311	4	3	6496358	6828698	4
	17		79020897	79997119	21	4	78895931	80077759	22
	29	**A**	22523122	23217070	6	5	22370951	22609138	5
	31		7659497	7654406	3	0	7278750	7741436	2
Glucose-OST	5	**C**	63447777	63549216	3	0	62823019	64472035	13
	16		86361940	NA	1	1	86311940	86411940	1
	28	**A E G H**	34434081	34960948	5	1	34271949	35138699	9

Table 3.14: Prioritization of the GWA results of section A and B Welsh ponies based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
NEFA	5		17682899	NA	1	0	17632899	17732899	2
	6		74667806	74721945	4	0	74462718	75137224	42
	7		5590146	NA	1	0	5540146	5640146	5
	7	**C**	7744001	NA	1	0	7694001	7794001	1
	7		90384141	90387298	2	0	90103889	90535803	6
	9	**A**	48876850	50428786	66	10	48031329	51265805	37
	14	**C**	33144705	33289979	13	0	32851327	33404045	5
	18		21325941	22264264	2	0	20477464	22315518	9
	20		31639261	NA	1	0	31589261	31689261	2
	22	**C**	19009107	19028315	2	0	17869740	19275273	17
	28	**A E G**	34865969	34877252	7	0	34727198	35029308	5
TG	1		47645272	NA	1	0	47595272	47695272	0
	2		98328483	NA	1	0	98278483	98378483	1
	4		88686448	NA	1	0	88636448	88736448	1
	7		26533379	26635921	7	1	26358820	26960566	21
	9		73409149	73438018	4	0	72531878	73855746	15
	20		56347955	NA	1	0	56297955	56397955	1
Adiponectin	8		5894342	NA	1	0	5844342	5944342	2
	18		39196722	NA	1	0	39146722	39246722	1
	20		26633993	NA	1	0	26583993	26683993	1
	22	**C**	37957795	NA	1	1	37907795	38007795	1
Leptin	10	**C H**	872249	NA	1	0	822249	922249	1
	14		60295756	NA	1	0	60245756	60345756	1
	17		5633648	NA	1	1	5583648	5683648	1
ACTH	1	**A**	44284734	45133993	30	1	42944403	45232767	9
	5		17101043	17252354	4	0	16484006	18102564	25
	10	**C H**	79880592	80023665	2	1	79691144	80306613	25

Table 3.14: Prioritization of the GWA results of section A and B Welsh ponies based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
ACTH (cont.)	11		58160240	NA	1	0	58110240	58210240	5
	20		8331002	8355327	7	0	8092967	8599500	7
	24	**C**	38516095	39660384	23	0	38361711	40099306	28
NH	1		9347701	NA	1	0	9297701	9397701	1
	3		69624972	NA	1	0	69574972	69674972	1
	4	**A**	67618110	69482711	14	3	67130904	69873296	16
	10		10827320	NA	1	0	10777320	10877320	3
	14		73473354	NA	1	0	73423354	73523354	0
	21	**F**	23990259	24995726	2	0	23600027	25046226	16
GH	1		166271712	NA	1	0	166221712	166321712	1
	10		11229405	NA	1	0	11179405	11279405	9
	10		70502635	70536766	2	1	69524859	70587090	6
	12		20064456	NA	1	0	20014456	20114456	1
	17		27064422	NA	1	0	27014422	27114422	2
	21	**F**	23776930	23991948	5	0	23171361	24411682	12
	24		49764166	NA	1	0	49714166	49814166	0
	22		23930066	NA	1	0	23880066	23980066	2
	25		15030393	NA	1	0	14980393	15080393	2
LAM	3		77977500	NA	1	0	77927500	78027500	3
	8		45552432	NA	1	0	45502432	45602432	1
	10		15374259	15988198	15	2	14730688	16165003	66
	13		24882636	25740597	8	2	24242621	26399066	36
	14		58930834	59667233	4	0	58326568	60185720	17
	15	**C H**	50978261	51005138	3	1	50923563	51056110	0
	16	**C**	66471008	66521264	2	0	66309711	66691345	5
	18		15365144	NA	1	0	15315144	15415144	1
	19		54249861	54263396	4	0	54026036	54315136	2

Table 3.14: Prioritization of the GWA results of section A and B Welsh ponies based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
LAM (cont.)	20		43136147	43150142	3	0	43086126	43312141	9

Table 3.14: Prioritization of the GWA results of section A and B Welsh ponies (n=220) based on LD-defined regions. To be considered an ROI, at least five SNP had to exceed the suggestive threshold (1.0e-05) with one SNP exceeding the threshold for genome-wide significance (5.98e-08). Provided in the table is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP in the region, the min (Min_ROI) and maximum (Max_ROI) boundaries of the region based on LD, as well as the number of SNPs per region which exceeded the suggestive (Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. The total number of genes includes all protein-coding genes, pseudogenes, and RNA genes based on EquCab3. Black box in the summary column indicates the region did not meet the criteria to be considered an ROI and was not shared with another GWA cohort or trait. Letters in the summary column represent: (A) region was shared with another Welsh pony cohort GWA and at least one region was considered an ROI, (B) region was shared with the Morgan GWA and at least one region was considered an ROI, (C) region was shared with another Welsh pony cohort GWA but no regions met the criteria for an ROI, (D) region was shared with Morgan GWA but no regions met the criteria for an ROI, (E) region was shared with another trait in this GWA cohort and at least one region was considered an ROI, (F) region was shared with another trait in this GWA cohort but no regions met the criteria for an ROI, (G) region was shared across multiple traits in another GWA cohort, (H) region was identified as shared across breeds on metanalysis and was considered a MA-ROI. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.15: High Priority Region in Welsh ponies based on LD-defined regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	5	35409104	44806458	267	1	38	306
	8	69350844	75906595	32	0	23	55
	15	5748377	6612684	0	0	1	1
	18	78720858	79634082	2	0	4	6
	24	28451012	29887250	2	0	4	6
Insulin-OST	1	176773704	176873704	0	0	1	1
	8	73173455	73699198	1	0	1	2
	10	71967783	72438937	3	0	0	3
	28	39322188	39488807	8	0	1	9
Glucose	15	83728178	83828178	2	0	0	2
Glucose-OST	28	34271949	35138699	9	0	0	9
Adiponectin	1	171861236	178270042	25	0	24	49
	18	60060215	61349045	7	0	6	13
Leptin	5	39751797	50431769	207	0	32	239
	6	488137	4012580	15	0	10	25
	7	65678376	68117086	1	0	2	3
	10	692055	1068890	0	0	5	5
	21	22940681	23516697	1	0	0	1
NEFA	19	1005718	1105718	2	0	0	2
	28	32909542	35703535	65	0	11	76
ACTH	1	42944403	45232767	5	0	4	9
	1	69558737	70960589	7	0	16	23
	10	55060512	56255134	1	0	1	2
	10	78795710	80306613	20	0	5	25
	20	60381850	60481850	0	0	0	0
NH	4	67130904	69873296	8	0	8	16
	4	77298241	81186565	24	1	15	40
	4	83144842	83244842	1	0	0	1

Table 3.15: High Priority Region in Welsh Ponies based on LD-defined regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
NH (cont.)	7	93176991	93628686	0	0	1	1
	9	32632235	37587269	10	0	8	18
	11	18342117	19876247	55	1	4	60
	14	63702522	63847210	0	0	2	2
	20	40244007	41210876	3	0	11	14
	20	60723014	61735694	0	0	2	2
	21	5280993	6396786	2	0	6	8
	21	19515280	25046226	22	0	27	49
	24	31843480	36758215	47	0	10	57
GH	1	132184772	133716124	9	0	7	16
	4	68425678	69636837	5	0	1	6
	4	70026254	81648125	49	1	45	95
	4	82570011	86366835	49	1	25	75
	7	93191676	93628672	0	0	1	1
	11	15414337	16451463	24	0	1	25
	11	18613895	19317536	26	0	0	26
	18	79527484	81467661	13	0	12	25
	19	31204596	31799125	0	0	0	0
	20	29486630	30976763	54	0	8	62
	20	59464566	61015217	1	0	2	3
	20	64722427	65336095	1	0	3	4
	21	20611963	22057711	3	0	4	7
	22	41032889	41066045	0	0	0	0
	25	19435041	19535041	4	0	0	4
LAM	1	49391032	49491032	0	0	1	1
	2	35880861	36665473	8	0	6	14
	19	57082025	62825378	42	1	16	59
	28	9990892	10844823	4	0	0	4

Table 3.15: High Priority Regions in Welsh ponies based on LD-defined regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Total				**1146**	**6**	**415**	**1567**

Medium Priority Regions in Welsh ponies based on LD-defined regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	6	80572788	83599194	17	0	20	37
	9	60380309	63287617	7	0	12	19
	15	54076168	54634446	2	0	3	5
	23	45940500	46233500	0	0	2	2
Insulin-OST	9	51519922	52222979	1	0	1	2
Glucose	8	86176351	93189207	22	0	19	41
	17	78895931	80077759	13	0	9	22
	29	22370951	22764383	7	0	2	9
Glucose-OST	4	40143954	40782593	5	0	4	9
Leptin	1	71902092	78569116	34	0	23	57
NEFA	9	43402596	51140717	48	0	23	71
TG	7	26358820	26960566	19	0	2	21
ACTH	5	16534115	18234765	22	0	4	26
NH	15	73033562	73478127	1	0	2	3
	17	36704887	37121452	0	0	0	0
GH	13	4097503	6272661	8	0	10	18
	15	13131438	16662645	42	1	12	55
	16	25105634	30681811	31	1	12	44
	18	74790214	76353283	21	0	2	23
LAM	2	29447761	29803535	7	0	1	8
	10	14730688	16165003	56	0	10	66
	13	24242621	26399066	32	0	4	36
	19	34513667	34812456	4	0	0	4
	19	37990377	39825664	23	0	4	27

Table 3.15: Medium Priority Regions in Welsh ponies based on LD-defined regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
LAM (cont.)	30	11660801	12736188	7	0	5	12
	31	10611124	10918134	2	0	1	3
Total				**431**	**2**	**187**	**620**

Low Priority Regions in Welsh ponies based on LD-defined regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Leptin	4	47052514	48193459	3	0	3	6
NH	1	44348249	44448249	0	0	0	0
	6	903258	1451922	1	0	1	2
	20	30110893	30210893	1	0	0	1
GH	4	47576773	47975728	1	0	3	4
	6	903258	1734708	2	0	3	5
	21	23171361	24411682	10	0	2	12
Total				**18**	**0**	**12**	**30**

Table 3.15: Final boundaries of the regions based on LD and positional candidate genes of the prioritization GWA results for the Welsh ponies. Regions were categorized as high priority (regions found on metanalysis OR region was shared with another trait), medium priority (region was an ROI in at least one Welsh pony cohort but was not shared), or low priority (region was shared across traits but region was not an ROI). Final region boundaries of the region were based on LD-ROI and are indicated by the lowest base pair position (Min_ROI) and the highest base pair position (Max_ROI). The total number of genes includes all protein-coding genes, pseudogenes, and RNA genes identified for region based on EquCab3. Shared regions across prioritized traits are indicated by highlighted chromosomes.

Table 3.16: Prioritization of the GWA results of the Morgan horses based on LD-defined regions									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Insulin	2		117366086	117410894	5	1	117310352	117579882	4
	3		115316619	115326166	4	0	114849263	115698498	10
	4		97370223	NA	1	0	97370223	97420223	3
	5		88722709	NA	2	0	88722709	88772709	1
	8		36946690	NA	1	0	36946690	36996690	2
	8		62414695	62422169	3	0	61953438	62651012	5
	10		54997568	55022644	3	0	54821584	55225831	0
	18		38197723	NA	1	0	38197723	38247723	0
	19		20841248	NA	1	0	20841248	20891248	2
	20		4635861	4702640	7	0	4544080	5465175	12
	24	**H**	21134897	NA	1	0	21134897	21184897	1
	26		39653507	NA	1	0	39653507	39703507	2
Insulin-OST	2		22468309	22541921	4	1	21941652	22859290	15
	2		51548258	51661415	7	0	51173763	52005569	27
	4	**E**	28373202	NA	1	0	28373202	28423202	0
	4		57780431	57786154	2	0	57533782	57927057	9
	6	**E**	32931767	33694226	2	0	32751552	34029749	22
	8		10116471	NA	1	0	10116471	10166471	3
	10	**B H**	71996093	73613162	50	5	71666607	73534053	12
	11	**F**	18848207	19009809	7	0	18355073	19629302	53
	20		51914168	NA	1	0	51914168	51964168	1
	21		20781491	NA	1	0	20781491	20831491	1
Glucose	4	**E H**	17981325	18477651	33	11	17239374	19043831	11
	8		11530408	12159746	5	1	11193683	12404572	17
	16		42711571	NA	1	0	42711571	42761571	1
	28		36615983	NA	1	0	36615983	36665983	2
	29		9494870	NA	1	0	9494870	9544870	2

Table 3.16: Prioritization of the GWA results of the Morgan horses based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Glucose (cont.)	31		21504871	NA	1	0	21504871	21554871	0
Glucose-OST	2		62607747	NA	1	0	62607747	62657747	1
	3	**H**	56674808	58220254	85	53	55746338	58085997	21
	4	**E H**	27505119	28710128	39	4	26695616	29116058	9
	14		28998387	29000329	2	0	28709052	29055844	9
	25		18872032	NA	1	0	18872032	18922032	3
	26		22407530	23379414	23	2	21572162	23496516	5
NEFA	1		166669064	166888483	3	0	166406343	167009561	6
	1	**H**	185892360	186617146	25	15	184859013	187238015	41
	2		106012533	106052266	6	1	105664825	106542344	13
	7		86986401	87004808	3	0	86924655	87232954	1
	9		76549280	76571642	3	0	75789603	77130495	17
	15		66056425	NA	1	0	66056425	66106425	0
	17		13427110	14189583	14	1	12653835	14464765	6
	18		7685942	9565563	44	0	8293585	9790956	13
	19		48235446	NA	1	0	48235446	48285446	2
	24	**H**	20381260	20888104	2	1	20287835	20973401	16
	24		45325106	45675218	5	0	44139172	47064880	72
	30		6239856	6258423	5	0	5851204	6743672	9
	30	**H**	20974703	21044590	11	4	20915473	21380977	0
TG	1		126407798	127401777	6	0	126542590	128810519	46
	10		65383517	NA	1	0	65383517	65433517	1
	20		52368013	52589211	4	1	52145954	52997964	5
	21	**F**	49201984	49202284	2	0	48839667	49489807	2
Adiponectin	1		129650721	129653375	2	0	129419765	130122651	9
	1		138037003	NA	1	0	138037003	138087003	1
	2	**H**	16747148	17739125	38	27	16362904	18105119	42

Table 3.16: Prioritization of the GWA results of the Morgan horses based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
Adiponectin (con)	4	**H**	36557672	38544490	54	4	34723398	39321960	47
	6	**E H**	32601529	32727370	19	1	32486287	32841880	7
	6	**E H**	67997807	69847785	68	6	64297403	71493047	191
	7		21524454	21986901	14	0	19621101	22583950	53
	7		32963159	32963459	2	0	32448807	33202795	29
	8		3347264	3419299	6	0	2972877	3485969	18
	15		21830373	21834175	2	0	21702151	21904600	1
	15		66865469	66893151	4	0	66810113	66986537	2
	18	**F H**	41448414	NA	1	1	41448414	41498414	1
	18		49705278	49893633	7	0	48222088	50189162	36
	19		25833383	25859655	2	0	25269042	26285152	16
	20	**H**	3734902	3954772	12	0	3649052	4325872	11
	20		1882774	NA	1	0	1882774	1932774	2
	21	**F**	49478363	NA	1	0	49478363	49528363	1
Leptin	1		130957068	131062691	3	0	130419659	131677667	14
	4	**E**	52373692	52614368	22	0	51590680	52810437	9
	6		38446793	NA	1	0	38446793	38496793	3
	8		8682147	NA	1	0	8682147	8732147	2
	19	**H**	51360775	53132722	57	27	51286493	53959028	21
	21		16547954	16608200	3	0	14655783	16880737	21
	24	**H**	27275709	29038412	65	14	25564765	29384679	21
	25		27438558	27907420	14	2	26217071	29045128	65
ACTH	1	**E H**	83546191	83734040	17	4	82700933	84269783	24
	3	**D H**	43335201	44116411	13	0	42674448	44422013	10
	3	**H**	103056163	103438726	49	34	102944842	103801021	6
	5	**H**	25785666	27061038	32	10	25378878	27689002	28
	10		67992633	67997136	2	0	67173693	68509748	16

Table 3.16: Prioritization of the GWA results of the Morgan horses based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
ACTH (cont.)	10		70528773	NA	1	0	70528773	70578773	0
	11	**F**	18728679	18904099	4	0	17711712	19910206	80
	11		52897545	53669056	32	0	52809863	54320401	18
	13		25806289	NA	1	0	25806289	25856289	3
	16		31200001	NA	1	0	31200001	31250001	1
	18	**F**	41392781	NA	1	0	41392781	41442781	1
	20		29056288	NA	1	0	29056288	29106288	3
	21		11112604	NA	1	0	11112604	11162604	0
	21		24436227	24439739	3	0	23458912	25104737	16
	25		13299542	NA	1	0	13299542	13349542	0
	25		14989527	NA	1	0	14989527	15039527	1
	31		16965044	17737242	4	0	16852976	17943693	21
NH	1	**F**	78493587	79782621	37	0	78152399	80485573	21
	1	**E**	82958480	83232130	10	0	82097718	83618523	20
	2		93824111	93833011	2	0	93612698	93999072	2
	4	**E H**	52076906	53659651	149	110	52024470	54237747	20
	5		59796357	60233277	10	0	59986780	60283685	1
	5		65804297	65824216	3	0	65300990	66750795	4
	6	**E H**	64502443	65350057	44	12	60410647	70570773	172
	8		29756282	NA	1	0	29756282	29806282	0
	9		49062306	49078134	2	0	47678054	55125332	61
	14		74532493	NA	1	0	74532493	74582493	1
	18		2306238	NA	1	0	2306238	2356238	0
	19	**H**	1188889	1197320	3	2	661978	1345372	6
	19	**E H**	34421059	36247260	102	23	32962795	37391949	73
	19		46479290	47156982	5	1	46345791	47243745	16
	21		4745903	NA	1	0	4745903	4795903	0

Table 3.16: Prioritization of the GWA results of the Morgan horses based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
NH (cont.)	24		42026470	42450741	9	1	41516893	42504006	13
GH	1	**F**	79175507	79234421	5	0	79092549	79839480	7
	1		109778420	109819993	4	0	108645695	110793330	27
	1	**B H**	122383349	123036781	71	31	120644115	124691346	56
	2		85183513	86093522	20	10	84295572	88599903	36
	4		3032922	NA	1	1	3032922	3082922	1
	6	**F**	3139850	3453652	2	0	2318331	3601991	10
	6		6272129	6335115	8	0	6143412	6435255	8
	6		15729023	16202020	8	0	14200808	18578117	55
	7		26684853	26701040	6	0	26591740	26974624	18
	8		63557829	NA	1	0	63557829	63607829	0
	17	**H**	32020513	33031579	39	2	31806060	33720086	7
	18		2423391	NA	1	0	2423391	2473391	1
	22		45719751	48733979	2	0	48299638	49204093	12
	29		19108245	19432974	7	2	19255639	19488768	3
LAM	2		66192812	NA	1	1	66192812	66242812	1
	3		3294278	NA	1	0	3294278	3344278	2
	4	**E H**	17509325	19295909	52	4	17301415	19812653	16
	6	**F**	3466933	NA	1	0	3466933	3516933	1
	6		79661858	NA	1	0	79661858	79711858	1
	8		59199626	60121756	24	0	59149588	60266527	9
	12	**H**	33127411	34414133	53	27	32885278	34800986	45
	14		66311023	66688404	15	0	65014422	67256851	8
	14	**H**	88975206	90135630	48	9	87916190	91602875	58
	16		64556111	NA	1	0	64556111	64606111	0
	16		74667638	NA	1	0	74667638	74717638	2
	18	**H**	31710749	33317633	65	33	30095266	35177011	36

Table 3.16: Prioritization of the GWA results of the Morgan horses based on LD-defined regions (cont.)									
Trait	**Chr**	**Summary**	**Min_SNP**	**Max_SNP**	**Sugg_SNPs**	**Sign_SNPs**	**Min_Region**	**Max_Region**	**Total_Genes**
LAM (cont.)	19	**H**	30133826	NA	51	3	30133826	30183826	2
	22	**B H**	3616445	4853827	75	45	2843476	5225020	23
	22		13852015	NA	1	0	13852015	13902015	0
	22		23806850	NA	1	0	23806850	23856850	1
	23	**H**	11116499	12515439	51	46	7656404	12984095	34
	31		6804894	NA	1	0	6804894	6854894	1

Table 3.16: Prioritization of the GWA results of the Morgan horses based on LD-defined regions. To be considered an ROI, at least five SNP had to exceed the suggestive threshold (1.0e-05) with one SNP exceeding the threshold for genome-wide significance (7.61e-08). Provided in the table is the base pair position of the lowest (Min_SNP) and highest (Max_SNP) SNP in the region, the min (Min_Region) and maximum (Max_Region) boundaries of the region based on LD, as well as the number of SNPs per region which exceeded the suggestive (Sugg_SNPs) and genome-wide significance (Sign_SNPs) threshold. The total number of genes includes all protein-coding genes, pseudogenes, and RNA genes based on EquCab3. Black box in the summary column indicates the region did not meet the criteria to be considered an ROI, was not significant on metanalysis and was not shared with another or trait. Letters in the summary column represent: (B) region was shared with one or more Welsh pony cohorts and at least one region was considered an ROI, (D) region was shared with one or more Welsh pony cohorts but no regions met the criteria for an ROI, (E) region was shared with another trait in the Morgan horses and at least one region was considered an ROI, (F) region was shared with another trait in the Morgan horses but no regions met the criteria for an ROI, (H) region was identified as shared across breeds on metanalysis and was considered an MA-ROI. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 3.17: High Priority Regions in Morgan horses based on LD-defined regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	24	21134897	21184897	1	0	0	1
Insulin-OST	4	28373202	28423202	0	0	0	0
	6	32751552	34029749	12	0	10	22
	10	71666607	73534053	6	1	5	12
Glucose	4	17239374	19043831	6	0	5	11
	8	11193683	12404572	8	0	9	17
Glucose-OST	3	55746338	58085997	15	0	6	21
	4	26695616	29116058	5	0	4	9
NEFA	1	184859013	187238015	24	0	17	41
	17	12653835	14464765	4	0	2	6
	24	20287835	20973401	15	0	1	16
	30	20915473	21380977	0	0	0	0
Adiponectin	2	16362904	18105119	21	0	21	42
	4	34723398	39321960	36	1	10	47
	6	32486287	32841880	3	0	4	7
	6	64297403	71493047	168	1	22	191
	18	41448414	41498414	0	0	1	1
	20	3649052	4325872	8	0	3	11
Leptin	4	51590680	52810437	4	0	5	9
	19	51286493	53959028	7	0	14	21
	24	25564765	29384679	7	0	14	21
ACTH	1	82700933	84269783	18	1	5	24
	3	42674448	44422013	2	1	7	10
	3	102944842	103801021	2	0	4	6
	5	25378878	27689002	12	0	16	28
NH	1	82097718	83618523	14	0	6	20
	4	52024470	54237747	8	0	12	20

Table 3.17: High Priority Regions in Morgan horses based on LD-defined regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
NH (cont.)	6	60410647	70570773	144	2	26	172
	19	661978	1345372	4	1	1	6
	19	32962795	37391949	53	1	19	73
GH	1	120644115	124691346	36	0	20	56
	17	31806060	33720086	3	1	3	7
LAM	4	17301415	19812653	8	1	7	16
	12	32885278	34800986	29	0	16	45
	14	87916190	91602875	32	0	26	58
	18	30095266	35177011	23	0	13	36
	19	30133826	30183826	2	0	0	2
	22	2843476	5225020	13	0	10	23
	23	7656404	12984095	11	0	23	34
Total				**764**	**11**	**367**	**1142**

Medium Priority Regions in Morgan horses based on LD-defined regions							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin	2	117310352	117579882	2	0	2	4
Glucose_OST	26	21572162	23496516	0	0	5	5
NEFA	2	105664825	106542344	10	0	3	13
Leptin	25	26217071	29045128	59	2	4	65
NH	19	46345791	47243745	15	0	1	16
	24	41516893	42504006	8	0	5	13
GH	2	84295572	88599903	22	0	14	36
	29	19255639	19488768	3	0	0	3
Total				**119**	**2**	**34**	**155**

Table 3.17: Low Priority Regions in Morgan horses based on LD-defined regions (cont.)							
Trait	**Chr**	**Min_Region**	**Max_Region**	**Protein_Coding**	**Pseudogenes**	**RNA_Genes**	**Total_Genes**
Insulin_OST	11	18355073	19629302	50	1	2	53
TG	21	48839667	49489807	1	0	1	2
Adiponectin	21	49478363	49528363	1	0	0	1
ACTH	11	17711712	19910206	67	1	12	80
	18	41392781	41442781	0	0	1	1
NH	1	78152399	80485573	9	0	12	21
GH	1	79092549	79839480	2	0	5	7
	6	2318331	3601991	5	0	5	10
LAM	6	3466933	3516933	0	0	1	1
Total				**135**	**2**	**39**	**176**

Table 3.17: Final boundaries of the regions based on LD and positional candidate genes of the prioritization GWA results for the Morgan horses. Regions were categorized as high priority (regions found on metanalysis OR region was shared with another trait), medium priority (region was an ROI in the Morgan horses but was not shared), or low priority (region was shared across traits but region was not an ROI). Final region boundaries of the region were based on LD-ROI and are indicated by the lowest base pair position (Min_ROI) and the highest base pair position (Max_ROI). The total number of genes includes all protein-coding genes, pseudogenes, and RNA genes identified for region based on EquCab3. Shared regions across prioritized traits are indicated by highlighted chromosomes.

Chapter 4: Fine-mapping high priority LD-ROI from genome-wide association analyses using imputed whole genome sequencing

Summary: Fine-mapping of genome-wide associated regions seeks to refine the genomic localization of causal variants through statistical analyses, bioinformatics, or functional methods, with the main goal of differentiating between the causal variants and those merely correlated with the causal variant**.** The objectives of this chapter were to fine-map selected high priority ROI by increasing the marker density for association analysis and to interrogate positional candidate genes for putative functional alleles based on predicted impact from variant annotation and biological effect based on evidence in other species. We utilized imputed whole-genome sequencing (WGS) and linear regression analysis in order to fine-map selected high priority LD-ROI in both the Morgan horses and Welsh ponies. Five fine-mapped regions from each breed were further interrogated for predicted impact using variant annotation. All variants which exceeded the threshold for genome-wide significance mapped to non-coding regions of the genome, with 66.7% of the significantly associated SNPs being intronic, 17.0% intergenic and 10.3% within lncRNA. We further evaluated positional candidate genes with exonic variants in our fine-mapped region with a p-value <0.05 (i.e. "sub-threshold"). Protein-coding genes containing non-coding or coding variants within the fine-mapping region were then further prioritized based on known function and biological evidence in other species utilizing the PubMed search engine. A total of 19 positional candidate genes were identified as having biological evidence for a role in EMS including: *SSTR1, SEC23A, FBXO33, MIA2, EIF3D, CSF2B, IFT27, ACE, TACO1, ABCA13, NKAIN2, BBX, XXYLT1, BDH1, NCKAP5L, GPD1, LIAA1, METTL7A, SCL11A2.* These data provide intriguing biological evidence for the role of several coding genes in the pathogenesis of EMS.

Introduction:

Equine metabolic syndrome (EMS) is a clustering of risk factors leading to laminitis with the key component being insulin dysregulation, manifesting as baseline hyperinsulinemia, an exaggerated or prolonged insulin or glucose response post carbohydrate challenge, tissue insulin resistance, or dyslipidemia [21]. In chapter 2, we provided the first concrete evidence that EMS is a complex trait with a strong genetic basis. In chapter 3, we utilized genome-wide association analyses (GWA) with single nucleotide polymorphism (SNP) genotype data to identify high priority regions of interest (ROI) harboring the risk alleles associated with EMS phenotypes in both Welsh ponies and Morgan horses, two breeds considered high risk for this syndrome. Although these results provided valuable information about the genetic architecture of EMS by identifying the number of loci contributing to EMS, estimating their effect size and allele frequencies, and providing evidence that risk alleles are both shared and unique across breeds, they do not identify the specific risk alleles contributing to EMS.

Fine-mapping of GWA regions seeks to refine the genomic localization of causal variants through statistical analyses, bioinformatics, or functional methods, with the main goal of differentiating between the causal variants and those merely correlated with the causal variant [457]. One of the key principles of fine-mapping is that all variants within the region must be represented in order to capture the causal variant. Thus, increasing marker density is essential for accurately fine-mapping a region with genotype imputation being considered a cost-effective and precise method to achieve this goal [457,458].

The first objective of this chapter was to fine-map selected high priority ROI identified from GWA in chapter 3 by increasing the marker density for association analysis within these regions through the use of imputed whole genome sequencing (WGS). The second objective was to interrogate positional candidate genes for putative functional alleles based on predicted impact from variant annotation and biological effect based on evidence in other species.

Material and Methods

Samples: Horses used in this study were a part of a large, across breeds study evaluating the EMS phenotype [20]. From this dataset, 264 Welsh ponies (194 females and 70 males with a mean age of 11.7 years) and 286 Morgan horses (184 females and 102 males with a mean age of 12.3 years) were included in this analysis. Samples were collected from 31 and 28 farms throughout the United States and Canada for the Morgan horses and Welsh ponies, respectively.

Phenotype data collected on all horses included: signalment, medical history, laminitis status, environmental management (feed, supplements, turnout and exercise regimen), and morphometric measurements (body condition score (BCS), wither height, and neck and girth circumference). After an eight hour fast, an oral sugar test (OST) was performed using 0.15mg/kg Karo lite corn syrup as previously described [406]. Biochemical measurements at baseline included insulin, glucose, non-esterified fatty acids (NEFA), triglycerides (TG), adiponectin, leptin and ACTH. Biochemical measurements 75 minutes after the OST included insulin (INS-OST) and glucose (GLU-OST).

For inclusion in the study, each farm had to have at least one control and one horse with clinical signs consistent with EMS (including horses with regional adiposity, hyperinsulinemia or an exaggerated response to the OST, elevations in TG, and decreased levels of adiponectin at the time of sampling) under the same management. Horses with a history or phenotypic appearance of pars pituitary intermedia dysfunction (PPID) were excluded from the study. The category of previously laminitic was defined as a horse who had been diagnosed with pasture-associated or endocrinopathic laminitis by a veterinarian, had radiographic evidence of laminitis, or had laminitis rings or clinical signs of laminitis as the time of sampling. Horses in which laminitis could have been caused by another inciting factor (history of illness, grain-overload, corticosteroid administration or PPID), or who were clinical for acute laminitis at the time of sampling, were also excluded from the study.

Whole Genome Sequencing: DNA was isolated from whole blood or hair roots using the Puregene Blood Core Kit, (Qiagen) per manufacturer's instructions. 19 Welsh ponies and 18 Morgan horses, representing both insulin sensitive and insulin dysregulated horses (individuals with a history of laminitis and at the breed-specific upper range for the EMS

biochemical measurements baseline insulin, insulin concentrations post oral sugar challenge, or triglycerides, or breed-specific lower range for adiponectin concentrations) were chosen for WGS (**Tables 4.1 and 4.2**). DNA samples were submitted for WGS at the University of Minnesota Genomics Center using an Ilumina HighSeq 2500 sequencer for 100 base pair paired-end sequencing, with an average read coverage of 6-12X over the 2.7Gb of the equine genome per sample. Quality control, processing and mapping of reads to EquCab3, the reference genome at the time of this analysis, was performed using the PALEOMIX pipeline [459]. Briefly, this pipeline utilizes the AdapterRemoval software tool for initial quality control and processing of the raw reads, Burrows Wheeler alignment (BWA-MEM) software to map processed reads to the reference genome, SAMtools and Picard tools for quality and duplicate filtering, and GATK's Indel Realigner for local re-alignment around small insertions and deletions [459]. Insertions, deletions, and SNPs were called for both breeds using multi-sample variant detection in three variant calling software programs: Platypus [460], HaplotypeCaller [461], and Samtools [462]. Variants were filtered to remove those called exclusively in one program, yielding a total of 19,722,966 variants in the final VCF.

SNP Genotype Data: Genome-wide single nucleotide polymorphism (SNP) genotyping was performed with horses genotyped either on the Illumina EquineSNP50 BeadChip (54K array) containing 54,602 SNPs, the Axiom Equine MCEc670 array (670K array) containing 670,795 SNPs, or the Axiom Equine MCEc2M array (2M array) containing 2,011,826 SNPs across the equine genome including the 31 autosomes and X chromosomes. For the Morgan horses, 40 individuals were genotyped on the 2M array and 246 individuals were genotyped on the 54K array of which 18 horses also had WGS. For the Welsh ponies, 44 ponies were genotyped on the 2M array, of which 6 ponies also had WGS, and 220 ponies were genotyped on the 670K array, of which 12 were also sequenced for WGS. Base pair locations for all SNP data were remapped from EquCab2 to EquCab3 using the script from Beeson *et. al.* [427].

Imputation to WGS: Haplotype phasing and genotype imputation up to WGS was performed on horses genotyped on the SNP arrays using Beagle software [407]. Based on published recommendations, WGS data from the 19 Welsh ponies (one horse with

suspected PPID was included in the reference population for imputation but removed for subsequent analyses) and 18 Morgan horses were combined for a cross breed reference population [408]. Prior to phasing, quality control was performed to remove variants with a genotyping rate of <90%, non-biallelic SNPs, insertions and deletions, and variants mapping to chromosomes unknown yielding a total of 16,056,906 variants in the reference population. Within breed imputation of data from the 2M arrays were performed followed by imputation of the 670K array and then the 54K array, using the following protocol (see **Figure 4.1**):

(i) For horses with both WGS and genotype data from the SNP array being imputed (test file), 5 (2M array) or 10 individuals (lower density arrays) were randomly chosen to be removed from the reference file in order to test for concordance. The remaining duplicated individuals were excluded from the test file.

(ii) For SNPs arrays in which individuals did not have concurrent WGS, 5 (2M array) or 10 individuals (lower density arrays) were randomly selected from the WGS data to be masked down to variants present on the SNP array.

(iii) Imputation of the test file was completed using Beagle software [407].

(iv) Imputation concordance was calculated as the percentage of calls that matched between the reference and imputed files.

(v) The test and reference files were merged together and all SNPs which did not achieve 100% concordance were removed from the final reference file. At this point, the reference file contained the WGS and imputed data which was then used for imputation of the next SNP array.

After imputation of all three genotyping arrays, the final file contained 264 Welsh ponies and 286 Morgan horses with a total of 12,787,473 variants. This file was parsed by breed and quality control was performed to remove non-informative SNPs and SNPs with a minor allele frequency <0.01, yielding 6,098,487 SNPs in the Morgan horses and 6,695,837 in the Welsh ponies. Variants for each breed were annotated using the software program SnpEff [463].

Fine Mapping ROI: Selected high priority ROI in the Welsh ponies and Morgan horses identified in chapter 3 were chosen for fine mapping with boundaries of the ROI defined

by the breakdown of linkage disequilibrium (LD-ROI; see chapter 3 for further details). LD-ROI were selected for additional fine-mapping if the region contained at least 5 SNPs with one SNP exceeding the threshold for genome-wide significance providing an initial list of fine-mapped LD-ROI (**Tables 4.3 and 4.4**). Linear regression analysis was performed with the EMS phenotypic trait as the outcome variable, the subset of imputed WGS SNPs for the LD-ROI as the response variables, sex and age as fixed effects, and farm as a random effect. For these analyses, eleven traits significantly associated with EMS [20] were treated as quantitative response variables in this including: insulin, glucose, adiponectin, leptin, NEFA, TG, ACTH, insulin-OST, glucose-OST, and measures of obesity (neck circumference to whither height ratio (NH), and girth circumference to whither height ratio (GH). Laminitis status was treated as a binary response variable. Threshold for significance was determined using the Genetic Type 1 Error Calculator [434], which calculated a Bonferroni corrected p-value based on the effective number of SNPs for the ROI (**Tables 4.3 and 4.4**).

Based on the results of the linear regression, five LD-ROI for the Welsh ponies and five LD-ROI for the Morgan horses were chosen for further analysis (**Table 4.5**). Regions were chosen for this additional analysis if the region contained one or more SNPs that exceeded the threshold for genome-wide significance, appeared to have clear delineation between baseline SNPs and those that exceeded the threshold, and regions where significant SNPs appeared to be tightly clustered. Regions were fine-mapped based on the base pair position of SNPs which exceeded the threshold for genome-wide significance and variants in the region were interrogated for predicted effect based on the results of variant annotation. Positional candidate genes were identified using the Ensembl genome browser with EquCab3 as the reference genome and investigated for predicted biological effect and evidence across species by performing a literature search using the PubMed search engine with the gene identifier and key words: obesity, metabolic, metabolism, diabetes, fat, and the EMS trait of interest.

Results

WGS: EMS phenotypic data for Welsh ponies and Morgan horses selected for whole genome sequencing is presented in **Tables 4.1 and 4.2**, respectively. In the Welsh ponies,

9 horses had a history of laminitis, 4 had insulin concentrations above one-standard deviation from the breed-specific mean (10.9µIU/mL; SD: 17.2), 7 had INS-OST concentrations above one-standard deviation from the breed-specific mean (36.2µIU/mL; SD: 45.4), 5 had triglyceride concentrations above one-standard deviation from the breed-specific mean (28.0.mg/dL; SD: 34.8), and 6 horses had adiponectin concentrations less than one standard deviation from the breed-specific mean (6.9µg/mL; SD: 5.5). One laminitic pony met the criteria for all four biochemical measurements, three laminitic ponies met three of the four criteria, three laminitic ponies met two of the four criteria, and one non-laminitic pony met two of the four criteria for biochemical measurements. Notably, all 9 laminitic horses would have been diagnosed with insulin dysregulation based on published criteria of insulin concentrations >45µIU/mL post oral sugar test [21] (**Table 4.1**).

In the Morgan horses, seven horses had a history of laminitis, seven had insulin concentrations above one-standard deviation from the breed-specific mean (8.5µIU/mL; SD: 8.9), five had insulin-OST concentrations above one-standard deviation from the breed-specific mean (33.2µIU/mL; SD: 59.7), three had triglyceride concentrations above one-standard deviation from the breed-specific mean (25.0.mg/dL; SD: 16.4), and nine horses had adiponectin concentrations less than one standard deviation from the breed-specific mean (5.0µg/mL; SD: 3.0). One laminitic horse met the criteria for all four biochemical measurements, four laminitic horses met three of the four criteria, and three laminitic horses met two of the four criteria for biochemical measurements. Notably, all seven laminitic horses and one non-laminitic horse would have been diagnosed with insulin dysregulation based on published criteria of insulin concentrations >45µU/mL post oral sugar test [21] (**Table 4.2**).

The number of WGS reads, average read length, number unique reads which mapped to nuclear genome, and average sequencing depth for each individual with WGS is presented in **Appendix C: Supplemental Table C1 and Supplemental Table C2**.

Variant Annotation for Imputed WGS: In the Welsh ponies, 6,695,837 variants were annotated using the software program SnpEff. On average, there was one variant every 359 bases, with a genome effective length of 2.4 billion base pairs, and a

transitions/transversions ratio of 2.09. The largest percentage of variants were identified in introns (65.2%) or intergenic (19.9%) regions of the genome, and 1.4% of the variants were identified in exons. The majority of the variants were predicted to be modifiers (99.11%), followed by those with a low (0.56%), moderate (0.32%) or high impact (0.008%). Of those with a predicted impact, 56.7% were called as silent mutations, 43% missense mutations, and 0.28% nonsense mutations.

In the Morgan horses, 6,098,487 variants were annotated using the software program SnpEff. On average, there was 1 variant every 395 bases, with a genome effective length of 2.4 billion base pairs, and a transitions/transversions ratio of 2.08. The largest percentage of variants were identified in introns (65.4%) or intergenic (20%) regions of the genome, and 1.4% of the variants were identified in exons. The majority of the variants were predicted to be modifiers (99.13%), followed by those with a low (0.55%), moderate (0.32%) or high impact (0.007%). Of those with a predicted impact, 57.1% were called silent mutations, 42.6% missense mutations, and 0.28% nonsense mutations.

Fine Mapping Welsh Pony ROI: In the Welsh ponies, 41 LD-ROI were fine-mapped with the imputed WGS with the results summarized in **Table 4.3**. Of these regions, 26 included SNPs which exceeded the threshold for genome-wide significance on regression analysis and five were chosen for additional follow-up **Table 4.5**.

1. *Adiponectin concentrations on chromosome (ECA) 1*: This region spanned ~6.4 megabases (Mb) and included 20,024 SNPs (3,837 effective SNPs) from base pair positions 171,861,236 to 178,270,042. Of these SNPs, 117 exceeded the calculated threshold for genome-wide significance of <1.3e-05 and were between base pair positions 175,155,905 to 177,764,563. Within the fine mapped region of 175,000,000 to 178,000,000 base pairs, there were 14 long noncoding RNAs (lncRNAs), 1 small nuclear RNA (snRNA), and 16 protein coding genes. Significantly associated SNPs were either intergenic (35.9%), intronic, intragenic, upstream or downstream from 3 pseudogenes, a lncRNA (LOC111769213) or the protein coding genes: paired box 9 (*PAX9*), solute carrier family 25 member 2 (*SLC25A2*), tetratricopeptide repeat domain 6 (*TTC6*), somatostatin receptor 1 (*SSTR1*), SEC23 homolog A/coat complex II component (*SEC23A*), gem nuclear organelle associated protein 2 (*GEMIN2*), MIA

SH3 domain ER export factor 2 (*MIA2*), and F-box protein 33 (*FBX033*) (**Figure 4.2** and **Table 4.6**). A total of 20 low impact, 26 moderate impact, and one high impact variant were identified in the fine-mapped region, of which 16 had p-values <0.05 and are listed in **Table 4.6.**

2. *Leptin concentrations on ECA7*: This region spanned ~1.4Mb and included 7,144 SNPs (3,966 effective SNPs) from base pair positions 65,678,376 to 68,117,086. Of these SNPs, 46 exceeded the calculated threshold for genome-wide significance of <1.3e-05 and were between base pair positions 67,940,623 to 67,971,228. All significantly associated SNPs were located in the introns of the protein coding gene teneurin transmembrane 4 (*TENM4*) (**Figure 4.3** and **Table 4.7**). Seven variants with a predicted low impact were identified in the *TENM4* gene, of which one SNP had a p-value of <0.05 and is listed in **Table 4.7**.
3. *GLU-OST and NEFA concentrations on ECA28*: The shared region between GLU-OST and NEFA spanned ~866.8 kilobases (kb) and included the entire region for GLU-OST, which included 2,036 SNPs (415 effective SNPs) from base pair positions 34,271,949 to 35,138,699. Of these, 85 SNPs exceeded the threshold for genome wide significance of 1.2e-04 and were between base pair positions 34,368,987 to 35,136,611. Significantly associated SNPs were either intergenic (4.7%), or intronic, upstream, downstream, or within the 3'UTR of the protein coding genes: RNA binding fox-1 homolog 2 (*RBFOX2),* FAD dependent oxidoreductase domain containing 2 (*FOXRED2),* eukaryotic translational initiation factor 3 subunit D (*EIF3D),* or calcium voltage-gated channel auxiliary subunit gamma 2 (*CACNG2).* Within the fine-mapped region of 34,350,000 to 35,007,000 there were nine protein coding gene, one lncRNA, and two pseudogenes (**Figure 4.4** and **Table 4.8)**. A total of four low impact and one moderate impact variants were identified in the fine-mapped region, of which one low impact variant within the lncRNA (LOC11177109) had p-values <0.05 and is listed in **Table 4.8.**

   For NEFA concentrations, this region spanned ~2.8Mb and included 6,227 SNPs (1,282 effective SNPs) from base pair positions 32,90,9542 to 35,703,535. Of these, 30 SNPs exceeded the calculated threshold for genome-wide significance of <3.9e-05 and were between base pair positions 33,271,314 to 35,288,863. Based on the fine-

mapped region for GLU-OST, the four significant SNP between base pair positions 3,3271,314 to 3,4102,222 were excluded. Statistically associated SNPs were intergenic (19.2%), intronic or downstream of the protein coding genes: *RBFOX2* and neutrophil cytosolic factor 4 (*NCF4)* (**Figure 4.5** and **Table 4.9**). Within the fine-mapped region between base pairs 34,050,000 to 35,350,000, there were nine low impact, six moderate impact, and one low impact variants, of which 10 had a p-value of <0.05 and are listed in **Table 4.9.**

4. *GH ratio on ECA11*: This region spanned ~1.0Mb and included 20,156 SNPs (1,077 effective SNPs) from base pair positions 15,414,337 to 16,451,463. Of these SNPs, 19 exceeded the calculated threshold for genome-wide significance of <4.6e-05 and were between base pair positions 15,700,606 to 16,403,791. Statistically significant SNPs were within the intron, 3' UTR, upstream, or downstream from the protein coding genes: mitogen-activated protein kinase kinase kinase 3 (*MAP3K3*), translational activator of cytochrome C oxidase 1 (*TACO1*), DDBI and CUL4 associated factor 7 (*DCAF7*), angiotensin I converting enzyme (*ACE*), and membrane associated ring-CH-type finger 10 (*MARCH10*) (**Figure 4.6** and **Table 4.10**). Within the fine-mapped region between base pair positions 15,680,000 to 16,430,000, there were 10 protein coding genes and one lncRNA. A total of nine variants within this region were predicted to have a moderate impact and 11 variants were predicted to have a low impact, of which six SNPs had a p-value of <0.05 and are listed in **Table 4.10**.
5. *ACTH concentrations on ECA 1:* This region spanned ~1.4Mb and included 3,940 SNPs (652 effective SNPs) from base pair positions 69,558,737 to 70,960,589. Of these SNPs, nine exceeded the calculated threshold for genome-wide significance of <7.6e-05 and were between base pair positions 70,264,921 to 70,272,614. All significantly associated SNPs were intergenic (**Figure 4.7** and **Table 4.11**). Within the LD-ROI, there were nine lncRNA, two rRNA, and seven protein coding genes. A total of four variants within this region were predicted to have a moderate impact and three variants were predicted to have a low impact, of which zero variants had a p-value of <0.05.

Fine Mapping Morgan ROI: In the Morgan horses, 25 LD-ROI were fine mapped with the imputed WGS with the results summarized in **Table 4.4**. Of these regions, 18 included

SNPs which exceeded the threshold for genome-wide significance on regression analysis and five were chosen for additional follow-up (**Table 4.5**).

1. *Laminitis status and basal glucose concentrations on ECA4*: The shared region between laminitis status and basal glucose concentrations spanned ~1.7Mb from base pair positions 17,301,415 to 19,043,831. For laminitis status, this region spanned ~2.5 Mb and included 9,251 SNPs (5,032 effective SNPs) from base pair positions 17,301,415 to 19,812,653. Of these, 25 SNPs exceeded the calculated threshold for genome wide significance of <9.94e-06 and were between base pair positions 17,436,000 to 18,650,000. Within this region of the genome, there were five protein coding genes, two lncRNA, and one snRNA. Significantly associated SNPs were intergenic (2.7%) or located within the introns of the protein-coding genes: tensin3 (*TNS3*), polycystin 1 like 1 (*PDK1L1*), HUS1 checkpoint clamp component (*HUS1),* or ATP-binding cassette transporter A13 (*ABCA13*) (**Figure 4.8** and **Table 4.12**).

   For basal glucose concentrations, this region spanned ~1.8Mb and contained 6,495 SNPs (3,156 effective SNPs) from base pair position 17,239,374 to 19,043,831. Of these SNPs, one SNP exceeded the calculated threshold for genome-wide significance of <1.6e-05 and was located within intron of the protein-coding gene *ABCA13* (**Figure 4.9** and **Table 4.13)**. A total of 15 low impact and 15 moderate impact variants were identified in the *ABCA13* gene, of which three had p-values <0.05 and are listed in **Table 4.13.**
2. *INS-OST on ECA10*: This region spanned ~1.9Mb and included 4,784 SNPs (1,827 effective SNPs) from base pair positions 71,666,607 to 73,534,053. Of these SNPs, 18 exceeded the calculated threshold for genome-wide significance of <2.74e-05 and were between base pair positions 72,939,355 to 72,945,989. All significantly associated SNPs were located within a single intron of the protein coding gene sodium/potassium transporting ATPase interacting 2 (*NKAIN2*) (**Figure 4.10** and **Table 4.14**). No variant with a predicted low, moderate, or high impact were identified in the *NKAIN2* gene.
3. *Leptin concentrations on ECA19*: This region spanned ~2.6Mb and included 8,454 SNPs (3,631 effective SNPs) from base pair positions 51,386,493 to 53,959,028. Of these SNPs, 35 exceeded the calculated threshold for genome-wide significance of <1.3e-05 and were between base pair positions 51,727,537 to 52,286,046. Within the

fine mapped region of 51,580,000 to 52,450,000 base pairs, there were seven pseudogenes, two lncRNAs and two protein coding genes. Significantly associated SNPs were either intergenic (34.3%), intronic, intragenic, or downstream from 2 pseudogenes, a lncRNA (LOC111769112) or the protein coding gene Bobby sox homolog (*BBX)* (**Figure 4.11** and **Table 4.15**). No exonic variants were identified in the protein coding genes within the fine-mapped region.

4. *NH concentrations on ECA19*: This region spanned ~4.4Mb and included 10,479 SNPs (5,321 effective SNPs) from base pair positions 32,962,795 to 37,391,949. Of these SNPs, 16 exceeded the calculated threshold for genome-wide significance of <9.4e-06 and were between base pair positions 33,7007,27 to 35,859,978. Within the fine mapped region of 33,400,000 to 35,900,000 base pairs, there were eleven lncRNAs, one miscRNA and 31 protein coding genes. Significantly associated SNPs were either intergenic, intronic, intragenic, or upstream from one lncRNA (LOC111769074) or the protein coding genes: leucine rich repeat containing 15 (*LRRC15)*, ATPase 13A3 (*ATP13A3),* and xyloside xylosyltransferase 1 (*XXYLT1*) (**Figure 4.12** and **Table 4.16**). Within the fine-mapped region, 46 exonic or splice site variants were identified in protein-coding genes, of which 30 were predicted have a low impact and 16 were predicted to have a moderate impact; 12 of these variants had p-values <0.05 and are listed in **Table 4.16**.
5. *Adiponectin concentrations on ECA6:* This region spanned ~7.2Mb and included 17,106 SNPs (8,812 effective SNPs) from base pair positions 64,297,403 to 71,493,047. One intergenic SNP exceeded the calculated threshold for genome-wide significance of <5.7e-06 (**Figure 4.13** and **Table 4.17**). Within the fine mapped region of 68,250,000 to 70,001,000 base pairs, there were 38 protein coding genes and one lncRNA; 48 exon or splice site variates were identified in protein-coding genes, of which 46 were predicted have a low impact, 21 were predicted to have a moderate impact, and one was predicted to have a high impact; 20 of these variants had p-values <0.05 and are listed in **Table 4.17**.

Discussion

We fine-mapped ten regions of interest identified on GWA using imputed WGS and linear regression analysis. Within these regions, variants were identified within several positional candidate genes which were further prioritized based on known functional effect and biological evidence in other species utilizing the PubMed search engine.

Biologic Evidence for Candidate Genes.

Adiponectin concentrations on ECA1 in Welsh ponies: This region contained a number of protein-coding genes with either intronic or exonic variants (**Table 4.6)**. *SSTR1* is a G protein-coupled membrane receptor for somatostatin which has been shown to be differentially expressed in the presence or absence of adipocytokines [464]. Somatostatin is neuropeptide which is primarily known for its role in inhibiting the secretion of hormones such as insulin, growth hormone, glucagon, and cortisol, but it has also been shown to inhibit circulating levels of adiponectin and leptin in lean, healthy males [465]. Interestingly, infusions of somatostatin decreased circulating levels of adiponectin in obsess individuals while leptin levels remained unchanged [466]. Although the mechanism underlying somatostatin's inhibitory effect on adiponectin is unknown, activation of somatostatin receptors has been proposed [466]. In our cohort, 1 genome-wide significant variant was identified downstream of *SSTR1* (**Table 4.6**). *SEC23A* encodes a component of the coat protein complex II-coated vesicles that transports secretory proteins from the endoplasmic reticulum to the Golgi apparatus. In GWA, a variant in *SEC23A* was found to be associated with decreased vitamin D concentrations in a population of ~79,000 individuals of European ancestry [467]. Interestingly, numerous studies have suggested a role for low vitamin D3 levels in the development of MetS although the exact mechanism is unknown [468-471]. Additionally, vitamin D has been correlated with adipokines and it was suggested that there is a connected mechanism between vitamin D and adiponectin binding proteins [472]. In the Welsh ponies, 5 genome-wide significant variants were identified in the intron of *SEC23A* (**Table 4.6**). Variants in *FBXO33* were associated with concentrations of advanced glycation end-products, which have been found to elevated in individuals with hyperglycemia and diabetic complications [473]. In the Welsh ponies 11 intronic variants were identified which exceed the threshold for genome-wide significance (**Table 4.6**). *MIA2* was identified as part of a chimeric protein TANGO1-like (TAL1)

which interacts with apolipoprotein B (ApoB), a protein involved in the metabolism of lipids (see chapter 1), and TAL1 is required for the recruitment of ApoB-containing lipid particles to the endoplasmic reticulum [474]. However, no direct association with adiponectin, obesity or metabolic syndrome has been identified for *MIA2*. In our cohort, 17 intronic, 8 downstream, and one variant in the 5' UTR which was predicted to cause the gain of a start codon (low impact) exceeded the threshold for genome wide significance for *MIA2*. In addition, one missense mutation (moderate impact) was identified in exon 3 of *MIA2* (p-value=1.54e-02) (**Table 4.6**).

Leptin concentrations on ECA7 in Welsh ponies: Within this region, significantly associated SNPs were located with the intron of *TENM4*. *TENM4* encodes a gene that has an essential role in establishing proper neuronal connectivity during development and a causal variant for essential tremors has been identified [475]. *TENM4* has been associated with Schizophrenia [476] and prostate cancer [477]. A role of *TENM* in leptin, fatty acid metabolism, obesity, diabetes, or metabolic syndrome has not been established. In the Welsh ponies, 46 *TENM* intronic variants exceeded the threshold for genome-wide significance, and a splice site variant (low impact) was identified between exon 3 and 4 with a p-value of 4.06e-03 (**Table 4.7**).

GLU-OST and NEFA concentrations on ECA28 in Welsh ponies: Based on the boundaries of the LD-ROI, this region on ECA28 was identified as shared between GLU-OST and NEFA concentrations. Given the more precisely fine-mapped region for GLU-OST, the region between 34,350,000 to 35,007,000 was further evaluated. *EIF3D* encodes the largest-subunit of one of the most complex translation initiation factors and is required for the initiation of protein synthesis of several mRNA via assistance in the recruitment of ribosomes to the mRNA [478]. Using transcriptomics, proteomics and metabolomic profiling, it was identified that cells lacking EIF3D were unable to synthesize components of the mitochondrial electron transport chain, leading to a shift in energy balance with increased glucose uptake, upregulation of glycolytic enzymes, and fermentation of carbon sources, suggesting a role for EIF3D in glucose metabolism [479]. Further, in models of gall bladder cancer, EIF3D was shown to activate PI3K/AKT signaling by blocking the degradation of the G-protein coupled receptor kinase 2 (GRK2). As discussed in chapter

1, the PI3K/AKT signaling pathway is essential for glucose homeostasis and lipid metabolism [480], and GRK2 has been suggested to have a relevant role in insulin resistance and obesity [481]; providing evidence for a role for *EIF3D* in both glucose homeostasis and fatty acid metabolism. In the Welsh ponies, 8 intronic SNPs in *EIF3D* exceeded the threshold for genome-wide significance for GLU-OST concentrations (**Table 4.8).** In addition, protein coding genes with a predicted functional impact in the association analysis for NEFA concentrations (**Table 4.9**) include: one splice site variant in intron 3-4 and one synonymous variant in exon 14 (p-values=2.43e-02; both with a predicted low impact) of *CSF2B* which was previously found to be correlated with BMI and upregulated in the subcutaneous white adipose tissue of obese individuals [482], and a missense variant (p-value=7.13e-03; predicted moderate impact) in exon 2 of *IFT27* which encodes a protein that is a core component of the intraflagellar transport, and mouse knockout models of this gene results in an obese phenotype secondary to alterations in ciliary function [483].

GH ratio on ECA11 in Welsh ponies: ACE encodes a zinc metallopeptidase which is involved in the conversion of angiotensin I into the biologically active peptide angiotensin II, which acts on the central nervous system to regulate renal sympathetic nerve activity, renal function, and blood pressure. High levels of angiotensin II have been proposed to play a key role in glucose and insulin regulation and studies have shown an increased risk of diabetes [484]. In a population of Native Americans, significant genetic associations were identified between a variant in ACE and insulin resistance and fasting hyperinsulinemia [485]. Further, large-scale studies across multiple populations have identified associations between alleles within the *ACE* gene and obesity, metabolic syndrome and type II diabetes in patients with and without hypertension [486-489]; however, a few studies have found no difference between the prevalence of an *ACE* insertion/deletion dimorphism in patients with type II diabetes versus controls [484,490,491]. Nonetheless, this gene is an intriguing biological candidate gene and warrants further investigation. In the Welsh ponies, two intronic and two upstream variants were identified in *ACE*. Further, 1 missense mutation (predicted moderate impact) was identified in exon 15 with a p-value of 1.11e-02 (**Table 4.10**). In addition, *TACO1* encodes a mitochondrial protein that functions as a translational activator of cytochrome c oxidase 1. In mice, this protein was one of several mitochondrial proteins that were upregulated in

obese mice fed a Western diet, as well as in obese mice who exercised on a wheel over those that were sedentary, indicating a role for mitochondrial proteins in obesity and promoting skeletal muscle health during exercise-induced weight loss [492]. In the Welsh ponies, two intronic and one upstream variant were identified in *TACO1*. Further, 1 5'UTR variant with a premature start codon (p-value=6.71e-03) and one synonymous mutation in exon 15 (p-value=5.62e-04; both predicted low impact variants) were identified for *TACO1* (**Table 4.10**).

ACTH on ECA 1 in the Welsh ponies: The region was not assessed for biological candidate genes given that all significantly associated SNPs were intergenic and no exonic SNPs within the full LD-ROI had a p-value of <0.05 (**Table 4.11**).

Basal glucose concentrations and laminitis status on ECA4 in Morgan horses: Based on the boundaries of the LD-ROI, this region on ECA4 was identified as shared between basal glucose concentrations and laminitis status. Given the more precisely fine-mapped region for glucose concentrations, *ABCA13* was identified as a candidate gene. In our cohort, *ABCA13* intronic variants were identified in the Morgan horses for glucose concentration (one SNP; **Table 4.12**) and laminitis status (10 SNPs; **Table 4.13**). In the horse, *ABCA13* has 62 exons and is associated with 5,097 variant alleles based on data from dbSNP, including the 2 synonymous variants in exon 34 and 46, the one missense variant in exon 42, and seven of the 11 intronic variants we identified (**Tables 4.12** and **4.13**). Further, our missense variant had a SIFT score of 1.0, indicating a well-tolerated amino acid substitution. Although the function of *ABCA13* has yet to be elucidated, it belongs to a subfamily of cell-membrane transporters with known roles in lipid metabolism [493]. *ABCA1* has been studied for its role in metabolic syndrome based on its known role in regulating high-density lipoprotein biogenesis, very-low-density lipoprotein production, and triglyceride lipolysis, [494]. Studies utilizing mouse models have identified that *ABCA1* has an important role in (i) β-cell insulin secretion and cholesterol homeostasis [495], (ii) adipocyte lipid metabolism and body weight [496], and (iii) GLUT4 trafficking and glucose uptake in skeletal muscle [497]. Less is known about the role of *ABCA13* although it has been linked to autism [498] and Schwachman-Diamond syndrome, a rare genetic disorder affecting the bone marrow, skeletal muscles, and pancreatic tissue [499].

Thus, additional investigation is required to determine if *ABCA13* has a role in metabolic syndrome and the impact of the variants that we identified.

INS-OST on ECA10: A total of 18 SNPs exceeded our threshold for genome-wide significance for this region, all of which were intronic SNP within the protein-coding gene *NKAIN2* (**Table 4.14**). This gene is a transmembrane protein that interacts with the β subunit of the sodium/potassium transporting ATPase; however, the role of this gene is not well established but is thought to be primarily involved in neuronal development and function [500]. This was supported by initial studies which linked variants in *NKAIN2* with mental disorders such as Schizophrenia and depression [501,502]. In addition, evidence suggests that *NKAIN2* acts as tumor suppressor [503], with down regulation or loss of function mutations associated with prostate cancer [504]. However, recent studies have also suggested a role for *NKAIN2* in obesity and lipid metabolism. Using longitudinal exome-wide association analysis, *NKAIN2* was identified as one of three novel SNPs associated with body mass index in a population of Japanese subjects [505]. In addition, investigators found evidence that SNPs within *NKAIN2* were interacting between total fat intake to influence the variation of low-density lipoproteins [506]. Thus, the role of *NKAIN2* is still being investigated and a specific biological function in glucose and insulin homeostasis has not been identified.

Leptin concentrations on ECA19 in Morgan horses: There is currently no literature available on the pseudogenes or lncRNA identified with statistically significant variants in this region (**Table 4.15)**. The protein coding gene *BBX* is a member of a superfamily of high-mobility group of architectural transcription factors, which are responsible for transcription, replication and DNA repair [507]. Specifically, *BBX* belongs to the family of high-mobility group box (*HMGB*) which bind, bridge, and loop DNA for transcription [508]. There is limited information about the function of *BBX*; however, variants in high-mobility AT-hook 2 (*HMGA2*) has been associated with metabolic syndrome in humans and correlated with metabolic phenotypes in Welsh ponies (see **Chapter 5**). Thus, members of this superfamily warrant further investigation. In the Morgan horses, three downstream and nine intronic variants were identified in this gene (**Table 4.15).**

NH concentrations on ECA19 in Morgan horses: This region contained a number of protein-coding genes with either intronic or exonic variants (**Table 4.16)**. Of these genes, *XXYLT1* has been associated with obesity and *BDHI* has a known function in fatty acid and ketone production. *XXYLT1* is an integral membrane protein and belong to the GT8 family of glycosyltransferases and has been shown to have an essential role in glucose biology via targeting Notch proteins [509]. This gene was also found to be in one of seven potentially pleiotropic loci associated with osteoporosis and obesity in humans [510]. In our cohort, 11 intronic SNPs were identified in *XXYLT1* which exceed the threshold for genome-wide significance (**Table 4.16).** *BDHI* gene belongs to a family of dehydrogenase/reductases enzymes and encodes a homotetrameric lipid-requiring enzyme of the mitochondrial membrane. *BDHI* is responsible for beige fat differentiation [511] and has a role in mitochondrial production of ketone bodies during fatty acid catabolism via the reduction of acetoacetate to beta-hydroxy butyrate. In a mouse model of type II diabetes and diabetic cardiomyopathy, *BDHI, HMGCS2*, and *PDK4* were found to be upregulated in response to PPAR-γ activation secondary to obesity [512]. Upregulation of all three ketogenic enzymes correlated with obesity, lipotoxity and cardiac dysfunction [512]. In the Morgan horses, one synonymous variant was identified in exon 7 of this gene which had a p-value of 3.59e-02 (**Table 4.16)**. Given that NH ratio is a measurement of obesity, these genes were considered as biological candidate genes.

Adiponectin concentrations on ECA6 in Morgan horses: This region contained a single intergenic SNP which exceeded the threshold for genome-wide significance and 20 exonic variants within protein coding genes that had a p-value of <0.05 which were further assessed for biological function (**Table 4.17)**. NCK-associated protein 5 like (*NCKAP5L)* has an important role in regulating microtubule organization and stabilization and causal variants have been identified for Retinitis Pigmentosa. In a study evaluation differential gene expression in tissue from obese individuals, *NCKAP5L* was found to be upregulated in adipose tissue and downregulated in the blood and was located within 1Mb of a known obesity susceptibility SNP, providing evidence for a novel candidate gene for obesity [513]. In the Morgan horses, one synonymous variant (low impact) was identified in exon 1 with a p-value of 3.80e-02 (**Table 4.17**). Glycerol-3-phosphate dehydrogenase 1 (*GPD1)* is a member of the NAD-dependent glycerol-3-phospate dehydrogenase family and plays

a critical role in carbohydrate and lipid metabolism. The enzyme is responsible for catalyzing the reversible conversion of NADH to glycerol-3-phospate and NAD+, as well as facilitating the movement of glycerol from the cytosol to the mitochondria. Mutations in this gene are the cause of transient infantile hypertriglyceridemia, and decreased concentrations have been reported in insulin-resistant individuals [514]. Comparison of adipose tissue depots showed decreased expression of lipoprotein lipase, adiponectin, and GPD1 in omental fat which was accentuated in the presence of glucose intolerance [515]. Similarly, investigation of adipose tissue dysfunction secondary to insulin-resistance revealed that *GPD1* served as a canonical marker of adipogenesis, and both adiponectin and *GPD1* correlated with decreased expression of ErbB1 [516], potentially indicating a direct link between adiponectin and *GPD1* in insulin resistance. In our cohort one missense variant (moderate impact) was identified in exon 1 of *GPD1* with a p-value of 3.51e-02 (**Table 4.17**). Several genes with variants identified in this fine-mapped region have been associated with lipid metabolism although are not directly related to adiponectin, obesity or metabolic syndrome. Variants in *LIMA1* have been found to lower low-density lipoprotein-cholesterol levels and inhibiting gastrointestinal cholesterol absorption [517], microRNAs have been shown to target *METTL7A* during lipid droplet formation suggesting that these micoRNAs may act as a biomarker for obesity or MetS [518], and the non-heme transporter *SLC11A2* has been shown to be upregulated in the intestinal cell of obese patients with type II diabetes potentially explaining the elevated iron levels in many of these patients [519]. In the Morgan horses we identified missense variants (moderate impact) in exon 1 of *LIMA1* (p-value=3.17e-02), exon 2 of *METTL7A* (p-value=3.17e-02), and exon 12 of *SLC11A2* (p-value=2.97e-02; **Table 4.17**) Finally, aquaporin 6 (*AQP6*) belongs to a family of membrane water channels which are involved in water/salt homestasis, exocrine fluid secretion and epidermal hydration [520]. Mouse knockout models of *AQP7*, an aquaglyceroporin expressed in adipocytes, have been shown to have increased fat mass and adipocyte hypetrophy, indicating an essential role for plasma membrane glycerol permeability in adipose glycerol and triglyceride accumulation [521]. AQP7 and APQ9, an aquaporin important for hepatic glycerol uptake, have both been suggested as metabolic regulator in diabetes and obesity [522]. Less is known about the function of APQ6 but based on sequence similarities it belongs to the subfamily of

orthodox aquaporins which are mainly selective for the transfer of water across the plasma membrane, and a role for this subfamily in metabolic syndrome or obesity has not been established [523]. In our cohort, we identified a premature start codon in the 5'UTR (low impact; p-value=3.17e-02) of this gene (**Table 4.17)**. Hence, this region contains several positional candidate genes with the top biological candidate genes being *GPD1* and *NCKAP5L*.

Non-Coding Variants and Future Directions

Notably, all variants which exceeded the threshold for genome-wide significance mapped to non-coding regions of the genome, with 66.7% of the significantly associated SNPs being intronic, 17.0% intergenic (including the entire fine-mapped region for ACTH on ECA1 in the Welsh ponies) and 10.3% within lncRNA. Consistent with what we found, in human association studies ~90% of the phenotype-associated SNPs were found to intergenic or within the introns, promoters, or 5' or 3' UTR of coding genes or small non-coding RNAs [524-527]. This could be explained by several factors. First, association studies using WGS follow the same principals of GWA indicating that the statistical power to detect sample size is dependent on (i) sample size, (ii) effect size of the causal variant, (iii) the allele frequency of the causal variant, and (iv) the LD between correlated and causal variants [528]. Therefore, the causal variant may not have reached statistical significance in our population but was tagged by variants within LD of the causal variant. Based on the conservative Bonferroni correction imposed on most association studies, several studies using multi-omics data have shown that the causal variant does not reach the set threshold for significance but are rather "sub-threshold" [529]. In order to capture potential causal variants within protein-coding genes, we further evaluated positional candidate genes with exonic variants in our fine-mapped region with a p-value <0.05.

Second, the causal variant may not have been represented in our population. The principles of fine-mapping requires a complete catalog of all variants in the associated region in order to capture the causal variant, which is highly dependent on accurate genotyping or imputation of the region [458]. For our analysis, we imputed SNP array data to WGS from a reference population of 19 Welsh ponies and 18 Morgan horses. The reference population was chosen to be representative of both insulin sensitive and insulin

dysregulated horses; however, complex traits are highly polygenic, and each individual likely has a unique combination of risk alleles, indicating that not all causal variants may be present in our reference population. Further, imputation is a statistical best guess of missing genotypes with an error rate of ~2-6% [530]. In order to minimize imputation errors, we calculated concordance on individuals with both WGS and SNP array data or we masked down individuals from WGS to the SNP array, and SNPs without 100% concordance were excluded. Evaluation of the percentage of intergenic, intronic, and exonic variants and the transition/transversion ratios for the within breed variant annotation was also consistent with what we would expect for WGS [531,532]. However, there is still a chance that an undetected error would lead to a missed causal variant. In addition, mapping of WGS is dependent on the quality of the reference genome. We utilized EquCab3, released in 2018, which improved the count of non-N bases from 2.33 Gb in EquCab2 to 2.41 Gb in EquCab3 [445]. However, this still indicates that the reference genome is not fully annotated and contains region that cannot be mapped by our WGS. Finally, a limitation of Beagle software is that it does not impute copy number variants, small insertions/deletions (indels), or tri-allelic SNPs, and so these variants were not represented in our analyses. Therefore, a complete investigation of these types of variants in the entire cohort of Welsh Ponies and Morgan horses was not possible using imputation. In future these variant types will be further evaluated by performing cytogenetic analysis or genotyping our population using array comparative genomic hybridization. Although the relative impact of indels and copy number variants on complex traits has not been thoroughly invested, it has been proposed that they account for a portion of the missing heritability in GWA studies [533]. In a study evaluating gene expression, 17.7% of the total genetic variation in gene expression was captured by copy number variants [534], and copy number variants have been associated with several neurological complex diseases including Schizophrenia [535]. Further, the evaluation of 89 trait associated loci including 1,319 SNPs and 88 indels revealed that indels were the most likely causal variant in seven loci [536]. Thus, causal variants due to one of these variants would not have been represented in our analyses but could have a functional effect on EMS.

Third, the causal variant resides within the non-coding region of the genome with functional consequences on protein-coding genes. Interestingly, assessment of 21 human

tissues associated with 392 diseases revealed that 85% of the transcript blocks contained novel, non-coding transcripts [537]. Although these regions were originally thought of as non-essential DNA, projects such as the Encyclopedia of DNA Elements (ENCODE) and Functional Annotation of the Animal Genomes (FAANG) have provided a large dataset of experimental evidence for a functional role of non-coding regulatory elements [524-527,538,539]. Further, evaluation of association studies has shown that statistically significant variants are enriched within regulator sequencing including enhancer elements, DNase hypersensitivity regions and chromatin marks [526,540,541]. Specifically, intronic variants have been shown to have a functional effect by activating abhorrent splice sites, creating a novel acceptor or donor splice site, altering splicing regulatory elements, or disrupting transcription regulatory motifs and non-coding RNA genes [542], and over 180 deep intronic pathogenic variants have been identified for 77 different disease [542]. In addition, lncRNAs are defined as non-protein coding transcripts with greater than 200 nucleotides that structurally resemble mRNA [543], and it has been estimated that 54% of lncRNAs are located in intergenic regions [544]. Long noncoding RNAs have roles in epigenetic regulation, chromosome-imprinting, cell-cycle control and cell apoptosis [545] and recent studies have implicated them in the pathophysiology and pathogenesis of endocrine, reproduction, metabolic, immune, nervous and cardiovascular diseases [546]. Thus, the functional effect of the non-coding variants in our analyses requires further investigation.

However, prioritization of non-coding variants poses a more difficult challenge then protein-coding-regions. Human studies have relied on publicly available data to prioritize causal variants by identifying those which overlap with accessible chromatin, transcription factor binding, or histone marks associated with regulatory activity [244]. However, there is a large gap in knowledge of the function of these regions in horses. FAANG seeks to provide this resource to the animal genetics community but this is a large, multi-collaborative project and data release has been slow. Further, across species non-coding regions pose a challenge in deciphering their biological effect due to interactions with both proximal and distal protein-coding genes [547]. Analysis of trait-relevant, multi-omics data (for example genomic and transcriptomic data) may provide the means necessary to identify the targets of non-coding variants [526].

In summary, this chapter provided intriguing biological evidence for the role of multiple coding genes in the pathogenesis of EMS but did not conclusively identify the causal variants and additional follow-up is required. Several methods could be utilized to further interrogate our regions for both protein-coding and non-coding causal variants. First, interrogation of the allele frequency of the variants identified in this chapter in a large database of mixed-breed horses would allow for the assessment of the frequency of these variants in healthy horses. Given that EMS can manifest at an older age, of particular interest would be assessment in breeds considered low-risk for EMS such as the Quarter Horse or Thoroughbred. Identification of variants at low frequency in these breeds would allow for the prioritization of specific biological candidate genes for interrogation through Sanger sequencing. Second, haplotype analysis can be utilized to find shared ancestral haplotypes to further fine-map the LD-ROI, prioritize variants, and identify additional horses for whole-genome or Sanger sequencing. Third, development of a custom high-throughput genotyping assay would allow for the validation of imputed genotypes as well as assessment of the statistically significant coding and non-coding variants in an independent population of Welsh ponies and Morgan horses phenotyped for EMS.

Importantly, the 10 regions described here for variant annotation and assessment of biological candidate genes were chosen if the fine-mapped region contained one or more SNPs that exceeded the threshold for genome-wide significance, appeared to have clear delineation between baseline SNPs and those that exceeded the threshold, and where significantly associated SNPs appeared to be tightly clustered. However, this only provided a criterion for initial analysis and does not exclude the remaining high priority LD-ROI. Therefore, future directions also include interrogation of the remaining high priority LD-ROI as well as fine-mapping the medium and low priority LD-ROI.

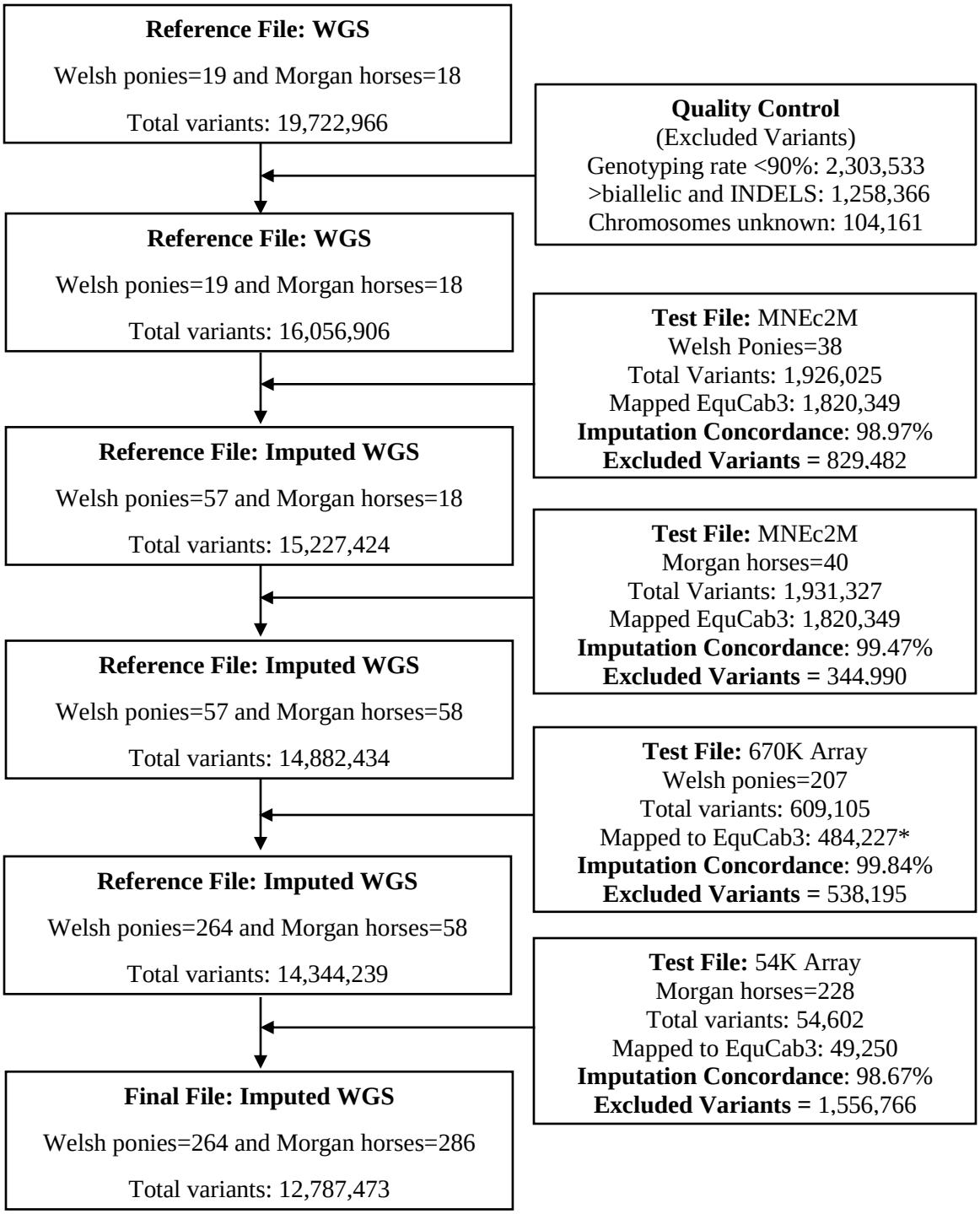


Figure 4.1: Imputation of SNP genotyping arrays to whole genome sequencing. Imputation concordance was determined from individuals which were genotyped for WGS and on the corresponding SNP array or masked down from WGS to the SNP array. SNPs without 100% concordance were removed and the imputed horses were added to the reference genome. *For the Welsh ponies genotyped on the 670K array, 584,301 SNPs mapped to EquCab3; SNPs which were noninformative or mapped to chromosome unknown were excluded when masking down horses from WGS.

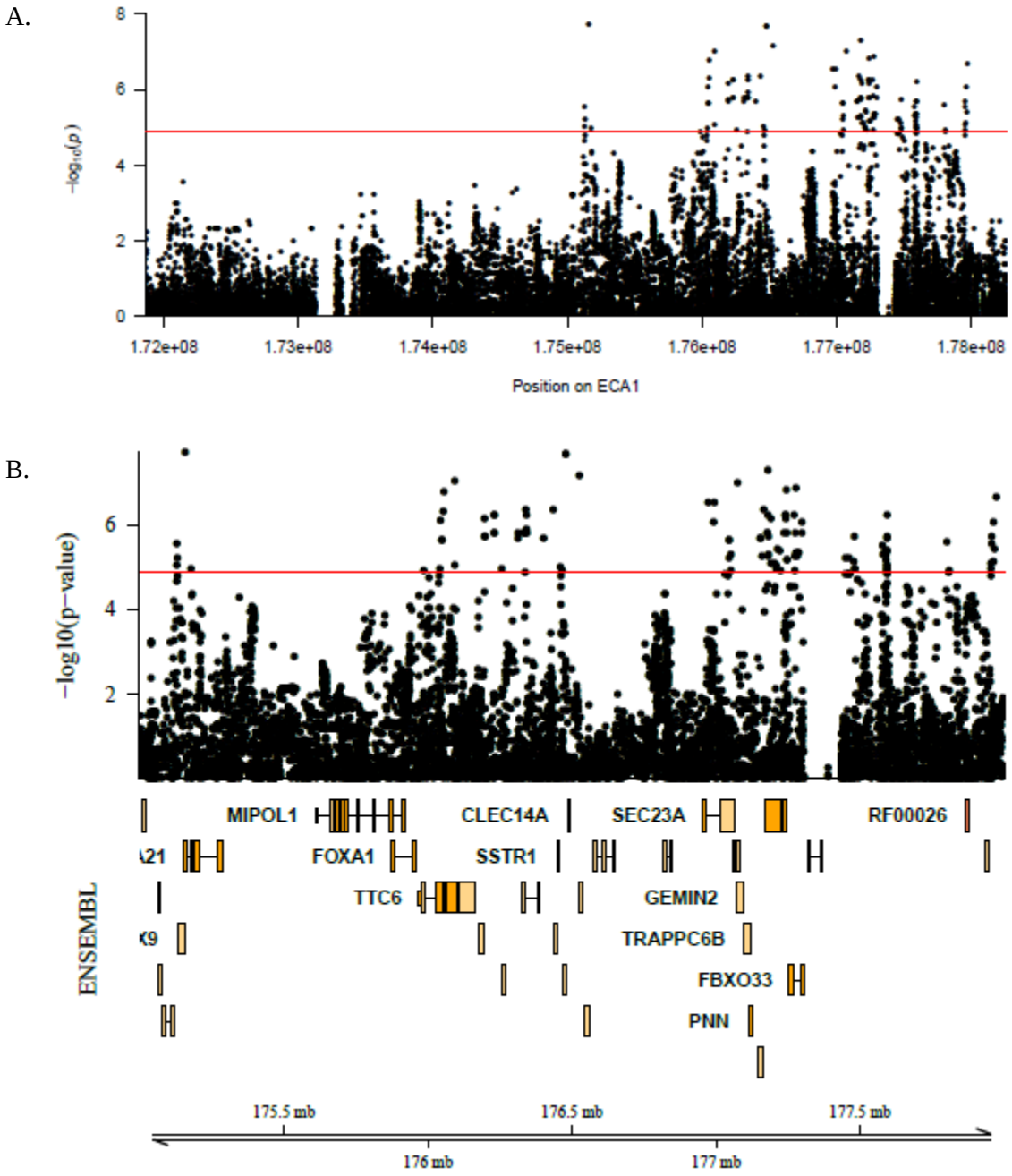


Figure 4.2: Fine-mapped region for adiponectin concentrations on chromosome 1 in the Welsh ponies. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 171861236 to 178270042, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <1.3e-05, of which 117 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 175000000 to 179000000. Aligning the Ensembl genome browser identified statistically significant variants in protein-coding genes including *PAX9, SLC25A2, TTC6, ABCA13, SSTR1, SEC23A, GEMIN2, MIA2, and FBX033.*

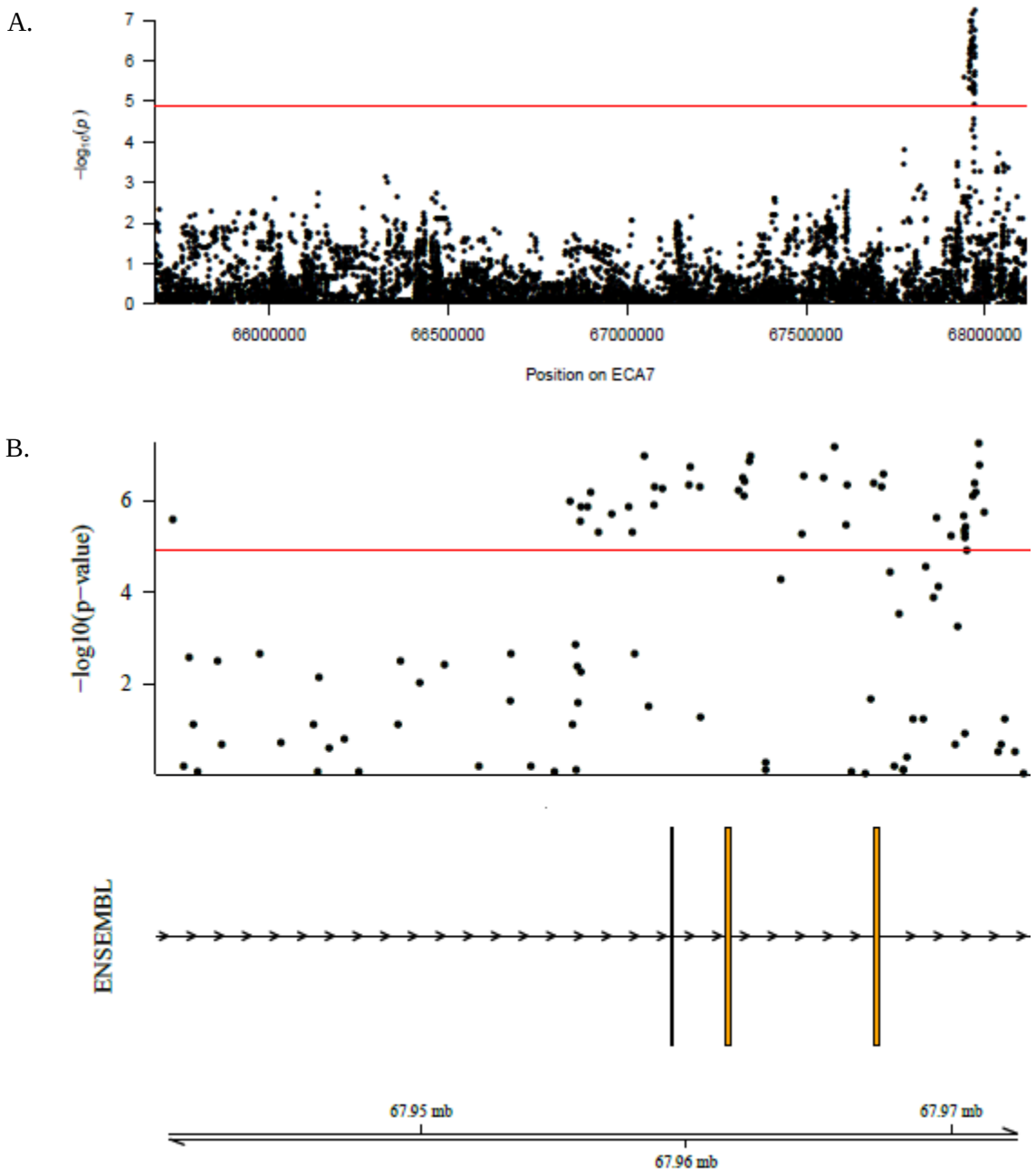


Figure 4.3: Fine-mapped region for leptin concentrations on ECA7 in the Welsh ponies. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 65678376 to 68117086, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <1.3e-05, of which 46 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 67940000 to 67973000. Aligning the Ensembl genome browser revealed that all statistically significant variants were in a single protein-coding gene, *TENM4*.

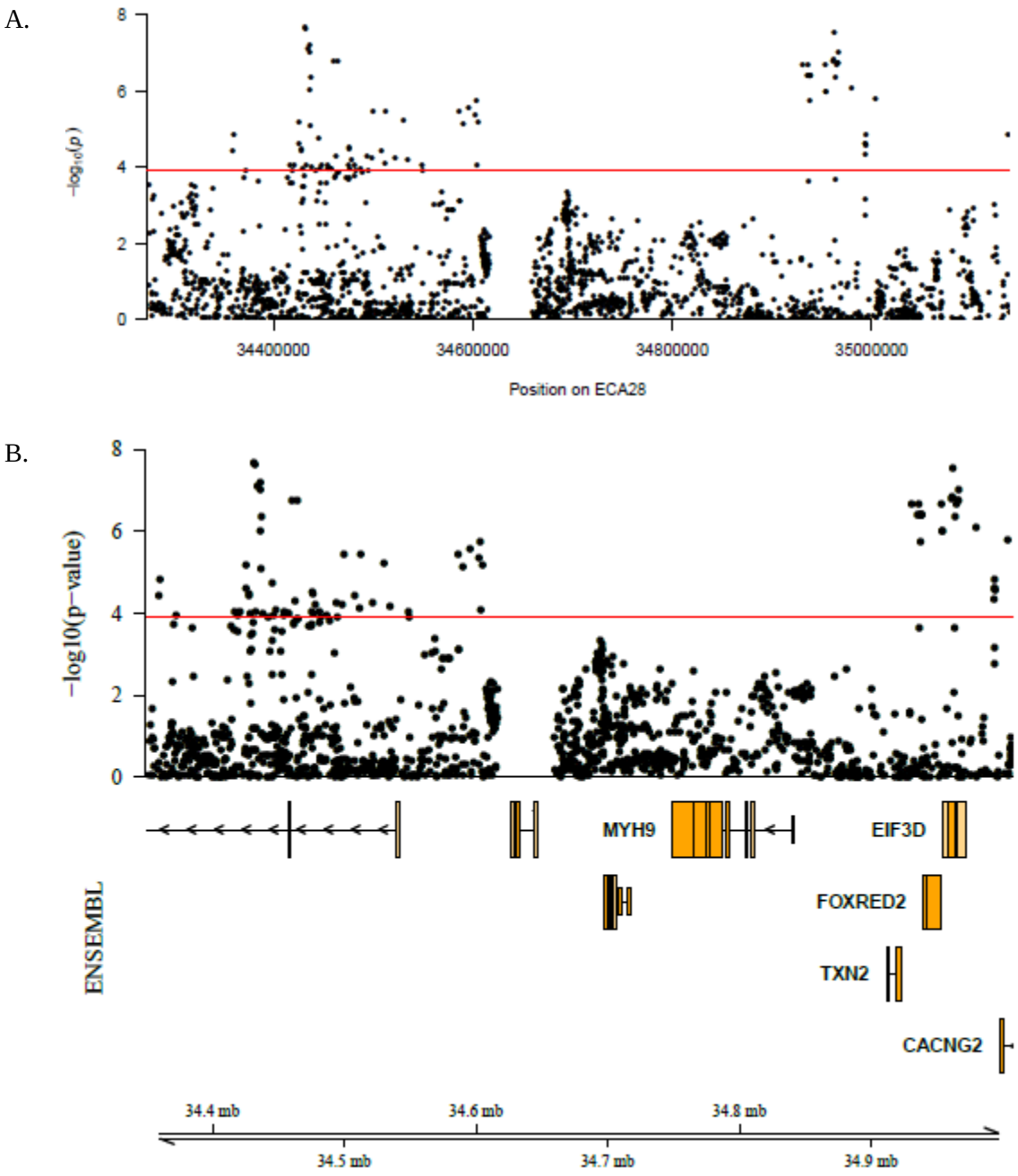


Figure 4.4: Fine-mapped region for glucose concentrations post oral sugar test on ECA28 in the Welsh ponies. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 34271949 to 35138699, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <1.2e-04, of which 85 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 34350000 to 35007000. Aligning the Ensembl genome browser revealed that statistically significant variants were in the protein-coding gene *RBFOX*, *FOXRED2*, *EIF3D*, or *CACNG2*.

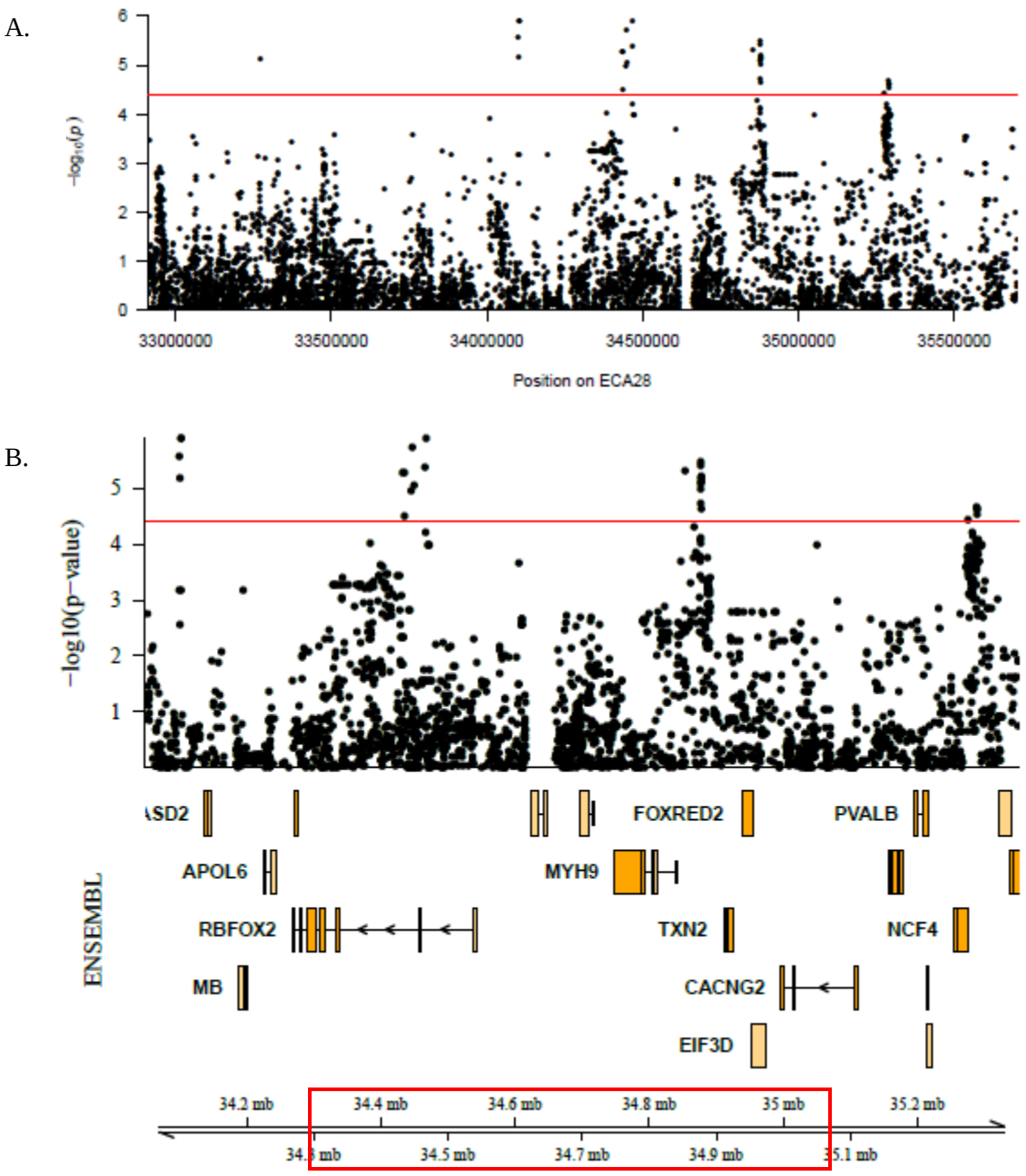


Figure 4.5: Fine-mapped region for glucose concentrations post oral sugar test on ECA28 in the Welsh ponies. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 32909542 to 35703535, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <3.9e-05, of which 30 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 34050000 to 35350000. Aligning the Ensembl genome browser revealed that statistically significant variants were in the protein-coding gene *RBFOX2* and *NCF4*. The red box around the base pair positions indicates the shared fine-mapped region for glucose concentrations post oral sugar test (see **Figure 4.4**).

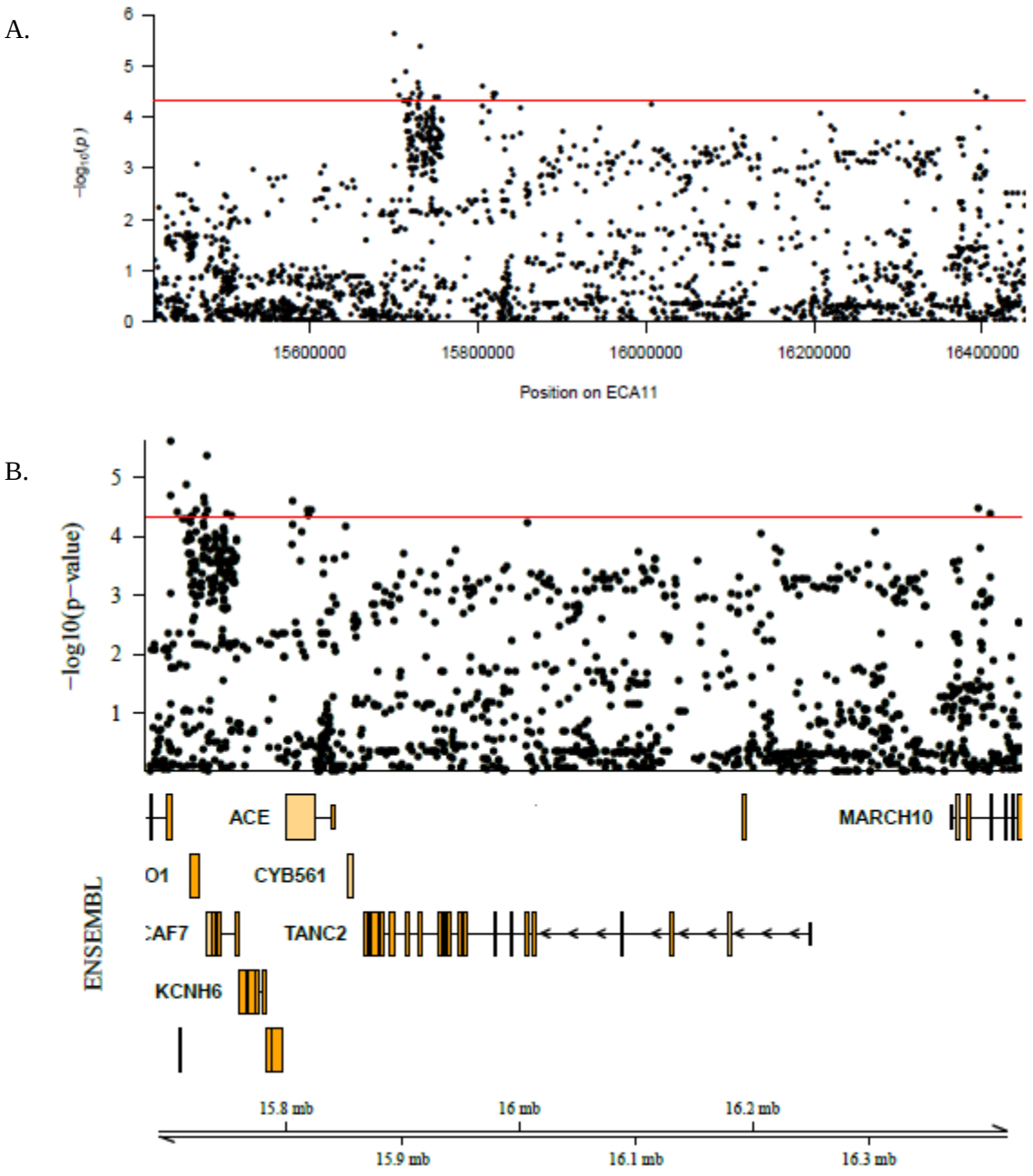


Figure 4.6: Fine-mapped region for glucose concentrations post oral sugar test on ECA28 in the Welsh ponies. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 15414337 to 16451463, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <4.6e-05, of which 19 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 15700606 to 16403791. Aligning the Ensembl genome browser revealed that statistically significant variants were in the protein-coding gene *MAP3K, TACO1, DCAF7, ACE,* and *MARCH10.*

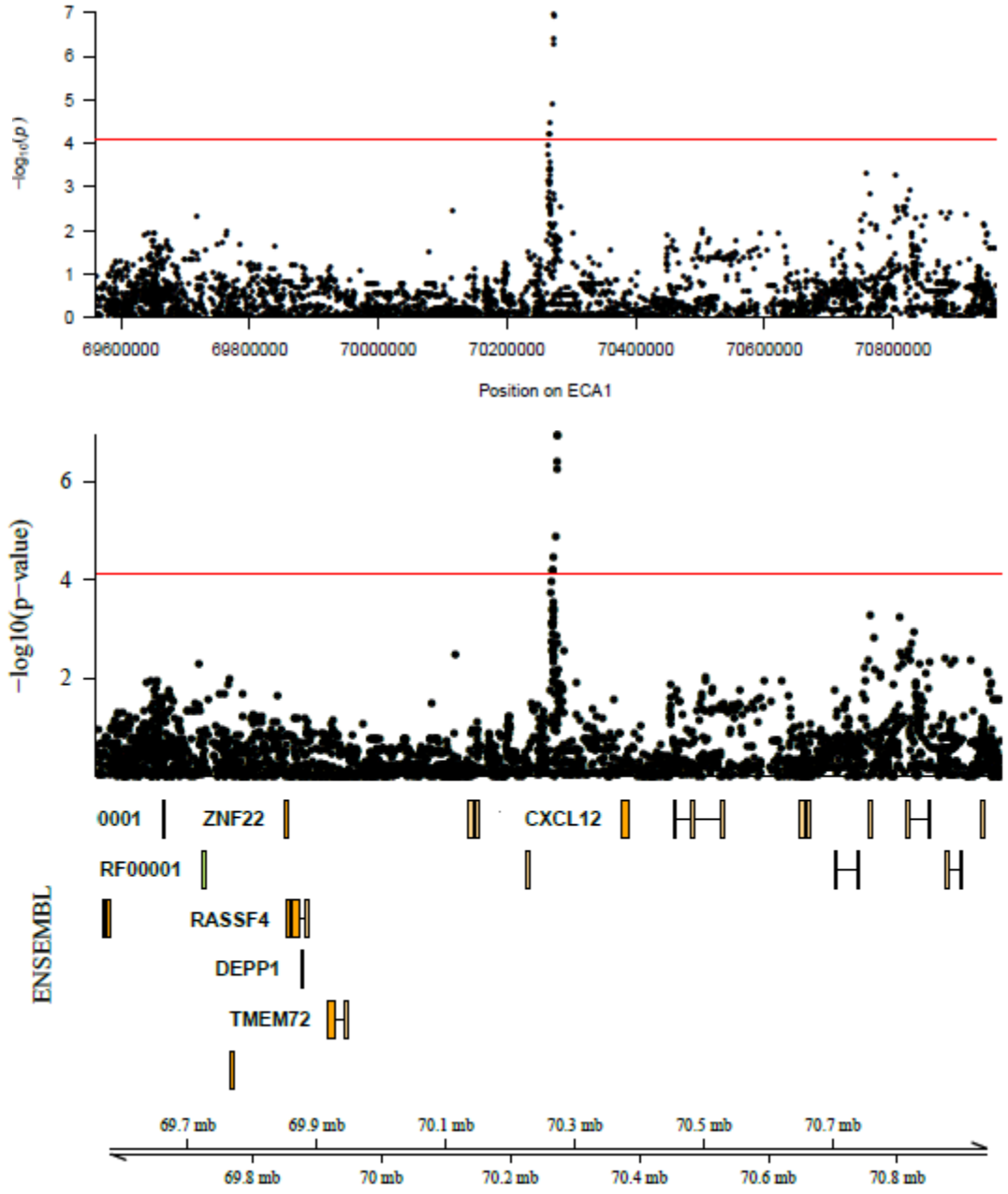


Figure 4.7: Fine-mapped region for ACTH concentrations ECA1 in the Welsh ponies. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 15414337 to 16451463, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <4.6e-05, of which 19 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI. Aligning the Ensembl genome browser revealed that statistically significant variants were all intergenic.

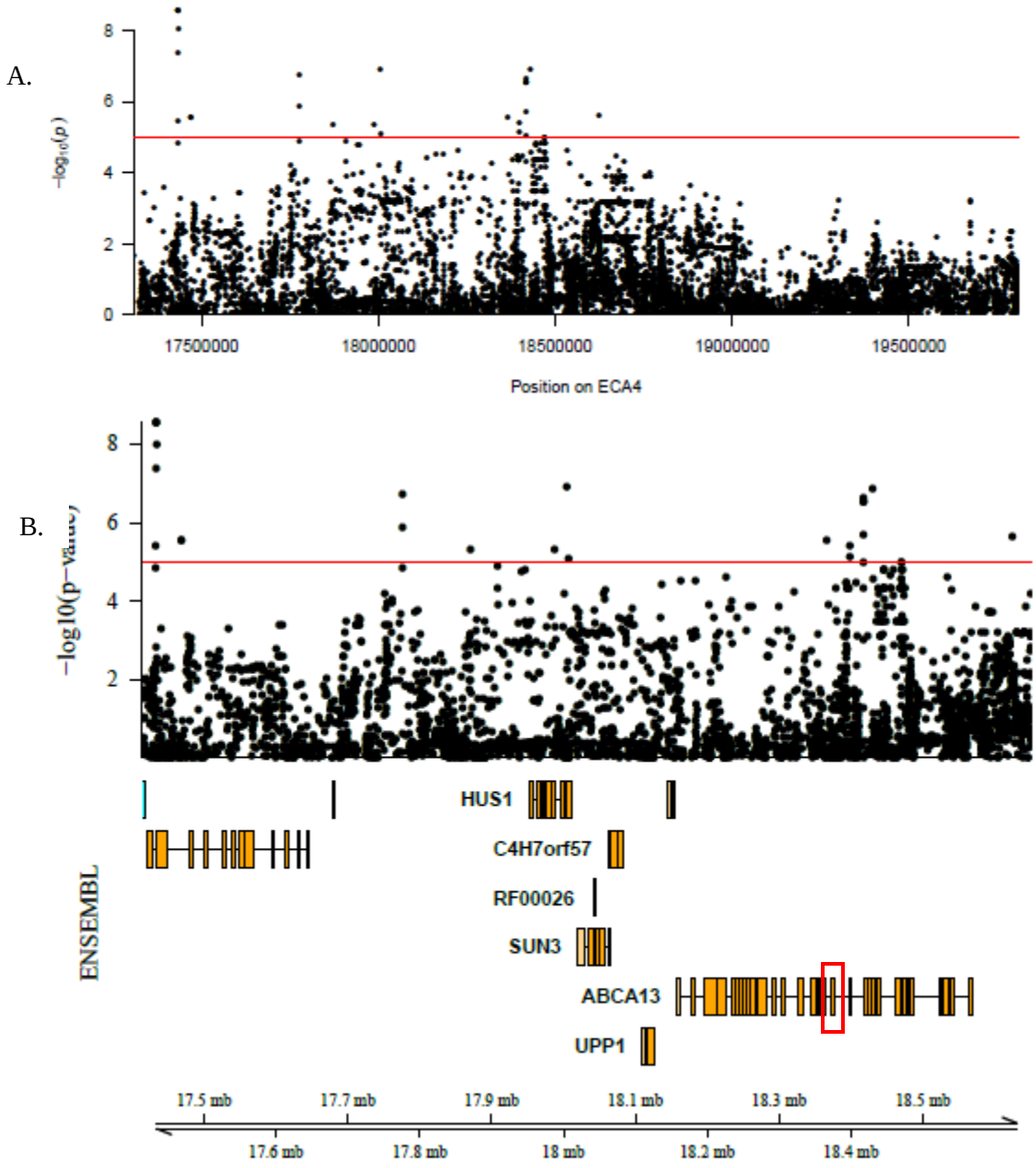


Figure 4.8: Fine-mapped region for laminitis status on chromosome 4 in the Morgan horses. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 17307352 to 19812647, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <9.9e-06, of which 25 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 17415000 to 18650000. Aligning the Ensembl genome browser identified statistically significant variants in protein-coding genes including *TNS3*, *PKD1L1*, and *HUS1*, *ABCA13*. The red box around *ABCA13* marks the approximate region for significant SNP for glucose concentrations on ECA4 SNPs (see **Figure 4.9**).

A.

B.

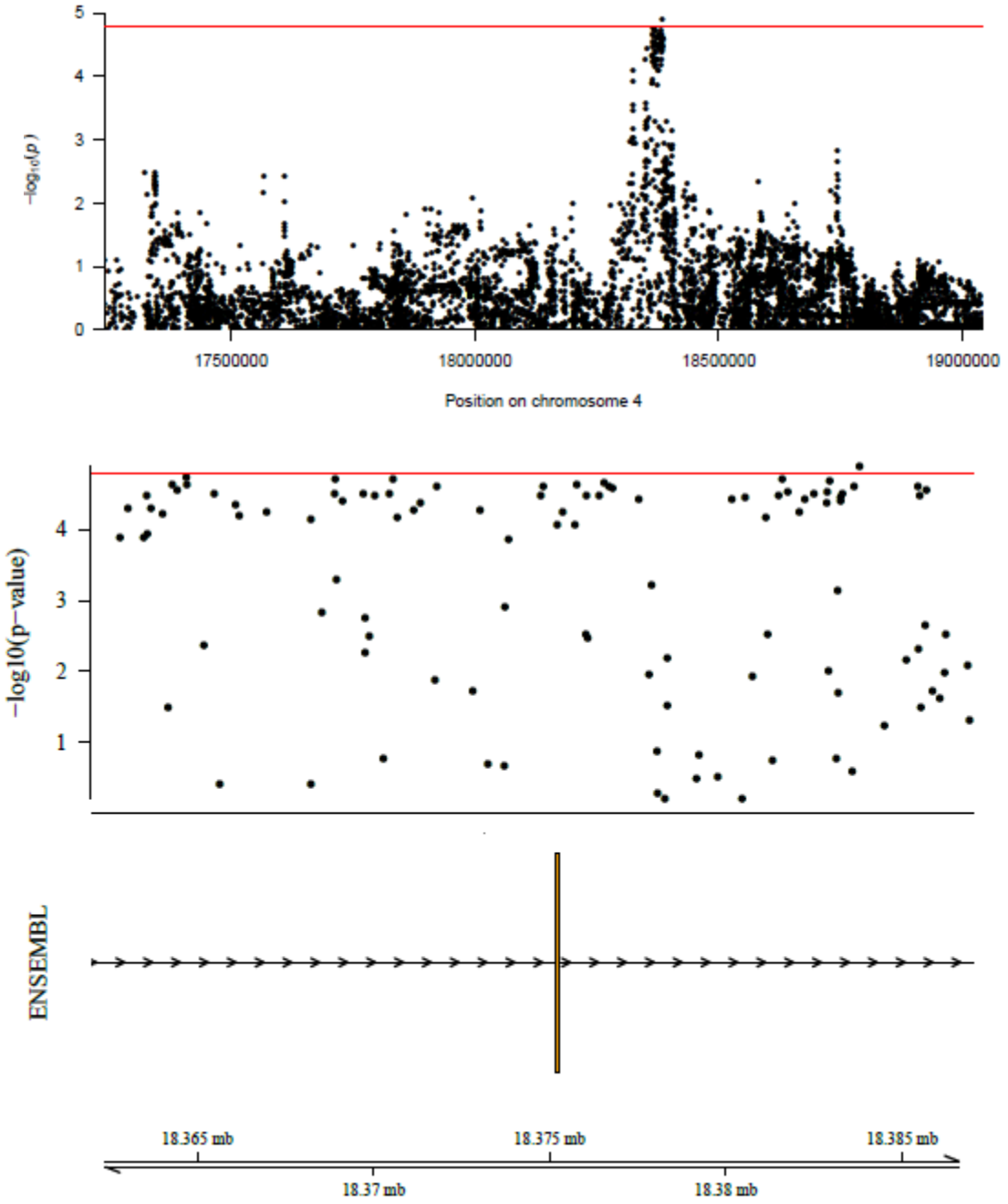
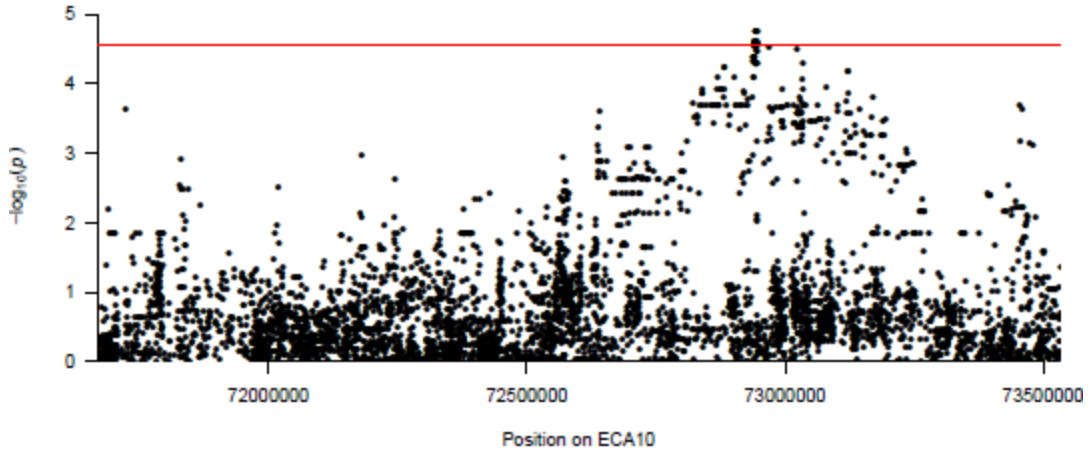


Figure 4.9: Fine-mapped region for basal glucose concentrations on ECA4 in the Morgan horses. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 17239374 to 19043831, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <1.5e-04, of which 64 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 18320000 to 18390000. Aligning the Ensembl genome browser revealed that all statistically significant variants were in a single protein-coding gene, *ABCA13*.

A.



B.

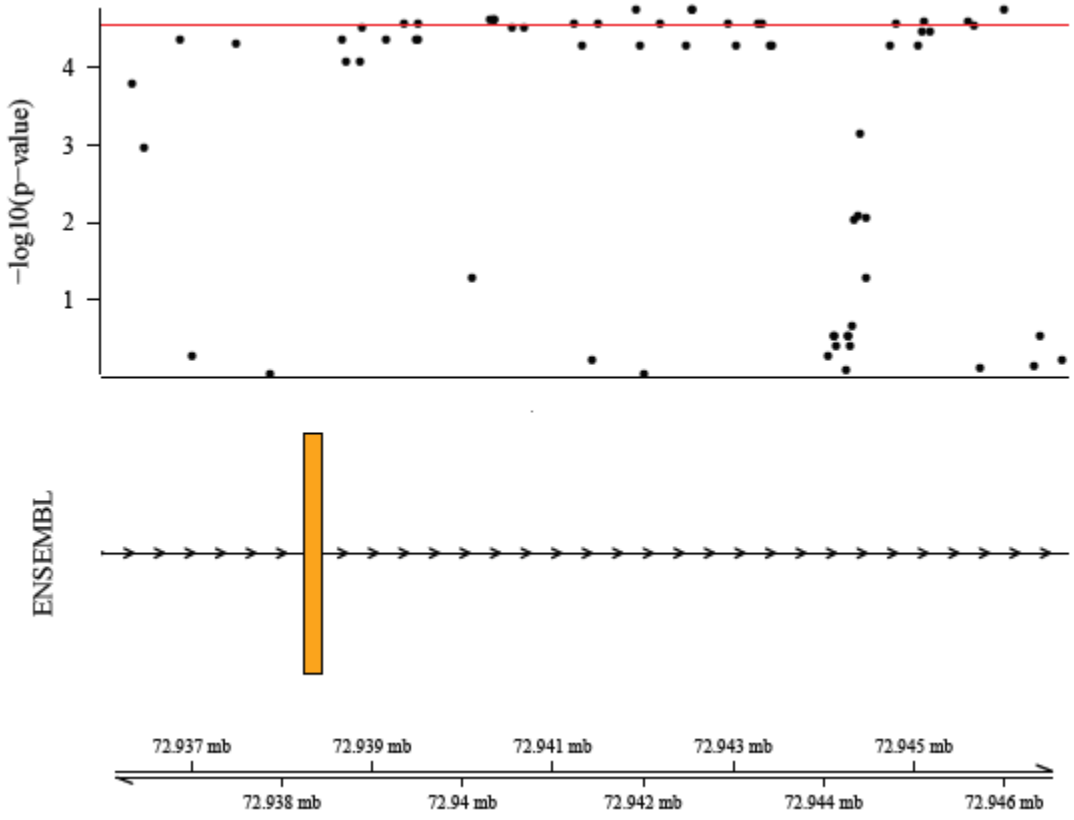


Figure 4.10: Fine-mapped region for insulin concentrations post oral sugar test on ECA10 in the Morgan horses. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 71666607 to 73534053, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <2.7e-05, of which 18 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 72936000 to 72946700. Aligning the Ensembl genome browser revealed that all statistically significant variants were in an intron of a single protein-coding gene, *NKAIN2*.

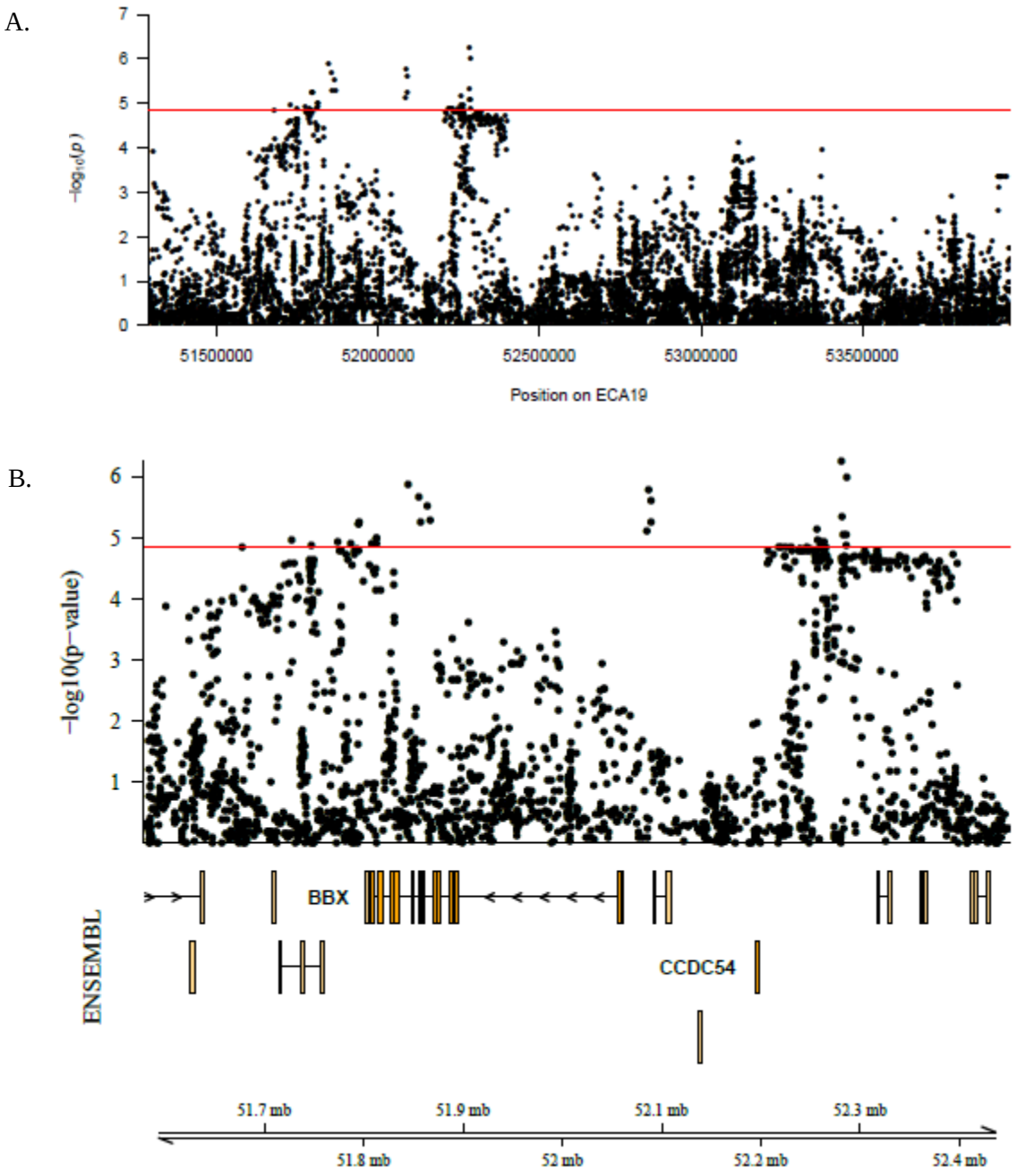


Figure 4.11: Fine-mapped region for leptin concentrations on ECA19 in the Morgan horses. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 51386493 to 53959028, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at < 1.3e-05, of which 35 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 51580000 to 52450000. Aligning the Ensembl genome browser revealed that all statistically significant variants were intergenic, and intragenic, intronic, or downstream of 2 pseudogenes, 1 lncRNA, and the protein coding gene *BBX*.

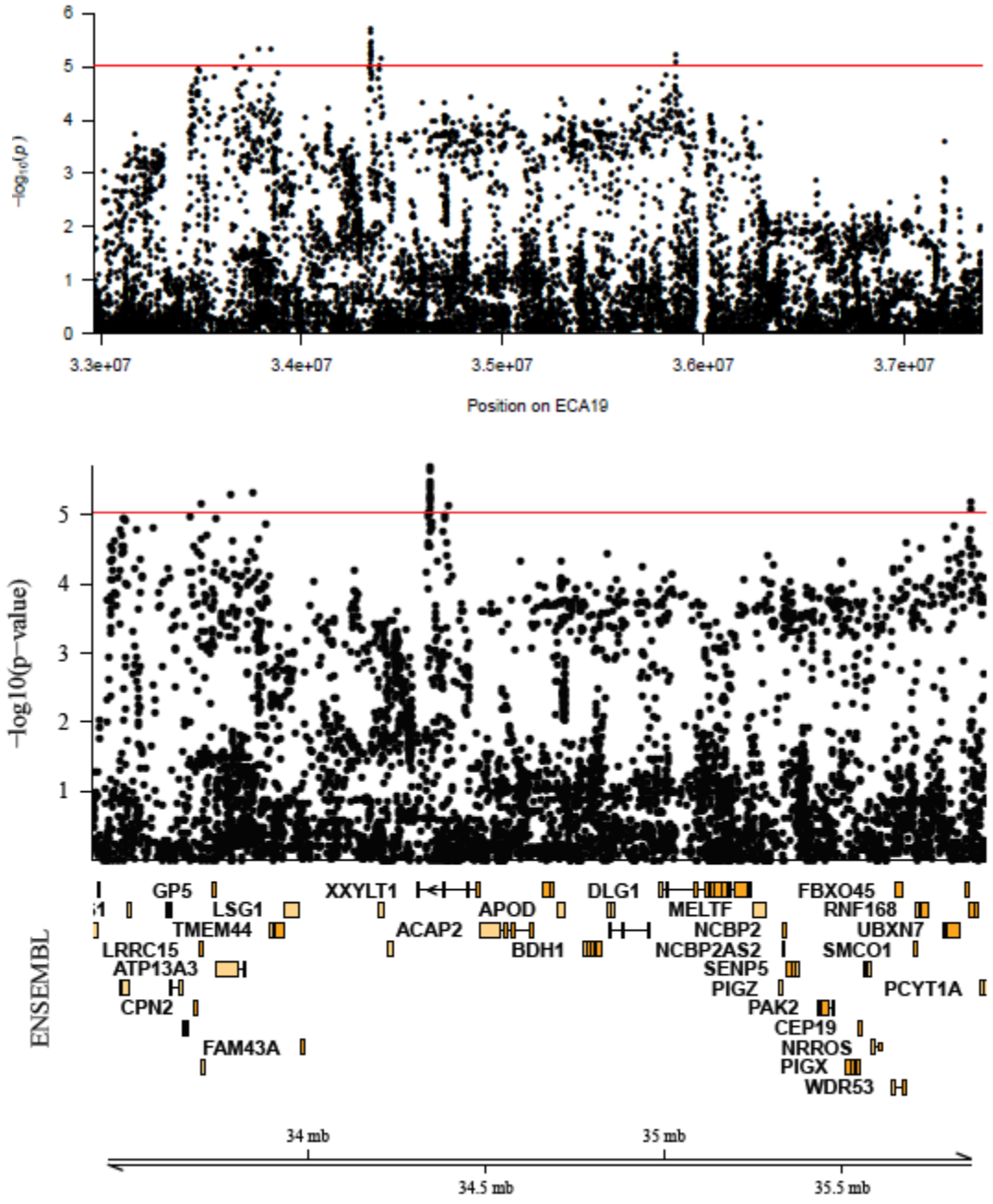


Figure 4.12: Fine-mapped region for NH ratios on ECA19 in the Morgan horses. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 32962795 to 37391949, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at < 9.4e-06, of which 16 SNPs exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 33400000 to 35900000. Aligning the Ensembl genome browser revealed that all statistically significant variants were intergenic, and intragenic, intronic, or upstream of 1 lncRNA, and the protein coding genes *LRRC15*, *ATP13A3*, and *XXYLT1*.

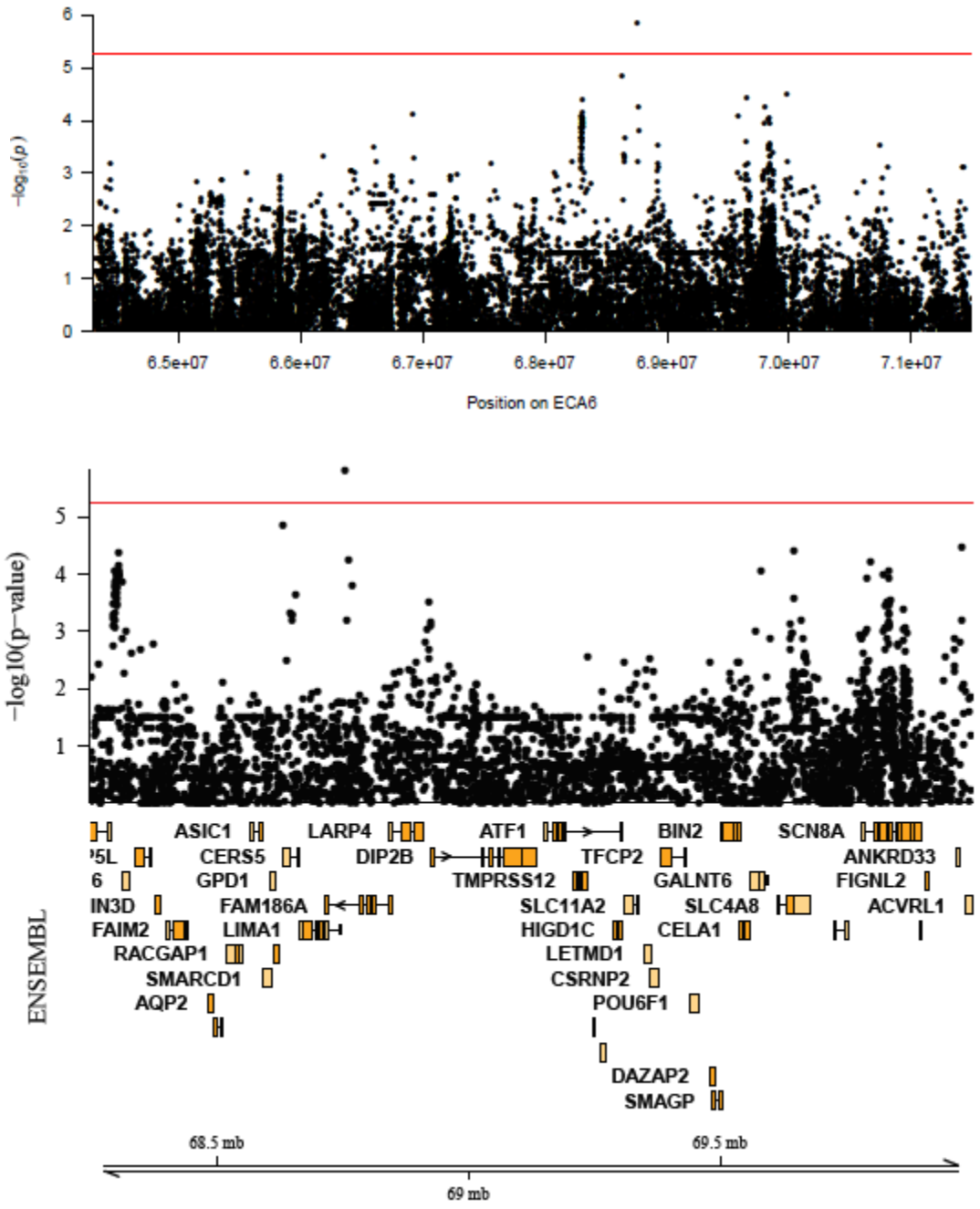


Figure 4.13: Fine-mapped region for adiponectin concentrations on ECA6 in the Morgan horses. (A) Manhattan plot of the full LD-ROI identified on GWA from base pair positions 64297403 to 71493047, with base pair positions on the x-axis and the -log10 of the p-value on the y-axis. Red line represents the threshold for genome significance set at <5.7e-06, of which 1 SNP exceeded this threshold. (B) Manhattan plot of the fine-mapped LD-ROI from base pair positions 68250000 to 70001000. Aligning the Ensembl genome browser revealed that the statistically significant variant was intergenic.

Table 4.1: EMS phenotypic data for Welsh ponies chosen for whole-genome sequencing					
EMS_ID	**Laminitis Status**	**Basal Insulin (µIU/mL)**	**Insulin-OST (µIU/mL)**	**Triglycerides (mg/dL)**	**Adiponectin (µg/mL)**
Mean (sd)		10.9 (17.2)	36.2 (45.4)	28.0 (34.8)	6.9 (5.5)
EMS_28	n	34.7	84.2	46.8	1.8
EMS_363	y	12.4	49.3	119.3	0.4
EMS_369	n	2.8	15.94	41.4	3.3
EMS_373	n	4.3	13.0	64.7	2.8
EMS_376	y	14.4	107.6	241.0	2.7
EMS_457	y	119.9	142.9	32.3	0.5
EMS_657	n	7.2	9.1	0.7	4.7
EMS_676	n	10.0	11.3	1.7	11.4
EMS_697	n	1.5	9.7	13.4	7.3
EMS_699	y	22.0	246.4	69.7	0.06
EMS_737	n	1.5	4.2	4.9	8.5
EMS_738	y	28.1	300.0	40.9	0.5
EMS_739	y	218.3	308.0	72.5	0.6
EMS_765	n	4.9	4.1	8.3	37
EMS_790	y	26.0	153.4	25.3	1.3
EMS_794	n	2.3	20.3	3.4	23.4
EMS_812	n	6.1	12.2	15.1	32
EMS_820	n	7.5	8.8	1.3	8.4

Table 4.1: EMS phenotypic data for Welsh ponies chosen for whole-genome sequencing. Ponies were chosen to represent both insulin sensitive and insulin dysregulated horses (individuals with a history of laminitis and at the breed-specific upper range for the EMS biochemical measurements baseline insulin, insulin concentrations post oral sugar challenge, or triglycerides, and breed-specific lower range for adiponectin concentrations. Breed-specific ranges are provided for each biochemical measurement. Abbreviations: Insulin-OST: insulin concentrations post oral sugar test.

Table 4.2: EMS phenotypic data for Morgan horses with whole-genome sequencing					
EMS_ID	**Laminitis Status**	**Basal Insulin (µIU/mL)**	**Insulin-OST (µIU/mL)**	**Triglycerides (mg/dL)**	**Adiponectin (µg/mL)**
Mean (sd)		8.5 (8.9)	33.2 (59.2)	25.0 (16.4)	5.0 (3.0)
EMS_9	n	1.5	2.2	8.1	1.1
EMS_49	n	8.9	12.9	11.9	14.9
EMS_50	n	1.5	4.5	16.7	4.4
EMS_91	y	37.7	665.9	36.8	0.2
EMS_93	y	102.2	587.1	41.5	0.3
EMS_134	y	33.4	200.7	22.7	3.5
EMS_246	y	19.9	87.9	54.1	0.9
EMS_259	n	1.5	1.5	6.6	4.3
EMS_265	n	1.5	8.3	15.2	3.9
EMS_279	n	4.9	11.9	35.3	0.6
EMS_333	y	38.3	73.3	27.9	1.4
EMS_336	n	12	7.6	29.5	6.4
EMS_355	y	12.0	71.6	26.9	1.4
EMS_395	y	14.0	136.5	66.2	0.5
EMS_479	n	1.5	11.8	15.7	3.9
EMS_595	y	32.4	81.6	29.4	2.7
EMS_605	n	2.8	3.1	21.7	4.1
EMS_611	y	18.3	104.3	36.9	0.8

Table 4.2: EMS phenotypic data for Morgan horses chosen for whole-genome sequencing. Horses were chosen to represent both insulin sensitive and insulin dysregulated horses (individuals with a history of laminitis and at the breed-specific upper range for the EMS biochemical measurements baseline insulin, insulin concentrations post oral sugar challenge, or triglycerides, and breed-specific lower range for adiponectin concentrations. Breed-specific ranges are provided for each biochemical measurement. Abbreviations: Insulin-OST: insulin concentrations post oral sugar test.

Table 4.3: Fine mapped high priority LD-ROI in the Welsh ponies							
Trait	**Chr**	**Min_ROI**	**Max_ROI**	**Total_SNPs**	**Effective_SNPs**	**Threshold**	**Sign_SNPs**
Insulin	5	35409104	44806458	16554	11488	4.35E-06	5
	8	69350844	75906595	21228	10497	4.76E-06	1
	15	5748377	6612684	1269	579	8.64E-05	0
	24	28451012	29887250	4942	935	5.35E-05	5
INS_OST	28	39322188	39488807	371	75	6.67E-04	15
GLU_OST	28	34271949	35138699	2036	415	1.20E-04	85
Adiponectin	1	171861236	178270042	20024	3837	1.30E-05	118
	18	60060215	61349045	3679	2237	2.24E-05	27
Leptin	5	39751797	50431769	22829	12790	3.91E-06	0
	6	488137	4012580	11480	6191	8.08E-06	9
	7	65678376	68117086	7144	3966	1.26E-05	45
NEFA	28	32909542	35703535	6227	1282	3.90E-05	30
ACTH	1	42944403	45232767	8200	1507	3.32E-05	17
	1	69558737	70960589	3940	652	7.67E-05	9
NH	4	67130904	69873296	9602	5077	9.85E-06	0
	4	77298241	81186565	13726	7639	6.55E-06	1
	4	83144842	83244842	263	167	2.99E-04	0
	7	93176991	93628686	2839	1869	2.68E-05	0
	9	32632235	37587269	16280	8508	5.88E-06	47
	11	18342117	19876247	2130	1222	4.09E-05	0
	14	63702522	63847210	457	291	1.72E-04	0
	20	40244007	41210876	2859	1518	3.29E-05	0
	20	60723014	61735694	5132	2566	1.95E-05	1
	21	5280993	6396786	3039	1568	3.19E-05	0

Table 4.3: Fine mapped high priority LD-ROI in the Welsh ponies (cont.)							
Trait	**Chr**	**Min_ROI**	**Max_ROI**	**Total_SNPs**	**Effective_SNPs**	**Threshold**	**Sign_SNPs**
NH (cont.)	21	19515280	25046226	15288	8217	6.08E-06	23
	24	31843480	36758218	11834	2330	2.15E-05	440
GH	1	132184772	133716124	4316	919	5.44E-05	1
	4	70026254	81648125	39176	21899	2.28E-06	113
	4	82570011	86366835	7965	4571	1.09E-05	7
	7	93191676	93628672	2690	1819	2.75E-05	2
	11	15414337	16451463	2056	1077	4.64E-05	19
	11	18613895	19317536	1001	510	9.80E-05	0
	18	79527484	81467661	5322	2904	1.72E-05	0
	19	31204596	31799125	1297	587	8.52E-05	0
	20	29486630	30976763	12762	4304	1.16E-05	6
	20	59464566	61015217	5688	3500	1.43E-05	0
	20	64722427	65336095	2754	1540	3.25E-05	0
	22	41032889	41066045	95	30	1.67E-03	0
LAM	2	35880861	36665473	1790	1033	4.84E-05	21
	19	57082025	62825378	15385	7954	6.29E-06	125
	28	9990892	10844823	2749	498	1.00E-04	8

Table 4.3: Fine mapped high priority LD-ROI in the Welsh ponies. Boundaries of the ROI were based on the breakdown of LD and are listed as the minimum SNP within (Min-ROI) and the maximum SNP (Max-ROI) within the region. The total number of SNP within the boundary, effective number of SNPs based on GEC, calculated Bonferroni corrected p-value for genome-wide significance (Threshold), and the total number of SNPs which exceeded that threshold (Sign_SNPs) are also listed in the table. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 4.4: Fine mapped high priority LD-ROI in the Morgan horses							
Trait	**Chr**	**Min_ROI**	**Max_ROI**	**Total_SNPs**	**Effective_SNPs**	**Threshold**	**Sign_SNPs**
INS-OST	10	71666607	73534053	4784	1827	2.74E-05	18
Glucose	4	17239374	19043831	6495	3156	1.58E-05	1
GLU-OST	3	55746338	58085997	4878	2474	2.02E-05	243
NEFA	1	184859013	187238015	5744	3045	1.64E-05	0
Adiponectin	2	16362904	18105119	3768	1936	2.58E-05	319
	4	34723398	39321960	9623	4353	1.15E-05	16
	6	32486287	32841880	547	167	2.99E-04	2
	6	64297403	71493047	17106	8812	5.67E-06	1
Leptin	19	51286493	53959028	8454	3631	1.38E-05	35
	24	25564765	29384679	11462	2104	2.38E-05	0
ACTH	1	82700933	84269783	3265	585	8.55E-05	1
	3	102944842	103801021	3179	1815	2.75E-05	0
	5	25378878	27689002	7546	3507	1.43E-05	0
NH	4	52024470	54237747	4057	2534	1.97E-05	0
	6	60410647	70570773	26417	13697	3.65E-06	1
	19	32962795	37391949	10479	5321	9.40E-06	16
GH	1	120644115	124691346	6147	1205	4.15E-05	31
	17	31806060	33720086	4769	1754	2.85E-05	0
LAM	4	17301415	19812653	9251	5032	9.94E-06	27
	12	32885279	34800986	4977	1358	3.68E-05	65
	14	87916190	91602875	10056	5294	9.44E-06	38
	18	30095266	35177011	11381	5767	8.67E-06	28
	19	30133826	30183826	140	112	4.46E-04	0
	22	2843476	5225020	6828	3351	1.49E-05	336

Table 4.4: Fine mapped high priority LD-ROI in the Morgan horses (cont.)							
Trait	**Chr**	**Min_ROI**	**Max_ROI**	**Total_SNPs**	**Effective_SNPs**	**Threshold**	**Sign_SNPs**
LAM (cont.)	23	7656404	12984095	14284	3368	1.48E-05	42

Table 4.4: Fine mapped high priority LD-ROI in the Morgan horses. Boundaries of the ROI were based on the breakdown of LD and are listed as the minimum SNP within (Min-ROI) and the maximum SNP (Max-ROI) within the region. The total number of SNP within the boundary, effective number of SNPs based on GEC, calculated Bonferroni corrected p-value for genome-wide significance (Threshold), and the total number of SNPs which exceeded that threshold (Sign_SNPs) are also listed in the table. Highlighted chromosomes (Chr) indicate regions which were shared with several traits.

Table 4.5: Selected fine-mapped high-priority LD-ROI for Welsh ponies						
Trait	Chr	Min_ROI	Max_ROI	Min_FMap	Max_FMap	Candidate Genes
Adiponectin	1	171861236	178270042	175000000	178000000	*SSTR1, SEC23A, FBX033, MIA2*
Leptin	7	65678376	68117086	67940000	67973000	None
GLU-OST/NEFA	28	34271949	35138699	34350000	35007000	*EIF3D, CSF2B, IFT27*
GH	11	15414337	16451463	15680000	16430000	*ACE, TACO1*
ACTH	1	69558737	70960589	70264921	70272614	None

Selected fine-mapped high-priority LD-ROI for Morgan horses						
Trait	Chr	Min_ROI	Max_ROI	Min_FMap	Max_FMap	
Glucose/ LAM	4	17239374	19043831	18320000	18390000	*ABCA13*
Insulin-OST	10	71666607	73534053	72936000	72946700	*NKAIN2*
Leptin	19	51286493	53959028	51580000	52450000	*BBX*
NH	19	32962795	37391949	33400000	35900000	*XXYLT1, BDHI*
Adiponectin	6	64297403	71493047	68250000	70001000	*NCKAP5L, GPD1, LIMA1, METTL7A, SLC11A2, AQP6*

Table 4.5: Selected fine-mapped high-priority LD-ROI for Welsh ponies and Morgan horses. Boundaries of the ROI were based on the breakdown of LD and are listed as the minimum SNP within (Min-ROI) and the maximum SNP (Max-ROI) within the region. For regions shared between traits, the smaller region size is listed. Fine-mapped boundaries (Min_FMap and Max_FMap) were based on the SNPs which exceeded the threshold for genome-wide significant. Candidate genes with a biological functional effect are listed under the Candidate Genes column.

Table 4.6: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
175126635	A	T	0.14	-0.59	0.12	2.88E-06	*PAX9*	UpStrm	Modifier
175126651	G	C	0.13	-0.57	0.13	9.40E-06	*PAX9*	UpStrm	Modifier
175129265	A	G	0.14	-0.56	0.12	6.06E-06	*PAX9*	UpStrm	Modifier
175155905	G	T	0.25	0.51	0.09	1.89E-08	*SLC25A21*	Intron	Modifier
175176688	T	G	0.13	-0.6	0.13	1.05E-05	*SLC25A21*	Intron	Modifier
175984052	C	T	0.18	-0.47	0.11	1.25E-05	*TTC6*	Intron	Modifier
176040706	T	C	0.18	-0.47	0.11	1.11E-05	*TTC6*	Intron	Modifier
176041751	C	T	0.12	-0.67	0.13	8.14E-07	*TTC6*	Intron	Modifier
176045167	G	C	0.12	-0.64	0.13	2.23E-06	*TTC6*	Intron	Modifier
176047376	G	A	0.12	-0.64	0.13	2.23E-06	*TTC6*	Intron	Modifier
176052940	C	T	0.11	-0.69	0.13	4.80E-07	*TTC6*	Intron	Modifier
176053778	T	C	0.11	-0.72	0.13	1.59E-07	*TTC6*	Intron	Modifier
176090641	T	C	0.25	0.43	0.1	8.82E-06	*TTC6*	Intron	Modifier
176092018	G	A	0.11	-0.72	0.13	9.22E-08	*TTC6*	Intron	Modifier
176194157	A	G	0.16	-0.56	0.12	1.91E-06	NA	Intergenic	Modifier
176195007	C	T	0.11	-0.68	0.13	7.02E-07	NA	Intergenic	Modifier
176196276	G	A	0.12	-0.65	0.13	1.96E-06	NA	Intergenic	Modifier
176228400	G	A	0.12	-0.68	0.13	5.83E-07	NA	Intergenic	Modifier
176229098	G	A	0.12	-0.65	0.13	1.60E-06	NA	Intergenic	Modifier
176229728	T	C	0.12	-0.65	0.13	1.60E-06	NA	Intergenic	Modifier
176230514	G	A	0.12	-0.68	0.13	5.83E-07	NA	Intergenic	Modifier
176254487	A	C	0.14	-0.55	0.12	1.12E-05	NA	Intergenic	Modifier
176310938	C	T	0.12	-0.64	0.13	2.04E-06	NA	Intergenic	Modifier
176311729	C	A	0.12	-0.65	0.13	1.60E-06	NA	Intergenic	Modifier
176329655	T	G	0.12	-0.65	0.13	1.60E-06	NA	Intergenic	Modifier
176334101	T	A	0.12	-0.65	0.13	1.60E-06	Pseudo	DwnStrm	Modifier
176335046	G	A	0.12	-0.65	0.13	1.60E-06	Pseudo	DwnStrm	Modifier
176337932	T	G	0.12	-0.65	0.13	1.60E-06	Pseudo	DwnStrm	Modifier
176338697	T	C	0.12	-0.66	0.13	4.26E-07	lncRNA	Intragenic	Modifier
176339665	T	C	0.12	-0.68	0.13	5.83E-07	lncRNA	Intragenic	Modifier
176340926	G	A	0.11	-0.67	0.13	1.24E-06	Pseudo	Intron	Modifier
176341057	A	G	0.12	-0.65	0.13	1.60E-06	Pseudo	Intron	Modifier
176400716	A	G	0.11	-0.68	0.14	2.02E-06	NA	Intergenic	Modifier
176432691	G	A	0.11	-0.68	0.13	4.52E-07	NA	Intergenic	Modifier
176457823	G	C	0.11	-0.64	0.14	9.79E-06	*SSTR1*	DwnStrm	Modifier
176464170	G	A	0.1	-0.65	0.14	1.15E-05	NA	Intergenic	Modifier
176477297	C	T	0.12	-0.75	0.13	2.22E-08	NA	Intergenic	Modifier
176477415	A	G	0.12	-0.75	0.13	2.22E-08	NA	Intergenic	Modifier
176524969	T	C	0.12	-0.72	0.13	7.07E-08	Pseudo	DwnStrm	Modifier

Table 4.6: SNPs which exceed the threshold for genome-wide significance (cont.)									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
176971726	C	G	0.12	-0.67	0.13	2.86E-07	NA	Intergenic	Modifier
176989550	A	G	0.11	-0.67	0.13	8.63E-07	NA	Intergenic	Modifier
176992467	C	T	0.11	-0.7	0.13	2.99E-07	NA	Intergenic	Modifier
177043601	C	A	0.11	-0.65	0.13	2.26E-06	*SEC23A*	Intron	Modifier
177044101	T	A	0.11	-0.65	0.13	2.26E-06	*SEC23A*	Intron	Modifier
177044819	C	T	0.11	-0.62	0.13	5.97E-06	*SEC23A*	Intron	Modifier
177049977	C	T	0.11	-0.6	0.14	1.23E-05	*SEC23A*	Intron	Modifier
177050619	T	C	0.11	-0.63	0.14	4.79E-06	*SEC23A*	Intron	Modifier
177072407	A	G	0.11	-0.74	0.14	1.02E-07	*GEMIN2*	UpStrm	Modifier
177152727	T	C	0.11	-0.66	0.14	2.07E-06	*MIA2*	5'UTRSC	Low
177157052	A	G	0.11	-0.63	0.14	5.55E-06	*MIA2*	Intron	Modifier
177164043	A	G	0.11	-0.7	0.14	4.43E-07	*MIA2*	Intron	Modifier
177177259	T	C	0.12	-0.66	0.13	1.58E-06	*MIA2*	Intron	Modifier
177177547	T	G	0.12	-0.66	0.13	1.58E-06	*MIA2*	Intron	Modifier
177178441	G	C	0.11	-0.69	0.13	5.64E-07	*MIA2*	Intron	Modifier
177178529	C	T	0.12	-0.73	0.13	5.15E-08	*MIA2*	Intron	Modifier
177178901	C	T	0.11	-0.63	0.14	5.55E-06	*MIA2*	Intron	Modifier
177180240	G	T	0.11	-0.66	0.14	1.90E-06	*MIA2*	Intron	Modifier
177180649	T	G	0.11	-0.63	0.14	5.06E-06	*MIA2*	Intron	Modifier
177188749	T	C	0.11	-0.7	0.14	6.91E-07	*MIA2*	Intron	Modifier
177189044	C	T	0.11	-0.63	0.14	7.82E-06	*MIA2*	Intron	Modifier
177200973	C	T	0.11	-0.63	0.14	7.82E-06	*MIA2*	Intron	Modifier
177208086	C	G	0.15	-0.53	0.12	9.00E-06	*MIA2*	Intron	Modifier
177208493	G	A	0.11	-0.61	0.14	1.08E-05	*MIA2*	Intron	Modifier
177212329	C	T	0.12	-0.61	0.13	3.58E-06	*MIA2*	Intron	Modifier
177219468	A	C	0.15	-0.55	0.12	4.12E-06	*MIA2*	Intron	Modifier
177223031	G	C	0.16	-0.51	0.11	1.22E-05	*MIA2*	Intron	Modifier
177234592	G	A	0.12	-0.66	0.13	1.58E-06	*MIA2*	DwnStrm	Modifier
177238090	T	C	0.11	-0.69	0.13	5.64E-07	*MIA2*	DwnStrm	Modifier
177241306	A	G	0.11	-0.69	0.13	5.64E-07	*MIA2*	DwnStrm	Modifier
177241680	A	G	0.15	-0.59	0.12	6.98E-07	*MIA2*	DwnStrm	Modifier
177242008	C	T	0.12	-0.7	0.13	1.44E-07	*MIA2*	DwnStrm	Modifier
177244749	T	G	0.12	-0.66	0.13	1.58E-06	*MIA2*	DwnStrm	Modifier
177244871	C	T	0.12	-0.61	0.13	6.57E-06	*MIA2*	DwnStrm	Modifier
177244882	T	A	0.12	-0.61	0.13	6.57E-06	*MIA2*	DwnStrm	Modifier
177268812	C	A	0.12	-0.66	0.13	5.77E-07	*FBXO33*	Intron	Modifier
177271487	T	C	0.12	-0.62	0.13	5.61E-06	*FBXO33*	Intron	Modifier
177272600	T	C	0.12	-0.58	0.13	1.22E-05	*FBXO33*	Intron	Modifier
177273743	A	G	0.12	-0.64	0.13	1.55E-06	*FBXO33*	Intron	Modifier

Table 4.6: SNPs which exceed the threshold for genome-wide significance (cont.)									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
177275526	A	G	0.16	-0.54	0.12	4.47E-06	*FBXO33*	Intron	Modifier
177276009	C	T	0.16	-0.61	0.11	1.37E-07	*FBXO33*	Intron	Modifier
177279813	A	G	0.12	-0.62	0.13	5.61E-06	*FBXO33*	Intron	Modifier
177283592	G	A	0.16	-0.54	0.12	4.47E-06	*FBXO33*	Intron	Modifier
177284558	G	A	0.16	-0.54	0.12	4.47E-06	*FBXO33*	Intron	Modifier
177296572	A	G	0.11	-0.69	0.14	8.76E-07	*FBXO33*	Intron	Modifier
177297230	T	A	0.12	-0.66	0.13	1.58E-06	*FBXO33*	Intron	Modifier
177446902	G	A	0.13	-0.6	0.13	5.98E-06	NA	Intergenic	Modifier
177461338	G	A	0.13	-0.6	0.13	5.98E-06	NA	Intergenic	Modifier
177476511	C	T	0.11	-0.63	0.14	7.55E-06	NA	Intergenic	Modifier
177478669	C	T	0.11	-0.67	0.14	1.88E-06	NA	Intergenic	Modifier
177479850	C	T	0.11	-0.61	0.14	1.05E-05	NA	Intergenic	Modifier
177480793	G	T	0.11	-0.61	0.14	1.05E-05	NA	Intergenic	Modifier
177577158	C	T	0.15	-0.56	0.12	2.97E-06	NA	Intergenic	Modifier
177577662	T	C	0.15	-0.57	0.12	4.88E-06	NA	Intergenic	Modifier
177585901	G	A	0.16	-0.56	0.12	4.26E-06	NA	Intergenic	Modifier
177588448	C	G	0.15	-0.55	0.12	6.85E-06	NA	Intergenic	Modifier
177589015	C	G	0.14	-0.56	0.12	5.23E-06	NA	Intergenic	Modifier
177590852	C	T	0.09	-0.67	0.15	9.33E-06	NA	Intergenic	Modifier
177591148	C	G	0.14	-0.54	0.12	1.25E-05	NA	Intergenic	Modifier
177592325	T	G	0.15	-0.54	0.12	1.08E-05	NA	Intergenic	Modifier
177592444	A	G	0.14	-0.59	0.12	2.02E-06	NA	Intergenic	Modifier
177592456	T	C	0.15	-0.55	0.12	8.74E-06	NA	Intergenic	Modifier
177592747	G	A	0.15	-0.57	0.12	4.18E-06	NA	Intergenic	Modifier
177593097	G	A	0.14	-0.57	0.12	4.24E-06	NA	Intergenic	Modifier
177593426	C	T	0.15	-0.57	0.12	1.94E-06	NA	Intergenic	Modifier
177593432	T	C	0.15	-0.6	0.12	6.15E-07	NA	Intergenic	Modifier
177593843	A	T	0.14	-0.55	0.12	1.06E-05	NA	Intergenic	Modifier
177800543	T	G	0.15	-0.59	0.12	2.53E-06	NA	Intergenic	Modifier
177806688	A	G	0.15	-0.55	0.12	1.27E-05	NA	Intergenic	Modifier
177808650	T	C	0.15	-0.53	0.12	1.23E-05	NA	Intergenic	Modifier
177952363	A	G	0.18	0.49	0.11	1.12E-05	Pseudo	Intron	Modifier
177953647	A	T	0.21	0.46	0.1	8.51E-06	Pseudo	Intron	Modifier
177954557	G	A	0.22	0.48	0.1	2.80E-06	Pseudo	Intron	Modifier
177954891	C	G	0.21	-0.51	0.1	1.95E-06	Pseudo	Intron	Modifier
177960815	A	G	0.21	0.48	0.1	7.70E-06	Pseudo	Intron	Modifier
177961297	C	T	0.21	0.52	0.1	8.37E-07	Pseudo	Intron	Modifier
177966158	A	G	0.21	0.49	0.1	3.86E-06	Pseudo	Intron	Modifier
177970858	T	C	0.21	0.55	0.1	2.13E-07	Pseudo	Intron	Modifier

Table 4.6: Fine-mapped coding SNPs with a p-value <0.05									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Effect**	**Impact**
175683703	C	G	0.16	0.29	0.11	1.06E-02	*MIPOL1*	Miss	Moderate
175683752	A	G	0.16	0.29	0.11	1.06E-02	*MIPOL1*	Miss	Moderate
175683836	A	C	0.14	0.27	0.11	1.74E-02	*MIPOL1*	Miss	Moderate
175683840	G	C	0.14	0.27	0.11	1.74E-02	*MIPOL1*	Miss	Moderate
175899508	A	G	0.44	0.25	0.09	3.43E-03	*MIPOL1*	Miss	Moderate
176036945	T	G	0.41	-0.26	0.08	1.90E-03	*TTC6*	Miss	Moderate
176037039	C	T	0.02	0.71	0.35	4.35E-02	*TTC6*	Miss	Moderate
176059721	C	T	0.48	0.24	0.08	4.69E-03	*TTC6*	Synon	Low
176080766	A	G	0.07	0.36	0.18	4.50E-02	*TTC6*	Miss	Moderate
176086066	C	T	0.29	0.3	0.09	8.57E-04	*TTC6*	Miss	Moderate
176092396	A	G	0.07	0.39	0.17	2.47E-02	*TTC6*	Miss	Moderate
176117140	C	T	0.29	-0.24	0.09	1.13E-02	*TTC6*	Synon	Low
176146547	C	T	0.5	-0.19	0.09	2.74E-02	*TTC6*	Synon	Low
177083177	A	G	0.07	0.38	0.15	1.34E-02	*GEMIN2*	Miss	Moderate
177167367	A	G	0.04	0.51	0.21	1.54E-02	*MIA2*	Miss	Moderate

Table 4.6: Summary table for SNPs which exceeded the threshold for genome-wide significance, and fine-mapped coding SNPs with a predicted low, moderate, or high impact (genome-wide significance>p-value <0.05) for the adiponectin concentrations on chromosome 1 in the Welsh ponies. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, and location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff. Additional abbreviations: lncRNA (long noncoding RNA), Pseudo (pseudogene), UpStm (upstream), DwnStrm (downstream), Synon (synonymous), Miss (missense), 5'UTRSG (gain of a 5' UTR premature start codon).

Table 4.7: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
67940623	T	G	0.36	-0.33	0.07	2.67E-06	*TENM4*	Intron	Modifier
67955613	C	A	0.45	-0.33	0.07	1.08E-06	*TENM4*	Intron	Modifier
67956000	G	A	0.46	-0.32	0.07	2.81E-06	*TENM4*	Intron	Modifier
67956021	G	A	0.44	0.33	0.07	1.33E-06	*TENM4*	Intron	Modifier
67956267	T	G	0.46	-0.34	0.07	1.38E-06	*TENM4*	Intron	Modifier
67956386	A	G	0.44	0.34	0.07	6.65E-07	*TENM4*	Intron	Modifier
67956667	G	A	0.46	-0.31	0.07	4.90E-06	*TENM4*	Intron	Modifier
67957167	G	T	0.45	-0.33	0.07	1.96E-06	*TENM4*	Intron	Modifier
67957813	G	A	0.45	-0.33	0.07	1.43E-06	*TENM4*	Intron	Modifier
67957957	C	G	0.45	-0.32	0.07	4.84E-06	*TENM4*	Intron	Modifier
67958411	G	A	0.44	-0.37	0.07	1.11E-07	*TENM4*	Intron	Modifier
67958767	G	A	0.45	-0.33	0.07	1.28E-06	*TENM4*	Intron	Modifier
67958804	G	C	0.44	-0.35	0.07	4.97E-07	*TENM4*	Intron	Modifier
67959095	G	A	0.45	-0.35	0.07	5.54E-07	*TENM4*	Intron	Modifier
67960089	A	C	0.46	-0.35	0.07	4.78E-07	*TENM4*	Intron	Modifier
67960134	G	A	0.45	-0.36	0.07	1.93E-07	*TENM4*	Intron	Modifier
67960495	C	T	0.46	-0.35	0.07	5.20E-07	*TENM4*	Intron	Modifier
67961971	T	G	0.46	-0.34	0.07	6.16E-07	*TENM4*	Intron	Modifier
67962121	A	C	0.44	0.35	0.07	3.22E-07	*TENM4*	Intron	Modifier
67962164	G	A	0.44	-0.34	0.07	8.01E-07	*TENM4*	Intron	Modifier
67962184	C	G	0.45	-0.35	0.07	3.90E-07	*TENM4*	Intron	Modifier
67962366	G	A	0.45	-0.36	0.07	1.41E-07	*TENM4*	Intron	Modifier
67962419	C	G	0.45	-0.37	0.07	1.06E-07	*TENM4*	Intron	Modifier
67964359	C	A	0.31	-0.32	0.07	5.53E-06	*TENM4*	Intron	Modifier
67964429	G	A	0.47	-0.35	0.07	2.96E-07	*TENM4*	Intron	Modifier
67965165	G	A	0.44	-0.35	0.07	3.33E-07	*TENM4*	Intron	Modifier
67965588	G	A	0.43	-0.36	0.07	7.08E-08	*TENM4*	Intron	Modifier
67966006	C	T	0.45	-0.32	0.07	3.48E-06	*TENM4*	Intron	Modifier
67966058	G	T	0.44	-0.34	0.07	4.50E-07	*TENM4*	Intron	Modifier
67967062	G	A	0.46	-0.35	0.07	4.38E-07	*TENM4*	Intron	Modifier
67967364	C	T	0.46	-0.34	0.07	5.13E-07	*TENM4*	Intron	Modifier
67967431	G	A	0.46	-0.35	0.07	2.60E-07	*TENM4*	Intron	Modifier
67969434	C	T	0.46	-0.32	0.07	2.31E-06	*TENM4*	Intron	Modifier
67969982	T	C	0.46	-0.32	0.07	5.70E-06	*TENM4*	Intron	Modifier
67970467	A	G	0.46	-0.33	0.07	2.25E-06	*TENM4*	Intron	Modifier
67970470	T	C	0.46	-0.32	0.07	4.38E-06	*TENM4*	Intron	Modifier
67970496	A	G	0.45	-0.31	0.07	5.39E-06	*TENM4*	Intron	Modifier
67970505	T	C	0.46	-0.31	0.07	6.34E-06	*TENM4*	Intron	Modifier
67970532	A	T	0.44	-0.32	0.07	3.71E-06	*TENM4*	Intron	Modifier

Table 4.7: SNPs which exceed the threshold for genome-wide significance (cont.)									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
67970581	G	A	0.44	-0.31	0.07	1.23E-05	*TENM4*	Intron	Modifier
67970792	T	C	0.45	-0.34	0.07	8.29E-07	*TENM4*	Intron	Modifier
67970877	G	A	0.44	-0.35	0.07	4.39E-07	*TENM4*	Intron	Modifier
67970907	A	T	0.46	-0.34	0.07	6.63E-07	*TENM4*	Intron	Modifier
67971020	G	C	0.43	-0.37	0.07	5.55E-08	*TENM4*	Intron	Modifier
67971042	G	A	0.44	-0.36	0.07	1.71E-07	*TENM4*	Intron	Modifier
67971228	C	T	0.46	-0.32	0.07	1.82E-06	*TENM4*	Intron	Modifier

Fine-mapped coding SNPs with a p-value <0.05									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Effect**	**Impact**
67610664	T	C	0.05	-0.43	0.15	4.06E-03	*TENM4*	SS	Low

Table 4.7: Summary table for SNPs which exceeded the threshold for genome-wide significance, and fine-mapped coding SNPs with a predicted low, moderate, or high impact (genome-wide significance>p-value <0.05) for leptin concentrations on chromosome 7 in the Welsh ponies. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, and location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff. Additional abbreviations: SS (splice site variant).

Table 4.8: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
34358987	A	G	0.11	-11.31	2.69	3.77E-05	*RBFOX2*	Intron	Modifier
34360156	T	C	0.12	-11.22	2.52	1.43E-05	*RBFOX2*	Intron	Modifier
34372193	T	G	0.12	-10.49	2.67	1.14E-04	*RBFOX2*	Intron	Modifier
34416581	G	A	0.17	-8.79	2.2	8.80E-05	*RBFOX2*	Intron	Modifier
34418871	C	T	0.17	-8.6	2.19	1.15E-04	*RBFOX2*	Intron	Modifier
34420571	C	T	0.17	-8.96	2.25	9.15E-05	*RBFOX2*	Intron	Modifier
34425365	T	G	0.21	-9.54	2.06	6.25E-06	*RBFOX2*	UpStrm	Modifier
34425522	G	C	0.21	-8.85	2.05	2.43E-05	*RBFOX2*	UpStrm	Modifier
34427465	T	C	0.2	-8.73	2.06	3.42E-05	*RBFOX2*	Intron	Modifier
34427606	A	C	0.2	-8.69	2.06	3.79E-05	*RBFOX2*	Intron	Modifier
34428714	T	C	0.16	-8.92	2.25	1.05E-04	*RBFOX2*	Intron	Modifier
34431090	A	G	0.22	-11	1.89	2.08E-08	*RBFOX2*	Intron	Modifier
34432082	C	A	0.22	-10.96	1.89	2.28E-08	*RBFOX2*	Intron	Modifier
34432531	C	T	0.17	-8.96	2.25	9.15E-05	*RBFOX2*	Intron	Modifier
34434081	T	C	0.22	-10.43	1.87	7.54E-08	*RBFOX2*	Intron	Modifier
34436070	A	G	0.22	-10.51	1.87	6.03E-08	*RBFOX2*	Intron	Modifier
34436211	C	T	0.22	-10.37	1.87	9.22E-08	*RBFOX2*	Intron	Modifier
34436355	G	A	0.13	-11.75	2.32	9.24E-07	*RBFOX2*	Intron	Modifier
34436762	T	C	0.13	-10.96	2.39	7.87E-06	*RBFOX2*	Intron	Modifier
34437058	G	A	0.14	-11.67	2.23	4.25E-07	*RBFOX2*	Intron	Modifier
34438111	G	A	0.14	-9.33	2.35	9.53E-05	*RBFOX2*	Intron	Modifier
34442061	G	A	0.17	8.62	2.2	1.19E-04	*RBFOX2*	Intron	Modifier
34445159	A	G	0.02	-26.06	5.92	1.73E-05	*RBFOX2*	Intron	Modifier
34446378	C	T	0.07	-12.54	3.18	1.11E-04	*RBFOX2*	Intron	Modifier
34447894	T	C	0.06	-13.64	3.4	8.38E-05	*RBFOX2*	Intron	Modifier
34453800	C	T	0.16	-8.91	2.26	1.12E-04	*RBFOX2*	Intron	Modifier
34454204	C	G	0.17	-8.96	2.25	9.15E-05	*RBFOX2*	Intron	Modifier
34457285	C	A	0.16	-8.92	2.25	1.05E-04	*RBFOX2*	Intron	Modifier
34457663	T	C	0.16	-8.92	2.25	1.04E-04	*RBFOX2*	Intron	Modifier
34460259	G	A	0.15	-11.86	2.19	1.65E-07	*RBFOX2*	Intron	Modifier
34462545	A	G	0.17	-9.13	2.2	4.92E-05	*RBFOX2*	Intron	Modifier
34464387	C	A	0.15	-11.86	2.19	1.65E-07	*RBFOX2*	Intron	Modifier
34475212	G	T	0.19	-8.64	2.16	8.69E-05	*RBFOX2*	Intron	Modifier
34475515	G	A	0.19	-9.03	2.11	2.88E-05	*RBFOX2*	Intron	Modifier
34475901	A	G	0.19	-9.3	2.19	3.18E-05	*RBFOX2*	Intron	Modifier
34476507	T	C	0.18	-8.81	2.23	1.05E-04	*RBFOX2*	Intron	Modifier
34478054	T	C	0.17	-9	2.2	6.20E-05	*RBFOX2*	Intron	Modifier
34480902	C	G	0.16	-9.07	2.27	8.70E-05	*RBFOX2*	Intron	Modifier
34482095	A	G	0.16	8.9	2.27	1.20E-04	*RBFOX2*	Intron	Modifier

Table 4.8: SNPs which exceed the threshold for genome-wide significance (cont.)									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
34486721	A	G	0.16	8.93	2.27	1.14E-04	*RBFOX2*	UpStrm	Modifier
34493686	A	G	0.16	9.37	2.27	5.25E-05	*RBFOX2*	Intron	Modifier
34494511	T	G	0.16	8.9	2.27	1.20E-04	*RBFOX2*	Intron	Modifier
34498254	T	G	0.17	9.21	2.24	5.80E-05	*RBFOX2*	Intron	Modifier
34499635	C	G	0.13	-11.06	2.32	3.39E-06	*RBFOX2*	Intron	Modifier
34507888	A	G	0.16	-9.54	2.27	3.74E-05	*RBFOX2*	Intron	Modifier
34511355	A	G	0.16	-9.11	2.26	7.65E-05	*RBFOX2*	Intron	Modifier
34512244	C	T	0.13	-11.06	2.32	3.39E-06	*RBFOX2*	Intron	Modifier
34521603	C	T	0.16	9.1	2.21	5.47E-05	*RBFOX2*	Intron	Modifier
34529874	T	C	0.14	-10.76	2.31	5.80E-06	*RBFOX2*	Intron	Modifier
34534466	C	G	0.16	-9.19	2.25	6.33E-05	*RBFOX2*	Intron	Modifier
34548667	G	T	0.16	-9.06	2.27	9.02E-05	Pseudo	UpStrm	Modifier
34586132	T	C	0.13	-11.06	2.32	3.39E-06	NA	Intergenic	Modifier
34589999	C	T	0.14	-10.44	2.27	7.29E-06	NA	Intergenic	Modifier
34595199	A	G	0.14	-11.07	2.29	2.71E-06	Pseudo	DwnStrm	Modifier
34602058	G	A	0.16	10.57	2.24	4.29E-06	Pseudo	DwnStrm	Modifier
34603017	C	T	0.14	-11.06	2.24	1.69E-06	Pseudo	Intron	Modifier
34603686	C	G	0.2	8.26	2.06	8.44E-05	Pseudo	Intragenic	Modifier
34605110	A	G	0.11	-11.31	2.44	6.25E-06	Pseudo	Intron	Modifier
34930591	T	A	0.09	-15.19	2.83	2.10E-07	NA	Intergenic	Modifier
34935008	T	G	0.09	-14.75	2.81	3.77E-07	*FOXRED2*	DwnStrm	Modifier
34935996	T	G	0.09	-15.19	2.83	2.10E-07	*FOXRED2*	DwnStrm	Modifier
34936212	A	C	0.09	-14.75	2.81	3.77E-07	*FOXRED2*	DwnStrm	Modifier
34937389	C	T	0.09	-14.1	2.86	1.69E-06	*FOXRED2*	DwnStrm	Modifier
34937614	C	G	0.09	-14.75	2.81	3.77E-07	*FOXRED2*	DwnStrm	Modifier
34938154	T	C	0.09	-14.75	2.81	3.77E-07	*FOXRED2*	3'UTR	Modifier
34938425	T	C	0.09	-14.75	2.81	3.77E-07	*FOXRED2*	3'UTR	Modifier
34953396	T	C	0.09	-15.19	2.83	2.10E-07	*FOXRED2*	UpStrm	Modifier
34953715	T	A	0.08	-14.54	2.88	9.82E-07	*FOXRED2*	UpStrm	Modifier
34953977	A	G	0.08	-14.54	2.88	9.82E-07	*FOXRED2*	UpStrm	Modifier
34960948	C	A	0.09	-15.17	2.8	1.62E-07	*EIF3D*	Intron	Modifier
34961593	C	T	0.08	-15.58	2.86	1.40E-07	*EIF3D*	Intron	Modifier
34961984	T	A	0.09	-16	2.77	2.83E-08	*EIF3D*	Intron	Modifier
34963692	C	T	0.09	-14.52	2.79	4.41E-07	*EIF3D*	Intron	Modifier
34964715	A	G	0.09	-15.19	2.83	2.10E-07	*EIF3D*	Intron	Modifier
34965892	A	G	0.09	-15.27	2.83	1.80E-07	*EIF3D*	Intron	Modifier
34965910	A	G	0.09	-15.27	2.83	1.80E-07	*EIF3D*	Intron	Modifier
34966194	G	T	0.08	-15.74	2.85	9.64E-08	*EIF3D*	Intron	Modifier
34979591	T	C	0.09	-14.51	2.85	7.85E-07	NA	Intergenic	Modifier

Table 4.8: SNPs which exceed the threshold for genome-wide significance (cont.)									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
34993332	G	C	0.15	-9.7	2.32	4.33E-05	*CACNG2*	DwnStrm	Modifier
34993431	A	C	0.15	-9.92	2.3	2.42E-05	*CACNG2*	DwnStrm	Modifier
34993538	T	C	0.15	-9.92	2.3	2.42E-05	*CACNG2*	DwnStrm	Modifier
34993712	C	T	0.15	-10.23	2.3	1.40E-05	*CACNG2*	DwnStrm	Modifier
34993878	T	C	0.15	-10	2.33	2.71E-05	*CACNG2*	DwnStrm	Modifier
35003752	C	T	0.06	-16.86	3.42	1.64E-06	*CACNG2*	Intron	Modifier
35136611	T	C	0.03	-18.67	4.2	1.41E-05	NA	Intergenic	Modifier

Fine-mapped SNPs with predicted low, moderate or high impact and p-value <0.05									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Effect**	**Impact**
34728950	A	G	0.21	5.58	2.12	8.99E-03	lncRNA	SS	Low

Table 4.8: Summary table for SNPs which exceeded the threshold for genome-wide significance, and fine-mapped coding SNPs with a predicted low, moderate, or high impact (genome-wide significance>p-value <0.05) for glucose concentrations post oral sugar test on chromosome 28 in the Welsh ponies. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, and location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff. Additional abbreviations: SS (splice site variant), DwnStrm (downstream), UpStrm (upstream), pseudo (pseudogene), lncRNA (long non-coding RNA).

Table 4.9: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
34432243	T	C	0.04	0.18	0.04	5.31E-06	*RBFOX2*	Intron	Modifier
34434781	G	T	0.04	0.18	0.04	5.31E-06	*RBFOX2*	Intron	Modifier
34434985	T	C	0.04	0.14	0.03	3.14E-05	*RBFOX2*	Intron	Modifier
34445042	C	T	0.04	0.16	0.04	1.08E-05	*RBFOX2*	Intron	Modifier
34446271	A	T	0.04	0.17	0.04	1.89E-06	*RBFOX2*	Intron	Modifier
34449423	T	A	0.04	-0.18	0.04	8.84E-06	*RBFOX2*	Intron	Modifier
34465152	G	A	0.04	0.17	0.04	4.26E-06	*RBFOX2*	Intron	Modifier
34466939	A	C	0.04	0.19	0.04	1.26E-06	*RBFOX2*	Intron	Modifier
34852931	C	T	0.03	-0.2	0.04	4.87E-06	NA	Intergenic	Modifier
34876505	C	T	0.22	-0.08	0.02	8.48E-06	Pseudo	DwnStrm	Modifier
34876660	C	T	0.23	-0.08	0.02	1.88E-05	Pseudo	DwnStrm	Modifier
34876768	A	C	0.22	-0.08	0.02	3.83E-06	Pseudo	DwnStrm	Modifier
34876773	G	A	0.22	-0.09	0.02	3.33E-06	Pseudo	DwnStrm	Modifier
34877001	C	T	0.22	-0.08	0.02	9.92E-06	Pseudo	DwnStrm	Modifier
34877033	A	G	0.22	-0.08	0.02	2.29E-05	Pseudo	DwnStrm	Modifier
34877252	C	A	0.23	-0.08	0.02	6.69E-06	Pseudo	DwnStrm	Modifier
34877400	G	A	0.23	-0.08	0.02	6.27E-06	Pseudo	DwnStrm	Modifier
34877681	T	C	0.22	-0.08	0.02	7.52E-06	Pseudo	DwnStrm	Modifier
34877683	G	T	0.22	-0.08	0.02	7.52E-06	Pseudo	DwnStrm	Modifier
34877723	G	T	0.22	-0.08	0.02	6.72E-06	Pseudo	DwnStrm	Modifier
35274846	A	C	0.5	-0.06	0.01	3.74E-05	*NCF4*	DwnStrm	Modifier
35287243	C	A	0.47	-0.06	0.01	2.12E-05	NA	Intergenic	Modifier
35288833	G	A	0.46	0.06	0.01	2.39E-05	NA	Intergenic	Modifier
35288838	A	G	0.46	0.06	0.01	2.39E-05	NA	Intergenic	Modifier
35288863	G	A	0.46	0.06	0.01	2.95E-05	NA	Intergenic	Modifier

Fine-mapped coding SNPs with a p-value <0.05									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Effect**	**Impact**
34284858	C	T	0.06	0.08	0.03	7.45E-03	*RBFOX2*	Miss	Moderate
34702062	G	A	0.01	-0.17	0.06	5.92E-03	Pseudo	Synon	Low
34728950	A	G	0.2	-0.05	0.02	1.97E-02	Pseudo	SS	Low
34922478	C	T	0.11	-0.08	0.02	1.64E-04	*TXN2*	Miss	Moderate
35164349	C	T	0.11	0.06	0.02	7.13E-03	*IFT27*	Miss	Moderate
35326823	T	C	0.08	0.06	0.03	2.43E-02	*CSF2RB*	SS	Low
35337318	A	G	0.08	0.06	0.03	2.43E-02	*CSF2RB*	Synon	Low
35349453	G	T	0.07	0.07	0.03	1.18E-02	Pseudo	Miss	Moderate
35349461	C	A	0.07	0.07	0.03	1.18E-02	Pseudo	Synon	Low
35349463	C	G	0.07	0.07	0.03	1.18E-02	Pseudo	Miss	Moderate

Table 4.9: Summary table for SNPs which exceeded the threshold for genome-wide significance, and fine-mapped coding SNPs with a predicted low, moderate, or high impact (genome-wide significance>p-value <0.05) for NEFA concentrations on chromosome 28 in the Welsh ponies. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, and location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff. Additional abbreviations: SS (splice site variant), DwnStrm (downstream), UpStrm (upstream), pseudo (pseudogene), lncRNA (long non-coding RNA).

Table 4.10: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
15700606	A	C	0.12	-0.04	0.01	1.96E-05	*MAP3K3*	Intron	Modifier
15700819	A	G	0.23	-0.03	0.01	2.30E-06	*MAP3K3*	Intron	Modifier
15706236	A	G	0.22	-0.03	0.01	3.73E-05	*MAP3K3*	Intron	Modifier
15714161	T	G	0.26	-0.03	0.01	1.30E-05	*MAP3K3*	UpStrm	Modifier
15719310	C	T	0.13	-0.04	0.01	4.34E-05	*TACO1*	Intron	Modifier
15722574	C	G	0.12	-0.04	0.01	3.40E-05	*TACO1*	Intron	Modifier
15729021	C	A	0.15	-0.04	0.01	2.06E-05	*TACO1*	UpStrm	Modifier
15729719	G	A	0.18	-0.03	0.01	2.68E-05	*DCAF7*	DwnStrm	Modifier
15730062	C	T	0.16	-0.03	0.01	4.23E-05	*DCAF7*	3'UTR	Modifier
15731895	A	G	0.16	-0.04	0.01	4.09E-06	*DCAF7*	3'UTR	Modifier
15732366	G	A	0.16	-0.03	0.01	3.58E-05	*DCAF7*	3'UTR	Modifier
15748734	G	A	0.18	-0.03	0.01	3.96E-05	*DCAF7*	Intron	Modifier
15753271	T	C	0.08	-0.05	0.01	4.24E-05	Pseudo	UpStrm	Modifier
15805251	A	G	0.07	-0.05	0.01	2.42E-05	*ACE*	Intron	Modifier
15818270	C	T	0.08	-0.05	0.01	3.46E-05	*ACE*	UpStrm	Modifier
15818753	G	T	0.08	-0.05	0.01	4.25E-05	*ACE*	UpStrm	Modifier
15821295	C	T	0.08	-0.05	0.01	3.46E-05	*ACE*	Intron	Modifier
16393525	G	A	0.37	-0.03	0.01	3.22E-05	*MARCH10*	Intron	Modifier
16403791	G	A	0.04	-0.06	0.01	4.04E-05	*MARCH10*	Intron	Modifier

Fine-mapped SNPs with predicted low, moderate or high impact and p-value <0.05									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Effect**	**Impact**
15720639	G	T	0.12	-0.03	0.01	5.62E-04	*TACO1*	Synon	Low
15724068	C	A	0.21	0.02	0.01	6.71E-03	*TACO1*	5'UTRSC	Low
15738303	A	G	0.12	-0.03	0.01	4.93E-04	*DCAF7*	Synon	Low
15792415	G	A	0.21	0.02	0.01	4.81E-03	lncRNA	Missen	Moderate
15816181	C	T	0.21	0.02	0.01	1.11E-02	*ACE*	Missen	Moderate
16375354	G	C	0.06	-0.03	0.01	2.39E-02	*MARCH10*	Missen	Moderate

Table 4.10: Summary table for SNPs which exceeded the threshold for genome-wide significance, and fine-mapped coding SNPs with a predicted low, moderate, or high impact (genome-wide significance>p-value <0.05) for GH ratios on chromosome 11 in the Welsh ponies. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, and location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff. Additional abbreviations: Synon (synonymous), DwnStrm (downstream), UpStrm (upstream), pseudo (pseudogene), 5'UTRSC (start codon gained in the 5'UTR), lncRNA (long non-coding RNA).

Table 4.11: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
70264921	C	T	0.01	-0.34	0.08	6.35E-05	NA	Intergenic	Modifier
70264929	C	A	0.01	-0.34	0.08	6.35E-05	NA	Intergenic	Modifier
70264944	T	C	0.01	-0.34	0.08	6.35E-05	NA	Intergenic	Modifier
70266067	A	G	0.04	-0.19	0.05	3.34E-05	NA	Intergenic	Modifier
70269910	G	A	0.01	-0.36	0.08	1.29E-05	NA	Intergenic	Modifier
70271957	G	A	0.02	-0.32	0.06	5.30E-07	NA	Intergenic	Modifier
70272029	T	C	0.01	-0.39	0.07	3.76E-07	NA	Intergenic	Modifier
70272056	A	G	0.01	-0.44	0.08	1.10E-07	NA	Intergenic	Modifier
70272614	C	T	0.01	-0.4	0.07	1.17E-07	NA	Intergenic	Modifier

Table 4.11: Summary table for SNPs which exceeded the threshold for genome-wide significance for ACTH concentrations on chromosome 1 in the Welsh ponies. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, and location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff.

Table 4.12: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
17432232	C	G	0.05	-0.31	0.06	3.73E-06	*TNS3*	Intron	Modifier
17432748	A	G	0.05	-0.38	0.06	2.70E-09	*TNS3*	Intron	Modifier
17433141	T	C	0.05	-0.35	0.06	4.09E-08	*TNS3*	Intron	Modifier
17433169	C	A	0.05	-0.38	0.06	2.70E-09	*TNS3*	Intron	Modifier
17433355	A	C	0.05	-0.38	0.06	2.70E-09	*TNS3*	Intron	Modifier
17433659	C	T	0.05	-0.38	0.06	9.72E-09	*TNS3*	Intron	Modifier
17468270	C	T	0.01	-0.56	0.12	2.72E-06	*TNS3*	Intron	Modifier
17468271	G	A	0.01	-0.56	0.12	2.72E-06	*TNS3*	Intron	Modifier
17775545	C	T	0.02	-0.61	0.11	1.80E-07	NA	Intergenic	Modifier
17775671	G	T	0.02	-0.54	0.11	1.36E-06	NA	Intergenic	Modifier
17870617	T	C	0.01	-0.67	0.14	4.77E-06	*PKD1L1*	Intron	Modifier
17987269	C	T	0.06	-0.28	0.06	4.60E-06	*HUS1*	Intron	Modifier
18003883	G	C	0.01	-0.71	0.13	1.20E-07	*HUS1*	Intron	Modifier
18006786	A	G	0.05	-0.28	0.06	8.69E-06	*HUS1*	Intron	Modifier
18365634	C	T	0.02	-0.44	0.09	2.67E-06	*ABCA13*	Intron	Modifier
18397812	C	T	0.01	-0.53	0.12	7.54E-06	*ABCA13*	Intron	Modifier
18397878	T	A	0.01	-0.66	0.14	3.91E-06	*ABCA13*	Intron	Modifier
18416711	G	A	0.02	-0.53	0.11	1.96E-06	*ABCA13*	Intron	Modifier
18416725	A	G	0.02	-0.42	0.09	9.72E-06	*ABCA13*	Intron	Modifier
18416750	C	T	0.02	-0.61	0.11	2.29E-07	*ABCA13*	Intron	Modifier
18416778	T	C	0.01	-0.69	0.13	2.78E-07	*ABCA13*	Intron	Modifier
18416783	C	T	0.01	-0.69	0.13	2.78E-07	*ABCA13*	Intron	Modifier
18416785	A	T	0.01	-0.69	0.13	2.78E-07	*ABCA13*	Intron	Modifier
18429150	C	T	0.02	-0.48	0.09	1.33E-07	*ABCA13*	Intron	Modifier
18623692	G	A	0.01	-0.68	0.14	2.35E-06	NA	Intergenic	Modifier

Table 4.12: Summary table for SNPs which exceeded the threshold for genome-wide significance for the laminitis status on chromosome 4 in the Morgan horses. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, gene, location of the SNP with the gene (Loc), and the predicted impact based on SnpEff.

Table 4.13: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
18383787	G	A	0.37	-3.79	0.85	1.24E-05	*ABCA13*	Intron	Modifier

Fine-mapped coding SNPs with a p-value <0.05									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Effect**	**Impact**
18331252	G	A	0.24	-1.97	0.91	3.19E-02	*ABCA13*	Synon	Low
18375205	T	G	0.43	-3.00	0.75	8.33E-05	*ABCA13*	Miss	Moderate
18430202	A	G	0.23	2.72	0.99	6.27E-03	*ABCA13*	Synon	Low

Table 4.13: Summary table for SNPs which exceeded the threshold for genome-wide significance, and fine-mapped coding SNPs with a predicted low, moderate, or high impact (genome-wide significance>p-value <0.05) for basal glucose concentrations on chromosome 4 in the Morgan horses. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, and location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff. Additional abbreviations: Synon (synonymous), Miss (missense).

Table 4.14: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
72939355	G	A	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72939516	A	G	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72940303	C	G	0.14	-0.23	0.05	2.40E-05	*NKAIN2*	Intron	Modifier
72940339	G	A	0.14	-0.23	0.05	2.40E-05	*NKAIN2*	Intron	Modifier
72940340	C	G	0.14	-0.23	0.05	2.40E-05	*NKAIN2*	Intron	Modifier
72941231	T	C	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72941505	C	T	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72941924	A	G	0.13	-0.24	0.06	1.72E-05	*NKAIN2*	Intron	Modifier
72942177	A	G	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72942533	A	C	0.13	-0.24	0.06	1.72E-05	*NKAIN2*	Intron	Modifier
72942549	A	T	0.13	-0.24	0.06	1.72E-05	*NKAIN2*	Intron	Modifier
72942931	A	G	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72943273	T	G	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72943314	G	C	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72944804	T	C	0.14	-0.24	0.06	2.57E-05	*NKAIN2*	Intron	Modifier
72945110	G	A	0.13	-0.24	0.06	2.45E-05	*NKAIN2*	Intron	Modifier
72945590	C	T	0.13	-0.24	0.06	2.55E-05	*NKAIN2*	Intron	Modifier
72945989	A	G	0.13	-0.24	0.06	1.71E-05	*NKAIN2*	Intron	Modifier

Table 4.14: Summary table for SNPs which exceeded the threshold for genome-wide significance for the insulin concentrations post oral sugar test on chromosome 10 in the Morgan horses. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, location of the SNP (Loc), and the predicted impact based on SnpEff.

Table 4.15: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
51727537	C	T	0.24	0.33	0.07	1.05E-05	Pseudo	DwnStrm	Modifier
51747148	A	G	0.27	0.32	0.07	1.28E-05	Pseudo	Intron	Modifier
51773433	T	C	0.23	0.35	0.08	1.17E-05	NA	Intergenic	Modifier
51786669	T	C	0.23	0.34	0.08	1.26E-05	NA	Intergenic	Modifier
51794467	C	T	0.24	0.35	0.08	5.80E-06	*BBX*	DwnStrm	Modifier
51794919	C	T	0.24	0.35	0.08	5.80E-06	*BBX*	DwnStrm	Modifier
51795524	C	T	0.24	0.35	0.08	5.65E-06	*BBX*	DwnStrm	Modifier
51808173	C	T	0.24	0.34	0.08	1.26E-05	*BBX*	Intron	Modifier
51812726	G	C	0.26	0.33	0.07	9.94E-06	*BBX*	Intron	Modifier
51812764	T	C	0.24	0.34	0.07	9.81E-06	*BBX*	Intron	Modifier
51813012	T	C	0.24	0.34	0.08	1.23E-05	*BBX*	Intron	Modifier
51844819	C	T	0.22	0.38	0.08	1.32E-06	*BBX*	Intron	Modifier
51855182	C	T	0.23	0.37	0.08	2.12E-06	*BBX*	Intron	Modifier
51857193	G	A	0.24	0.35	0.07	5.33E-06	*BBX*	Intron	Modifier
51863941	G	T	0.22	0.37	0.08	2.93E-06	*BBX*	Intron	Modifier
51867026	C	T	0.22	0.36	0.08	5.00E-06	*BBX*	Intron	Modifier
52084834	T	C	0.23	0.35	0.08	7.74E-06	NA	Intergenic	Modifier
52086447	A	C	0.24	0.36	0.07	1.66E-06	Pseudo	DwnStrm	Modifier
52088943	C	A	0.23	0.37	0.08	2.38E-06	Pseudo	DwnStrm	Modifier
52089422	A	G	0.24	0.35	0.07	5.46E-06	Pseudo	DwnStrm	Modifier
52217527	G	C	0.25	0.33	0.07	1.37E-05	NA	Intergenic	Modifier
52220209	A	G	0.25	0.33	0.08	1.37E-05	NA	Intergenic	Modifier
52224928	C	T	0.25	0.33	0.08	1.37E-05	NA	Intergenic	Modifier
52230428	T	C	0.25	0.33	0.07	1.37E-05	NA	Intergenic	Modifier
52245971	G	A	0.25	0.33	0.07	1.37E-05	NA	Intergenic	Modifier
52256174	T	C	0.31	0.32	0.07	7.03E-06	NA	Intergenic	Modifier
52256653	T	C	0.25	0.34	0.07	1.08E-05	NA	Intergenic	Modifier
52259907	T	C	0.26	0.34	0.07	1.16E-05	NA	Intergenic	Modifier
52263173	G	C	0.26	0.33	0.07	1.13E-05	NA	Intergenic	Modifier
52280838	G	A	0.3	0.35	0.07	5.60E-07	lncRNA	Intragenic	Modifier
52281253	C	T	0.27	0.34	0.07	4.58E-06	lncRNA	Intragenic	Modifier
52282507	T	C	0.27	0.34	0.07	8.58E-06	lncRNA	Intragenic	Modifier
52285378	A	G	0.26	0.33	0.07	1.30E-05	lncRNA	Intragenic	Modifier
52285633	A	G	0.29	0.33	0.07	8.62E-06	lncRNA	Intragenic	Modifier
52286046	G	A	0.31	0.35	0.07	1.01E-06	lncRNA	Intragenic	Modifier

Table 4.15: Summary table for SNPs which exceeded the threshold for genome-wide significance for leptin concentrations on chromosome 19 in the Morgan horses. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, location of the SNP (Loc), and the predicted impact based on SnpEff. Abbreviations: Pseudo (pseudogene).

Table 4.16: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
33700727	C	T	0.46	-0.02	0.004	6.66E-06	*LRRC15*	Intron	Modifier
33784070	C	T	0.47	0.02	0.004	4.92E-06	*ATP13A3*	Intron	Modifier
33844975	A	C	0.44	-0.02	0.003	4.78E-06	NA	Inter	Modifier
34340681	A	G	0.44	-0.02	0.004	5.80E-06	*XXYLT1*	Intron	Modifier
34341048	A	T	0.44	-0.02	0.004	4.27E-06	*XXYLT1*	Intron	Modifier
34341306	T	C	0.44	-0.02	0.004	5.71E-06	*XXYLT1*	Intron	Modifier
34341456	A	G	0.44	-0.02	0.004	2.02E-06	*XXYLT1*	Intron	Modifier
34341665	T	C	0.44	-0.02	0.004	7.60E-06	*XXYLT1*	Intron	Modifier
34342290	C	T	0.44	-0.02	0.004	3.72E-06	*XXYLT1*	Intron	Modifier
34342611	T	C	0.44	-0.02	0.004	2.23E-06	*XXYLT1*	Intron	Modifier
34343879	A	G	0.43	-0.02	0.004	4.94E-06	*XXYLT1*	Intron	Modifier
34344609	C	T	0.43	-0.02	0.004	3.33E-06	*XXYLT1*	Intron	Modifier
34345207	T	C	0.44	-0.02	0.004	6.57E-06	*XXYLT1*	Intron	Modifier
34394019	G	A	0.44	-0.02	0.004	7.22E-06	*XXYLT1*	Intron	Modifier
35859476	A	C	0.28	-0.02	0.004	8.33E-06	lncRNA	UpStrm	Modifier
35859978	G	A	0.27	-0.02	0.004	6.22E-06	lncRNA	UpStrm	Modifier

Fine-mapped coding SNPs a p-value <0.05									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Effect**	**Impact**
33698983	A	T	0.43	-0.01	0.004	3.76E-05	*LRRC15*	Miss	Moderate
33698984	G	A	0.43	-0.02	0.004	2.19E-05	*LRRC15*	Miss	Moderate
33699121	C	A	0.44	-0.01	0.004	5.72E-05	*LRRC15*	Miss	Moderate
33771781	C	T	0.42	-0.01	0.004	8.39E-05	*ATP13A3*	Miss	Moderate
33777160	C	T	0.5	-0.01	0.004	2.59E-04	*ATP13A3*	SS	Low
33829571	T	C	0.29	-0.01	0.004	1.71E-02	*id701397*	SS	Low
33913963	C	T	0.47	0.01	0.004	1.84E-02	*TMEM44*	Miss	Moderate
34449322	T	C	0.1	0.02	0.01	3.06E-03	*XXYLT1*	Miss	Moderate
34503919	C	A	0.08	0.01	0.01	3.17E-02	*ACAP2*	SS	Low
34820819	T	C	0.13	0.01	0.005	3.59E-02	*BDH1*	Synon	Low
35007250	C	T	0.03	0.02	0.01	2.37E-02	*DLG1*	Synon	Low
35675562	T	C	0.25	-0.01	0.004	1.51E-04	*WDR53*	Synon	Low

Table 4.16: Summary table for SNPs which exceeded the threshold for genome-wide significance, and fine-mapped coding SNPs with a predicted low, moderate, or high impact (genome-wide significance>p-value <0.05) for NH ratios on chromosome 19 in the Morgan horses. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, and location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff. Abbreviations: Synon (synonymous), Miss (missense), SS (splice site).

Table 4.17: SNPs which exceed the threshold for genome-wide significance									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Loc**	**Impact**
68755362	G	A	0.03	-0.78	0.16	1.47E-06	NA	Intergenic	Modifier

Fine-mapped coding SNPs with a p-value <0.05									
Pos	**REF**	**ALT**	**MAF**	**Beta**	**SE**	**P-Value**	**Gene**	**Effect**	**Impact**
68349618	G	A	0.16	0.15	0.07	3.80E-02	*NCKAP5L*	Synon	Low
68506154	T	C	0.01	-0.51	0.24	3.17E-02	*AQP6*	5'UTRSG	Low
68582345	C	A	0.02	-0.54	0.22	1.74E-02	*ASIC1*	5'UTRSG	Low
68608591	T	C	0.16	0.15	0.07	3.51E-02	*GPD1*	Miss	Moderate
68666697	T	G	0.01	-0.51	0.24	3.17E-02	*LIMA1*	Miss	Moderate
69123184	C	T	0.2	-0.14	0.07	4.79E-02	*DIP2B*	Synon	Low
69269736	C	G	0.01	-0.51	0.24	3.17E-02	*METTL7A*	Miss	Moderate
69321213	T	C	0.26	0.15	0.06	1.45E-02	*SLC11A2*	Miss	Moderate
69324290	G	A	0.22	-0.14	0.06	2.97E-02	*SLC11A2*	Synon	Low
69351495	G	C	0.29	0.16	0.06	8.77E-03	*LETMD1*	5'UTRSG	Low
69390944	C	T	0.01	-0.51	0.24	3.17E-02	*TFCP2*	Synon	Low
69420938	G	A	0.01	-0.51	0.24	3.17E-02	*TFCP2*	Miss	Moderate
69444682	C	T	0.01	-0.51	0.24	3.17E-02	*POU6F1*	Synon	Low
69512915	C	T	0.02	-0.44	0.21	3.86E-02	*BIN2*	Miss	Low
69649233	C	T	0.16	0.15	0.07	4.95E-02	*SLC4A8*	Synon	Low
69649251	A	G	0.04	-0.39	0.14	5.80E-03	*SLC4A8*	Synon	Low
69667716	T	G	0.02	-0.5	0.16	2.73E-03	*SLC4A8*	Synon	Low
69671998	G	A	0.16	0.17	0.07	2.31E-02	*SLC4A8*	SSDonor	High
69867709	C	T	0.48	0.14	0.06	1.25E-02	*SCN8A*	Synon	Low
69899012	C	G	0.16	0.15	0.07	4.57E-02	*SCN8A*	Synon	Low

Table 4.17: Summary table for SNPs which exceeded the threshold for genome-wide significance, and fine-mapped coding SNPs with a predicted low, moderate, or high impact (genome-wide significance>p-value <0.05) for the adiponectin concentration on chromosome 6 in the Morgan horses. The table includes base pair position (Pos) of the SNP, reference allele (REF), alternative allele (ALT), minor allele frequency (MAF), predicted SNP effect (beta), standard error (SE), P-value, location of the SNP (Loc) or predicted annotated effect (Effect), and the predicted impact based on SnpEff. Abbreviations: Synon (synonymous), Miss (missense), SS (splice site), SSDonor (splice site donor variant), 5'UTRSG (gain of a 5' UTR premature start codon).

Chapter 5: Evaluation of an *HMGA2* variant for pleiotropic effects on height and metabolic traits in ponies

Summary: Ponies are highly susceptible to metabolic derangements including hyperinsulinemia, insulin resistance and adiposity. We hypothesized that genetic loci affecting height in ponies have pleiotropic effects on metabolic pathways and increase the susceptibility to equine metabolic syndrome (EMS). Correlations between height and metabolic traits were assessed by Pearson's correlation coefficients and identified an inverse relationship between height and baseline insulin (-0.26) in ponies in a cohort of 294 Welsh ponies phenotyped for EMS. Using SNP genotype data from 264 Welsh ponies, genomic signature of selection and association analyses for both height and insulin identified the same ~1.3 megabase region on chromosome 6 that contained a shared ancestral haplotype between these traits. The ROI contributed ~40% of the heritability for height and ~20% of the heritability for insulin. *HMGA2* was identified as a candidate gene, and Sanger sequencing detected a c.83G>A (p.G28E) variant associated with height in Shetland ponies. In our cohort of ponies, the A allele had a frequency of 0.76, was strongly correlated with height (-0.75) and was low to moderately correlated with metabolic traits including: insulin (0.32), insulin after an oral sugar test (0.25), non-esterified fatty acids (0.19) and triglyceride (0.22) concentrations. This is the first report of a gene with a pleotropic effect for EMS and provided evidence for the underlying cause of the unique metabolic profiles and increased EMS susceptibility in ponies.

Note: This chapter was previously published in the *Journal of Veterinary Internal Medicine* as follows: Norton, E.M.; Avila, F, Schultz, N.E.; Mickelson, J.R., Geor, R.J.; McCue, M.E. Evaluation of an HMGA2 variant for pleiotropic effects on height and metabolic traits in ponies. *J Vet Intern Med* 2019; 33(2): 942-952.

Introduction

Equine metabolic syndrome (EMS) describes a clustering of metabolic disturbances including insulin dysregulation (hyperinsulinemia and/or insulin resistance) and dyslipidemia (elevated triglyceride low-density lipoprotein concentrations), and generalized obesity and/or regional adiposity (e.g. nuchal ligament, tail head).[3,4] EMS is an important health concern as affected horses and ponies are predisposed to laminitis.

Ponies (individuals with a wither height less than 58 inches) are more insulin insensitive then large breed horses,[36] and metabolic comparisons across breeds have consistently found ponies to be amongst the more insulin resistant groups.[38,51] Unlike many large breed horses, after domestication ponies have maintained a metabolically thrifty phenotype with seasonally-adaptive changes including suppressed metabolic rates and excessive fat storage.[56] However, the mechanisms underlying ponies' unique metabolic profiles and greater EMS susceptibility have not been identified.

A relationship between individuals of short stature and an increased risk of chronic disease has been well described in humans.[548-550] In particular, there are significant associations between height and the risk of developing Type 2 diabetes or metabolic syndrome (MetS),[551-556] with measured metabolic abnormalities more severe in shorter individuals.[551,552,557,558] Many negative correlations between height and specific derangements of the endocrine system include: obesity,[557,559,560] regional adiposity,[555] elevated triglycerides,[552,561] impaired glucose tolerance post oral sugar test,[558,562] and insulin resistance.[551-553,557] Several underlying mechanisms for these associations have been proposed, including a poor uterine environment, impaired nutrition, adverse social circumstances, and genetic factors.[551,563-566] The role of genetic factors is supported by the identification of pleiotropic effect between variants within the promoter of the *GAD2* gene and low birth weight, decreased length, impaired insulin secretion, and early onset obesity,[567] as well as associations between single nucleotide polymorphisms (SNPs) in the *LMNA* gene with short stature and elevated triglycerides, and obesity and increased waist circumference.[568]

We hypothesize that loci affecting height could also have pleiotropic effects on metabolic pathways in horses and ponies and increase the risk for EMS. Here we use genomic tools to identify a chromosomal locus associated with both height and fasting insulin concentrations in Welsh Ponies and demonstrate that a probable functional mutation in the high mobility group AT-hook 2 (*HMGA2)* gene is contributing to both height and metabolic traits.

Material and Methods

Samples: 294 Welsh ponies (213 females and 81 males) from 32 farms within the United States were included in the study, with ages ranging from 2 to 33 years (mean age of 11.7 years). As a breed, Welsh ponies are divided into six sections based on pedigree and height (**Appendix D: Supplemental Table D1**), which were represented in our cohort as follows: section A (n=74), section B (n=146), section C (n=3), section D (n=15), section H or P (n=26), and unregistered Welsh ponies (n=10). 529 individuals from four large-breed horses: Quarter horses (n=59), Arabians (n=64), Tennessee Walking horses (n=48), and Morgan horses (n=293); as well as 65 horses of other pure or mixed breeds, were also collected. These samples were obtained from farms throughout North America and represented 300 females and 229 males with an age range of 2 to 33 years old (mean age of 13 years).

Phenotype Data: Signalment, medical history, height at the withers, and biochemical measurements at baseline and after an oral sugar test (OST), were collected on all individuals. Baseline measurements and assays included: glucose (YSI 2300 STAT Plus glucose and lactate analyzer), insulin (Siemen's TKIN1 Insulin Coat-A-Count Kit), ACTH (Siemen's LKAC1 ACTH kits), leptin (Millipore Sigma's XL-85K Multi-Species Leptin RIA), adiponectin (Millipore Sigma's EZHMWA-64K Human High Molecular Weight Adiponectin ELISA), triglycerides (Millipore Sigma's TR0100 Serum Triglyceride Determination kit), and non-esterified fatty acids (NEFA; Wako Diagnostics' HR Series NEFA kit). OST measurements comprised insulin (INS-OST) and glucose (GLU-OST) levels 75 minutes after oral administration of 0.15mg/kg Karo lite corn syrup.

Genotype Data: Genomic DNA was isolated from whole blood or hair roots per manufacturer recommendations (Puregene Blood Core Kit, Qiagen). Welsh ponies were genotyped with either

the Axiom Equine MCEc670 (n=220 Welsh ponies) or MCEc2M (n=44 Welsh ponies) genotyping arrays, containing 670,805 SNP markers and 2,011,826 SNP markers [409], respectively. For the Welsh ponies not genotyped on the MCEc2M array, Beagle software [407,569] was used to perform genotype imputation and haplotype phasing, using an across-breed reference population of 516 horses of 14 different breeds, yielding a total of 1,931,311 SNPs.

Quality control (QC) measures were performed on the genotyping data using the PLINK software package.[410] This included SNP and individual missingness and genotyping rates, discordant sex information and abnormally high heterozygosity (≥3 standard deviations from the mean). All individuals passed QC and were kept in the study cohort. Individual SNPs with a genotyping success rate <90%, minor allele frequency <1.0%, or outside Hardy Weinberg equilibrium were pruned, leaving a total of 1,511,302 SNPs for subsequent analysis.

FST Based Statistic: Genomic regions of breed-specific population differentiation were identified in the Welsh ponies using SNP data from the 44 individuals genotyped on the MCEc2M. Calculation of the d_i statistic was performed using non-overlapping 10 kilobase (kb) windows across the 31 equine autosomes with a custom Python script (https://github.com/schae234/PonyTools) based on work previously described.[439,570] The d_i statistic detects locus-specific deviation in allele frequencies for the test population relative to the genome-wide average of pairwise F_{ST} summed across populations. The background population contained 463 individuals from 16 different breeds (**Appendix D: Supplemental Table D2**). Significant d_i windows were those corresponding to the top 0.1% of the empirical distribution and were considered regions of interest (ROI) for putative signatures of selection. Two or more contiguous significant d_i windows were considered as a single ROI.

Association Analysis: Association analysis for equine chromosome 6 (ECA6; total of 56,246 SNPs) was performed using imputed SNP genotype data from 264 Welsh ponies. Height and EMS traits were treated as quantitative phenotypes. Association analysis was performed using custom code for a mixed linear regression model that included a random polygenic term determined from a genomic relationship matrix calculated from select trait associated SNPs, random herd effect, and fixed covariates sex and age.[20] Analysis utilized a combination of the Bayesian Sparse Linear Mixed Model (BSLMM),[430] available in the software program Genome-wide Efficient

Mixed Model Association (GEMMA), [379] and a linear mixed model implemented in FaST-LMM[431] (additional description provided in **Appendix D: Supplemental Methods**).

The threshold for genome wide significance was based on the effective number of independent tests for the entire genome (*i.e.*, SNPs, after correction for linkage disequilibrium [LD]), as calculated using the Genetic Type 1 Error Calculator.[434] The effective number of independent tests was 841,750 resulting in a Bonferroni-corrected threshold for genome wide significance of 5.9e-08.

Estimation of Heritability: SNP chip heritability (h^2_{SNP}) for height in Welsh ponies was calculated from the imputed SNP genotype data (n=264) with the software program Linkage Disequilibrium Adjusted Kinship (LDAK),[411,413] including age, sex and section as covariates. Two separate techniques were used to estimate the genetic variance explained by our ROI. First, we used genomic partitioning as previously described.[413,571] The second technique fit the top SNPs from the association analysis as covariates in the analysis using LDAK's --top-preds function. Random subsetting of the data was performed in order to test the effect of a few cryptically related individuals on the h^2_{SNP} estimates (**Appendix D: Supplemental Methods)**.

Haplotype Analysis: Local haplotype sharing within the Welsh ponies used for association analysis (n=264) was calculated from the hapQTL program (http://www.haplotype.org) with default settings.[572] This approach relies on a statistical model for LD to infer ancestral haplotypes and their frequencies at each SNP marker for individuals within a population. For each analysis one expectation maximization run was used with 50 steps (-w 50), 3 upper clusters (-C 3), 10 lower clusters (-c 10), and with a prior LD length of .5 centiMorgan (-mg 200). Based on recommendations from Xu and Guan (2014), contiguous SNPs with $-\log_{10}$ Bayes factor (BF) >4 were considered significant ROI, and orphan signals were removed from the analysis. BF values were calculated for each of the 56,740 SNPs on ECA6 using height and baseline insulin as quantitative phenotypes.

HMGA2 and IRAK3 Reconstruction and Sequencing: PCR primers were designed for all exons within two candidate genes, *HMGA2* and interleukin 2 receptor associated kinase 3 (*IRAK3)*, using the Primer3 software.[573] Genomic sequences for primer design were retrieved using the

National Center for Biotechnology Information (NCBI) Gene tool (https://www.ncbi.nlm.nih.gov/gene); base pair (bp) position of equine exons were confirmed with NCBI's Nucleotide BLAST tool (https://blast.ncbi.nlm.nih.gov/) against the human genome. In some cases, the newly assembled EquCab3 version of the equine genome was queried using a local BLAST tool to confirm exon sequence identity. Details of all *HMGA2* and *IRAK3* exons, as well as the PCR primer sequences, are presented in **Appendix D: Supplemental Table D3 and D4.**

Genomic DNA from a panel of 56 individuals from 6 different breeds (6 Morgan horses, 6 Arabian horses, 6 Tennessee Walking horses, 12 Quarter horses, 3 Miniature horses, and 18 Welsh ponies) was amplified by standard PCR. The resulting products were submitted to the University of Minnesota Genomics Center for Sanger sequencing after enzymatic cleanup using the ExoSAP-IT™ PCR Product Cleanup Reagent (Thermo Fisher Scientific, Waltham, MA). Sequencing results were then analyzed, processed and aligned using the Sequencher software version 5.1 (Gene Codes Corporation, Ann Arbor, MI).

HMGA2 Exon 1 Variant Genotyping: Two methods were employed to genotype the *HMGA2* exon 1 mutation (*HMGA2* (c.83G>A) identified by Frischknecht, *et al*.[574] In the first method, standard PCR primers were designed to flank and Sanger sequence this exon (**Appendix D: Supplemental Table D3**) in 438 horses, including 150 ponies and 288 large breed horses. In the second method, a TaqMan SNP genotyping assay using the Bio-Rad CFX96 Real-Time System was designed as previously described [575] and per manufacturer's recommendations. Results were analyzed with BioRad's CFX Manager Software version 3.1 (see **Appendix D: Supplemental Methods** for a full description of this assay). Genotypes for this variant using the second genotyping assay were obtained for an additional 144 Welsh ponies and 241 large breed horses.

Statistical Analyses: Statistics were performed using functions within the software package R.[412] Metabolic traits were tested for normality using a normal probability plot and a Shapiro test; traits were log or square root transformed when appropriate. Correlations between height and EMS traits (insulin, INS-OST, glucose, GLU-OST, NEFA, triglycerides, leptin, adiponectin) and ACTH, were calculated using a Pearson's correlation coefficient. After adjusting for multiple testing using a Bonferroni correction (0.05/9), a p-value of <0.0056 was considered significant.

Analyses were performed as follows: all horses (n=824), Welsh ponies (n=294), all large breed horses (n=529), Quarter horses (n=59), Arabian horses (n=64), Morgan horses (n=293), and Tennessee Walking horses (n=48). Correlations between genotype for the *HMGA2* c83G>A variant and EMS traits, ACTH or height were calculated for the Welsh ponies (n=294) using Pearson's correlation coefficient and a Bonferroni corrected p-value (0.05/10; <0.005).

Least-square means were calculated with EMS traits, ACTH or height as the outcome variable, genotype as the response variable, and age and sex as predictors. The R statistical software package Linear and Nonlinear Mixed Effects Models (nlme) [576] fit the linear model using generalized least squares. The R statistical software package Least-Square Means (lsmeans) [577] was used to calculate the predicted marginal means and pair-wise comparisons.

Model comparison for modes of inheritance between the *HMGA2* c.83G>A variant and traits were performed using an ANOVA for an additive, dominant and recessive model. The p-values of the f-statistic were compared across all three models. The R statistical software package SNPassoc[578] was used to calculate the Akaike information criterion (AIC) and p-value between additive, recessive, dominant, and co-dominant models. Model selection was based on the lowest AIC values; however, models with less than 10 units difference between them were considered indistinguishable.

Results:

Correlations Between Height, EMS Traits, and ACTH: Correlation analyses between height and biochemical traits in the entire cohort (n=823), revealed statistically significant inverse correlations for insulin (-0.12), glucose (-0.11), , adiponectin (-0.23) and ACTH (-0.12); while positive correlations with height were found for triglycerides (0.14) and leptin (0.12) (**Table 5.1**). No statistically significant correlations between any of the traits and height were identified in the large breed horses as a whole (n=529), or within any individual breed (**Table 5.1**). However, within the Welsh pony population (n=294), a statistically significant inverse correlation with height was identified for insulin (-0.26), with the correlation coefficient between height and insulin higher than in the entire population (**Table 5.1**), indicating that the pony population was predominately driving the association observed for this trait in the full cohort.

FST-Based Statistic to Detect Signatures of Selection: 212,208 non-overlapping, 10kb windows across all 31 equine autosomes were analyzed in the Welsh pony cohort, with an average of 8.2 (+/- 3.2) SNPs per window. A total of 212 windows were within the top 0.1% of the empirical distribution of *di* values, which in turn represented 134 ROI. Among the significant *di* windows, 50 (24%) were located on ECA6 and corresponded to eight separate ROI (**Figure 5.1**). One of these ECA6 ROI comprised 42 (20%) of the total significant *di* windows and spanned an ~782kb segment. Based on EquCab2, the equine reference genome available at the time of this analysis, this segment ranged from bp positions 81,003,617 to 81,785,414 (**Appendix D: Supplemental Figure D1)**. The other seven significant ROIs on ECA6 were derived from singleton *di* windows, located at least one megabase (Mb) apart. 162 other significant *di* windows were distributed throughout all autosomes, except chromosomes 12, 16, 19, 30 and 31.

Association Analysis: For the Welsh pony cohort, p-values for 142 SNPs on ECA6 associated with height exceeded the threshold for genome-wide significance (**Figure 5.2A**). Based on EquCab2, all 142 SNPs were within the same ~1.3Mb region and included SNPs from bp position 80,501,273 to 81,808,008. For insulin, p-values for 58 SNPs on ECA6 exceeded the threshold for genome-wide significance and included SNPs from bp position 80,639,787 to 81,651,604 (**Figure 5.2B**). Significant SNPs within this ROI were not identified for any of the other EMS traits or ACTH.

Heritability and Genetic Variation: The h^2_{SNP} for height in the Welsh ponies was 0.87 (SD = 0.084). Using genomic partitioning for height, the percent of the genetic variation contributed by the ROI (SNPs from bp position 80,501,273 to 81,808,008) on ECA6 was 0.34 (SD = 0.083); *i.e.*, 39% of the total h^2_{SNP}. The top SNPs from association analysis were included in the h^2_{SNP} model as covariates to estimate the contribution of these SNPs to height in ponies. The 142 SNPs on ECA6 that exceeded the threshold for genome wide significance on association analysis were pruned at an LD of >.8 to avoid over fitting the h^2_{SNP} model, leaving 42 SNPs for analysis. The percent of genetic variation contributed by these 42 SNPs was estimated to be 0.41, *i.e.***,** 47% of the total h^2_{SNP}. After random subsetting of the data, the resultant mean values for h^2_{SNP} were not significantly different from the original estimates above as follows: .89 (SD = 0.087) for the overall h^2_{SNP} estimate of height, 0.38 (SD = 0.087) for genomic partitioning at the ROI, and 0.45 using the top SNPs from association analysis as covariates.

Within this cohort, we previously showed that baseline insulin had a h^2_{SNP} of 0.81 (SD=0.11), with a mean h^2_{SNP} of 0.82 (mean SE: 0.12) after random subsetting.[579] In this analysis, the h^2_{SNP} explained by genomic partitioning was 0.19 (SD=0.086), or 24% of the total h^2_{SNP} for baseline insulin. Of the 58 significant SNPs found on association analysis, 13 remained in our analysis after pruning for LD. Including these SNPs as top-predictors, the percent of genetic variation contributed by these SNPs was 0.13, or 16% of the total h^2_{SNP}. After random subsetting the data, the mean h^2_{SNP} for genomic partitioning at the ROI was 0.20 (SD = 0.086) and 0.14 using the top SNPs approach.

Haplotype Analyses for Height and Baseline Insulin: Nearly 40% (23,058) of all (56,740) ECA6 SNPs had a BF value >4 when analyzing height as the trait of interest. 107 SNPs had the highest BF values (>30) and were within the range of bp positions 81,012,766 to 81,782,298 (**Figure 5.3A**). Evaluation of all 652 SNPs within and flanking 1kb of the ROIs identified by association analysis and d_i statistic (SNPs from 80,499,826 to 81,809,066 bp), showed that all SNPs exceeded the BF value threshold, with values ranging from 4.17 to 40.12 (**Figure 5.3A**). When analyzing haplotypes using baseline insulin as the trait of interest, 290 SNPs on ECA6 had a BF value >4, which included 171 of the 652 SNPs comprising the ROI. The haplotypes consisted of two predominant regions where 46 SNPs were within bp positions 81,161,980 to 81,288,528 and 71 SNPs were within bp positions 81,381,221 to 81,583,507 (**Figure 5.3B**). The latter region also contained the SNPs with the highest BF values for the entire analysis (maximum BF of 7.5). HapQTL did not identify haplotypes on ECA6 for any of the other traits.

Candidate Gene Identification, Sequencing and Genotyping: The ROI identified in this study from association analysis and d_i statistics (ECA6: 80,499,826-81,809,066) was further analyzed for positional candidate genes. Using NCBI and the Ensembl genome browser with EquCab2 as the reference genome, a total of 16 positional candidate genes were identified, comprising three RNA genes, two pseudogenes, and 11 protein coding genes (**Figure 5.3C**). A search of the PubMed literature database for known biological function and relevance in other species resulted in the prioritization of *HMGA2* and *IRAK3* as biological positional candidate genes. *HMGA2* was the only protein-coding gene within the smaller 81,161,980 - 81,583,507 region fine mapped by haplotype analysis.

The *HMGA2* c.83G>A variant in exon 1 reported by Frischnecht *et al*[574] was identified in our 56 horse multi-breed cohort (6 Morgan horses, 6 Arabian horses, 6 Tennessee Walking horses, 12 Quarter horses, 3 Miniature horses, and 18 Welsh ponies); however, no additional *HMGA2* or *IRAK3* exonic variants were detected. All individuals (n=823) were then genotyped for the *HMGA2* c83G>A variant. In the Welsh pony (n=294) cohort, the A allele frequency was 0.76 and the G allele frequency was 0.24 **(Table 5.2)**. The *HMGA2* A allele frequencies across the sex sections of the Welsh pony present in our population were 1.0 for section A, 0.74 for section B, 0.83 for section C, 0.03 for section D, and 0.64 for section H/P (**Table 5.2, Appendix D: Supplemental Table D1**). In the large breed horses (n=529), there were only five horses heterozygous for the *HMGA2* A allele (2 Tennessee Walking horses, 1 Morgan horse, 1 Mustang, and 1 Kentucky Mountain horse); resulting in an overall A allele frequency of 0.005.

Correlations Between HMGA2 Genotype, EMS Traits, and ACTH: Correlation analyses between *HMGA2* genotype and the measured traits were performed in Welsh ponies. A negative (-0.75; 95%CI: -0.80 to -0.70; p-value <0.001) correlation was identified between the A allele and height. Pairwise comparisons of the least square means of height and *HMGA2* genotype revealed statistically significant differences between all three genotypes (**Figure 5.4A**). Although the ANOVA f-statistic did not differentiate between the three possible modes of inheritance, an additive model was favored over recessive and dominant based on AIC (**Appendix D: Supplemental Table D5**).

Positive correlations with p-values <0.005 were also identified between the *HMGA2* A allele and four of the nine measured EMS traits in the ponies: including insulin (0.32; 95%CI: 0.21 to 0.42), INS-OST (0.25; 95%CI: 0.14 to 0.35), NEFA (0.19; 95%CI: 0.075 to 0.30), and triglycerides (0.22; 95% CI: 0.10 to 0.32). Correlations for traits that were not statistically significant included: glucose, GLU-OST, leptin, adiponectin, and ACTH. Pairwise comparisons for insulin, INS-OST, and triglycerides revealed that the predicted marginal means for the A/A genotype were statistically different (p-value < 0.001) from both the G/G and G/A genotypes, but that the predicted marginal means for the G/G and G/A genotypes were not statistically different from each other, suggesting recessives model of inheritance for these measurements (**Figures 5.4B, 5.4C and 5.4D**). Although the p-values for the F-statistic linear regression modeling slightly favored recessive models for insulin, INS-OST and NEFA, the AIC values showed minimal separation

between additive and recessive models for all four biochemical measurements (**Appendix D: Supplemental Table D5**). Pairwise comparisons between the marginal means and genotype for NEFA also revealed statistically significant differences between the A/A and G/A genotypes (**Figure 5.4E**).

Discussion:

It is well recognized that ponies are at high risk for developing EMS; however, the mechanisms underlying this increased susceptibility, and the roles that genetic factors might play, are not understood. In this study we demonstrated that baseline insulin values, a major component of the EMS phenotype, were correlated to height in Welsh ponies. Using complementary genome-wide analysis methods with high-density SNP genotype data, we identified and fine-mapped a locus on ECA6 associated with both of these traits in Welsh ponies, which we estimated to be contributing ~40% and ~20% of the total h^2_{SNP} for height and insulin, respectively. The positional candidate genes *HMGA2* and *IRAK3* were prioritized based on known biological function and evidence in other species. Sequencing of the promoters, coding exons and flanking intronic regions revealed only a c.83G>A variant (p.G28E) in *HMGA2*, previously described in other small stature horse breeds.[574] Correlations between *HMGA2* genotype and critical metabolic measures of EMS in the Welsh ponies suggested a previously unrecognized pleiotropic effect of this locus and its candidate *HMGA2* functional variant.

Similar to what has been found in humans, an inverse correlation between height and five EMS measurements (insulin, glucose, triglycerides, leptin, and adiponectin) as well as ACTH were found in the large cohort of horses and Welsh ponies. However, we determined that the ponies were predominately driving the correlations in this cohort for baseline insulin, as statistically significant correlations were not identified for any of the four other individual breeds. This led us to investigate whether genetic loci for height, EMS measures, and ACTH in Welsh ponies could be one and the same.

High-density SNP genotype data enabled us to use an F_{ST}-based approach (*di*) to detect regions of low heterogeneity that exist due to selection for a phenotype, as well as identify genomic regions containing variants associated with both height and insulin on ECA6. We identified several breed-

specific loci undergoing selection in the Welsh pony; however, the region with the highest number of significant *di* windows, as well as those at the top of the empirical distribution, was a ~782kb segment on ECA6 that was within the boundaries of the 1.3Mb ROI identified by association analysis. Although the *di* statistic is blinded to phenotype, given the extensive breeding selection for short stature in ponies and the overlapping results with the association analysis, we surmised that selection for height was responsible for this genomic signature. Based on our cohort and the high heritability of height and baseline insulin, our association analysis had adequate power to identify alleles with moderate to high effect size [580] and readily detected the ECA6 locus in Welsh ponies for both traits.

With genomic partitioning, we estimated that the ROI (ECA6: 80,499,826-81,809,066) contributed to 39% of the genetic variation for height, and 24% for baseline insulin. However, this approach leads to inclusion of SNPs that were top predictors from association analysis, violating the effect size assumption when using a restricted estimated maximum likelihood analysis. Thus, we also performed a top predictors approach after pruning for highly correlated SNPs that resulted in an estimate of genomic contribution of 47% for height and 16% for baseline insulin. Although these estimates were not performed in an independent population, and can lead to over fitting of the data, it does suggest that the ECA6 locus is contributing ~40% of the genetic variation of height and ~20% for baseline insulin in our population. Unaccounted for population stratification or cryptic relatedness can lead to overestimation of h^2_{SNP}. However, the mean h^2_{SNP} estimates and standard deviations after randomly subsetting the data did not significantly differ from the original estimates, indicating that population substructure or cryptic relatedness was not significantly biasing our estimates (**Appendix D: Supplemental Methods**).

We identified a haplotype block that spanned the entire height ROI on ECA6 found by association analysis, while haplotype blocks in the same region for baseline insulin contained distinct major and minor peaks. This likely reflects differences in variant effect size, non-shared factors affecting the traits, and selection for height. We showed that 39-47% of the genetic variance in height could be explained by our ROI on ECA6; thus, the locus has a large effect on height in ponies. In contrast, the effect on insulin is smaller at 21-25% of the genetic variation. This is consistent with the results from the Pearson's correlation between height and insulin which was -0.26, indicating that not all the variation in insulin could be explained by its relationship to height with non-shared

factors present between the traits. Finally, short stature has been strongly selected for in ponies through extensive breeding; however, hyperinsulinemia is not a desirable trait. The long haplotype on height likely reflects extensive hitchhiking secondary to selective breeding for that trait. Thus, haplotype analysis allowed us to fine map our ROI for height and insulin to bp positions: 81,161,980-81,583,507, where *HMGA2* was the only annotated coding gene.

The *HMGA2* protein interacts with AT-rich regions of DNA through three DNA binding domains (AT hooks). This interaction alters the chromatin structure and promotes protein-protein interactions necessary for assembly and stabilization of the enhanceosome during initiation of transcription.[581] *HMGA2*'s main functional role is thought to be in cellular proliferation and differentiation, which has been supported by the numerous studies in humans linking *HMGA2* with height.[582-588] The *HMGA2* locus was also identified as being one of four loci explaining 83% of the genetic variation of height in horses, and one of six loci explaining 46-52.5% of the genetic variation of height in dogs.[395,589] Further, knockout mouse models for *HMGA2* result in a lean, pygmy phenotype;[590] whereas, gain of function mutations of this gene led to gigantism, excessive fat formation and lipomatosis in both mice and humans.[591,592] In addition to the alterations in fat metabolism noted above, *HMGA2* has been associated with other causes of metabolic derangements, particularly type II diabetes in humans.[593] Voight et al. hypothesized that an *HMGA2* variant was likely affecting insulin levels independent of an obesity driven mechanism.[593] Since then, both genome-wide association and meta-analyses have replicated this result.[594-596] The only *HMGA2* variant found in our panel of 48 horses was a missense mutation (c.83G>A) in exon 1, which was previously described as associated with decreased height in Shetland and other pony breeds.[574] The variant, with its glycine to glutamate substitution at residue 28, is predicted to affect the first AT hook, and the authors demonstrated that the mutant nucleotide sequence had decreased binding affinity for DNA. This is further evidence supporting the likely functional impact of this mutation.

In our pony cohort, the *HMGA2* variant had an allele frequency of 0.76, was distributed across the sections of the Welsh Pony breed consistent with their height distribution, was negatively correlated (-0.75) with height, and its effect was explained by an additive model of inheritance in the entire population of ponies. We also identified a negative correlation for the A allele with four

EMS traits, including insulin, NEFA, INS-OST and triglycerides. This provides evidence that *HMGA2* is having an effect on EMS traits beyond modulating height.

Notably, pairwise comparisons of NEFA between genotypes revealed that, although there was a statistical difference between the A/A and G/A genotypes, there was not a difference between either of the homozygous genotypes. This is most likely due to the large 95% confidence intervals identified when assessing the least square means for genotype and NEFA concentrations in the ponies, particularly those with the G/G genotype (**Figure 5.4E**). Pairwise comparisons between the least squared means for genotype and insulin, INS-OST, and triglycerides suggested a recessive model of inheritance; however, model analyses were unable to differentiate between an additive or recessive model. The lack of distinction is likely due to the large variation within EMS traits, as well as bias due to unequal sampling amongst our ponies, as our cohort only included three section Cs and 15 section Ds. Therefore, inclusion of more samples from these sections would likely improve our power to differentiate between an additive and recessive model.

IRAK3 was included as a biological candidate gene due to evidence in other species and its close proximity to the fine mapped ROI. *IRAK3* is down-regulated in individuals with obesity and metabolic syndrome, and is thought to be a key inhibitor of inflammation during metabolic derangements.[597] Further, *IRAK3* mutant mouse phenotypes include reduced body size, decreased femur diameter and abnormal bone morphology [598], as well as impaired glucose tolerance [599]. We sequenced the *IRAK3* gene in our sample panel of horses but did not find any variants. Although a predicted miRNA (MIR763) was within our refined ROI, its function is unknown and does not have any associated orthologues.

In conclusion, through genome-wide analyses we identified an allele for a known height gene, *HMGA2*, as contributing to both height and several EMS traits in a cohort of Welsh ponies. Additional functional analysis would determine if the *HMGA2* mutation has a pleiotropic effect on these traits, or if another unidentified variant within our ROI independently contributes to the EMS traits and has been inadvertently selected for due to genomic hitchhiking. Although this study focused on Welsh ponies, the *HMGA2* variant has been corelated with height in other pony breeds; thus, it is likely that this variant is also having an effect on metabolic traits in these individuals, as supported by the correlation analysis with the addition of three Shetland, two

Hackney, and three British Riding ponies to our cohort (**Appendix D: Supplemental Table D7**). Moreover, although height was not correlated with EMS traits in the large breed horses in this study, this does not rule out stature as contributing to these traits in that population. In humans, leg length-to-torso ratios are consistently correlated with metabolic traits over total height [552]. Therefore, length-to-torso ratios in large breed horses might reveal a correlation not identified in this analysis. These data are a major step forward towards understanding genetic influences on EMS that could also have implications for improving equine health and understanding contributors to MetS.

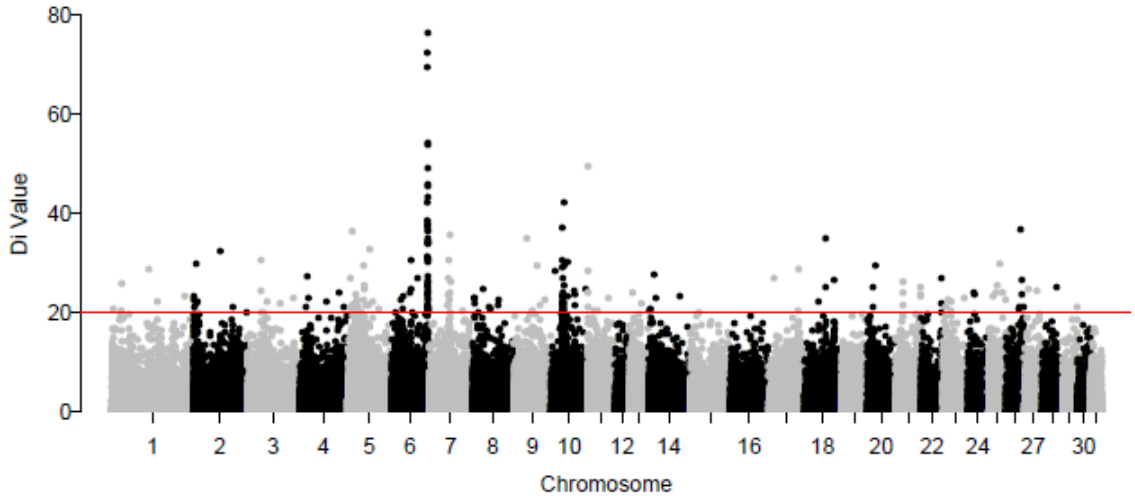


Figure 5.1 Genome-wide *di* values for Welsh Ponies. Each d_i value is plotted on the y-axis and each autosome is shown on the x axis in alternating colors. Each dot represents a 10 kb window. The red horizontal line represents the top 0.1% of the empirical distribution of d_i values. One region of interest on equine chromosome 6 (ECA6) spanned ~782kb segment, ranging from 81,003,617 to 81,785,414 bp.

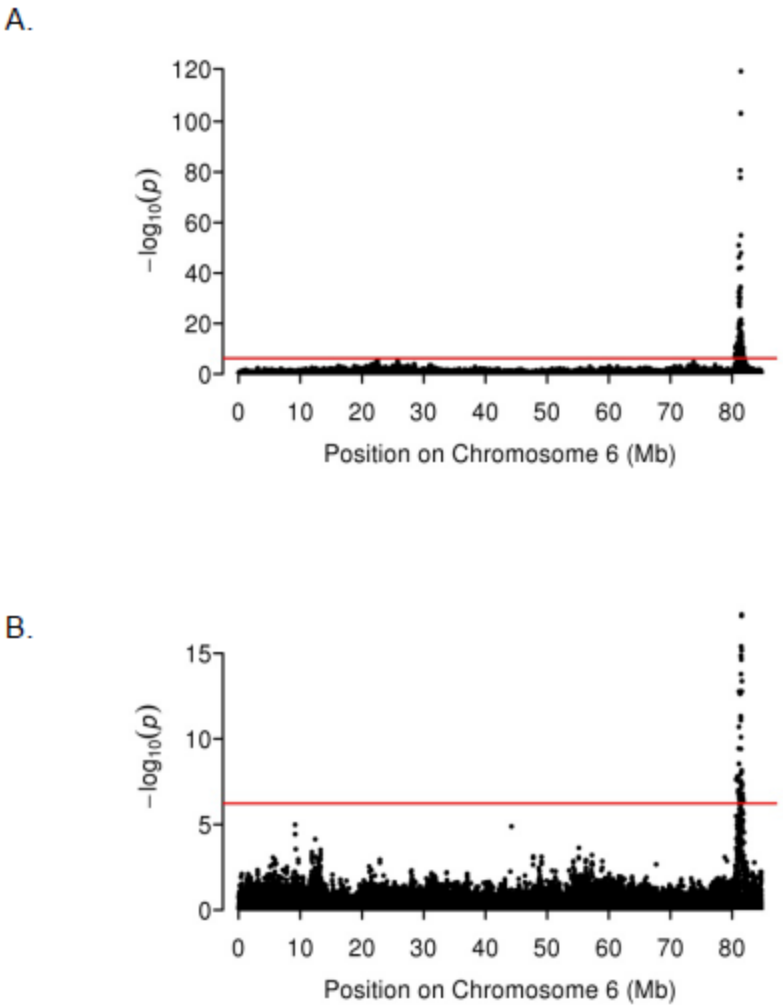


Figure 5.2: Plot of the association analysis for equine chromosome 6 (ECA6) in 264 Welsh ponies (WP). The base pair positions for chromosome 6 are plotted along the x-axis and the $-\log_{10}$ of the p-values are plotted on the y-axis. Individual circles represent single SNPs. A red line marks the thresholds for genome wide significance. (A) Results obtained in WP for height. Significant associations were noted on ECA6 with SNPs between 80,501,273 and 81,808,008 bp. (B) Results obtained in WP for baseline insulin. Significant associations were noted on ECA6 with SNPs between 80,639,787 to 81,651,604 bp.

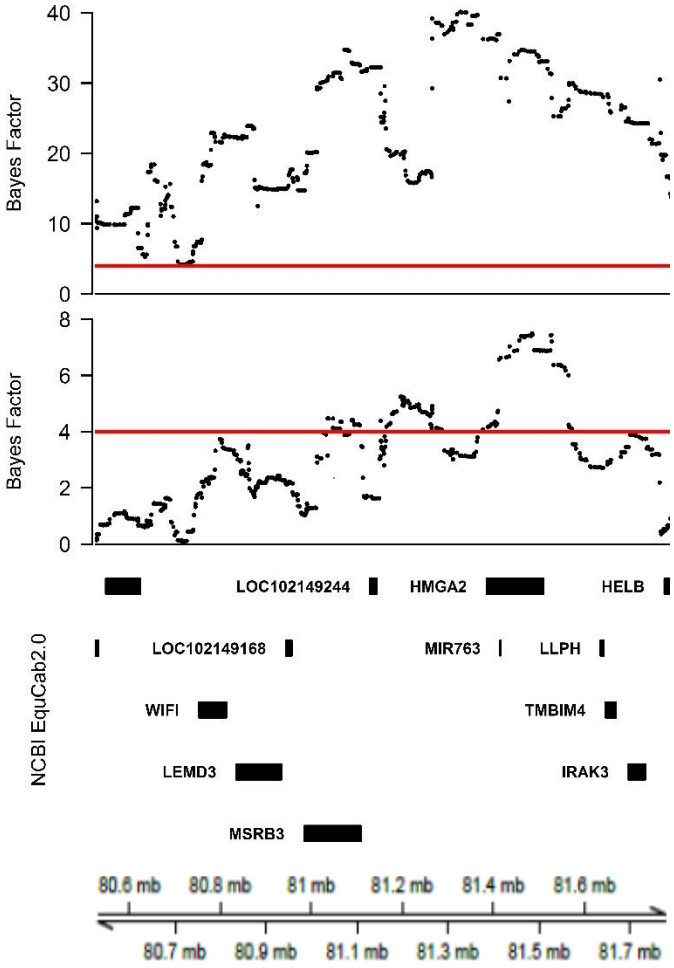


Figure 5.3. Fine-scale structure of the region of interest on equine chromosome 6 (ECA6). Regions of interest (ROI) identified from the results of the association analysis and di statistic were used for haplotype analysis for both height (A) and baseline insulin values (B) in Welsh ponies. Bayes Factor values above the red horizontal line are considered significant and represent an ancestral haplotype. Shared ancestral haplotypes between both traits are most predominant from base pair positions 81,161,980 to 81,288,528 and 81,381,221 to 81,583,507. Aligning the NCBI genome browser for the ROI identified *HMGA* (red circle) as a coding gene within the shared haplotype. *IRAK3* was also identified as a candidate gene based on proximity and biological data.

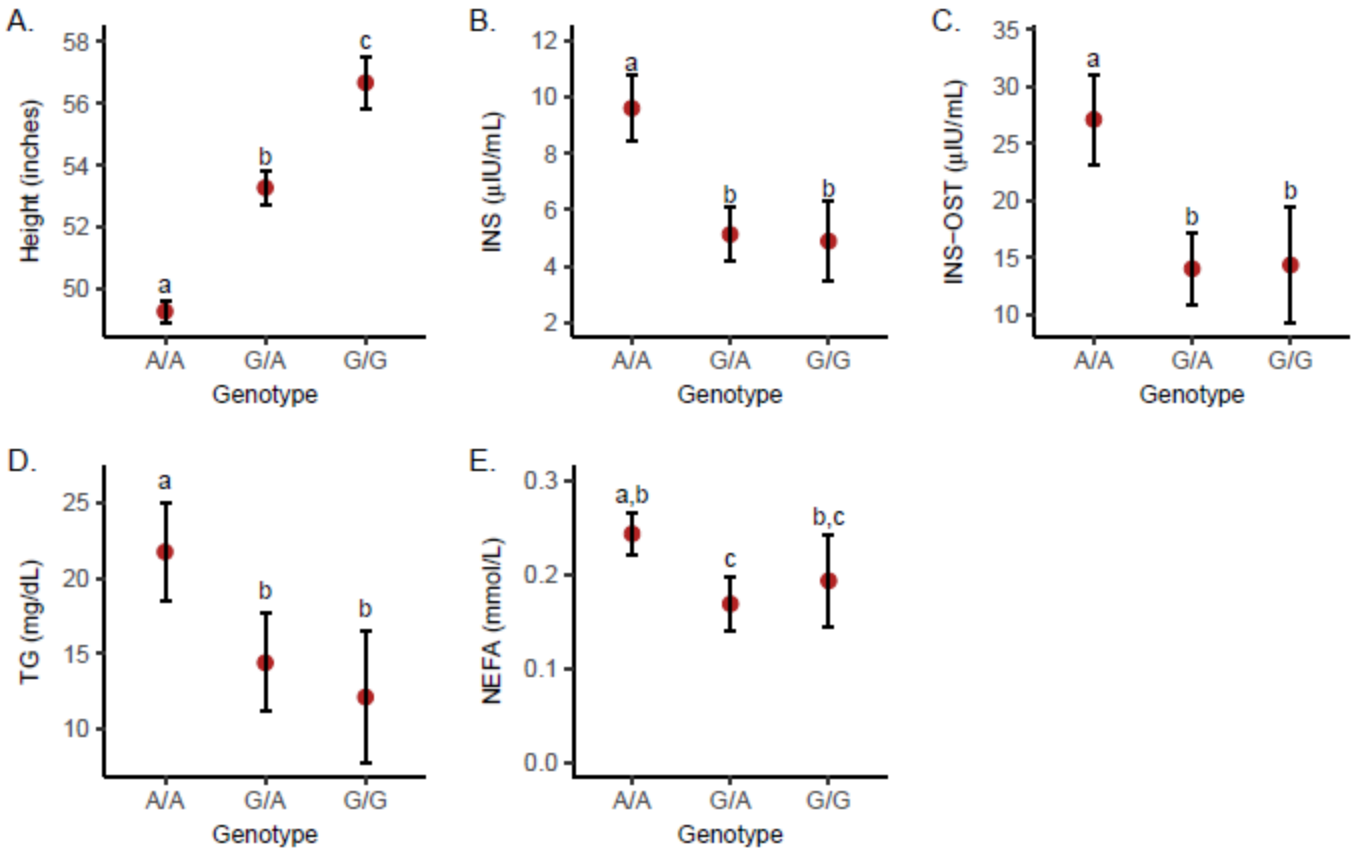


Figure 5.4: Least-square mean estimates and 95% confidence intervals for the *HMGA2* c.83G>A variant and various phenotypes in a population of 294 Welsh ponies. Height (A), insulin (B), INS-OST (C), triglycerides (D), and NEFA (E).

Breed	INS	INS-OST	GLU	GLU-OST	NEFA	TG	LEPTIN	ADIPON	ACTH
All Horses n=823	**-0.12 (-0.19,-0.05) p=<0.001**	-0.035 (-0.11,0.04) p=0.34	**-0.11 (-0.18,-0.04) p=0.002**	-0.07 (-0.14,0.003) p=0.006	-0.07 (-0.14,0.01) p=0.053	**0.14 (0.07,0.21) p=<0.001**	**0.12 (0.05,0.19) p=<0.001**	**-0.23 (-0.29,-0.16) p=<0.001**	**-0.12 (-0.18,-0.04) p=.0015**
Welsh ponies n=294	**-0.26 (-0.36,-0.15) p=<0.001**	-0.10 (-0.21,-0.02) p=0.12	-0.07 (-0.18,0.05) p=0.25	0.08 (-0.03,0.20) p=0.02	-0.13 (-0.24,-0.01) p=0.03	-0.12 (-0.23,-0.04) p=0.04	-0.07 (-0.18,0.04) p=0.23	0.06 (-0.06,0.17) p=0.35	-0.12 (-0.23,0.005) p=0.05
Large Breed n=529	-0.02 (-0.11,0.06) p=0.59	-0.07 (-0.16,0.02) p=0.15	0.02 (-0.07,0.10) p=0.72	-0.007 (-0.10,0.09) p=0.89	0.12 (0.03,0.20) p=0.008	-0.001 (-0.09,0.09) p=0.98	-0.06 (-.014,0.03) p=0.18	-0.05 (-0.13,0.04) p=0.29	0.03 (-0.06,0.11) p=0.55
Morgans n=293	-0.11 (-0.22, 0.0) p=0.05	-0.13 (-0.24,0.01) p=0.03	0.002 (-0.11,0.12) p=0.98	-0.10 (-0.21,0.02) p=0.12	0.07 (-0.05,0.18) p=0.25	-0.06 (-0.18,0.05) p=0.28	-0.06 (-0.17,0.06) p=0.32	-0.003 (-0.12,0.11) p=0.95	-0.05 (-0.16,0.07) p=.43
QH n=59	0.19 (-0.07,0.43) p=0.14	0.24 (-0.10,0.52) p=0.16	0.12 (-0.14,0.36) p=0.38	0.17 (-0.16,0.47) p=0.31	0.25 (-0.01,0.47) p=0.06	0.13 (-0.13,0.37) p=0.33	-0.005 (-0.26,0.25) p=0.97	-0.08 (-0.33,0.18) p=0.55	0.25 (-0.07,0.48) p=0.06
TWH n=48	0.23 (-0.06,0.48) p=0.12	0.08 (-0.28,0.42) p=0.66	0.14 (-0.15,0.41) p=0.35	0.13 (-0.23,0.46) p=0.47	-0.12 (-0.39,0.17) p=0.41	-0.08 (-0.36,0.17) p=0.60	0.04 (-0.26,0.32) p=0.81	-0.17 (-0.43,0.12) p=0.25	-0.08 (-0.35,0.21) p=0.61
Arabians n=64	-0.31 (-0.51,-0.06) p=0.02	-0.25 (-0.05,0.02) p=0.07	-0.19 (-.042,0.06) p=0.14	0.01 (-0.26,0.28) p=0.94	0.12 (-0.13,0.36) p=0.34	-0.21 (-0.44,0.04) p=0.10	-0.12 (-0.36,0.13) p=0.36	0.02 (-0.23,0.27) p=0.89	-0.12 (-0.04,0.13) p=0.34

Table 5.1: Correlations between height and biochemical traits across breeds. Pearson's correlation coefficients, 95% confidence intervals and p-values for height, eight EMS biochemical traits, and ACTH across breeds of horses. All traits were corrected for age and sex prior to analysis. Significant p-values (<.0056) are in bolded text. Abbreviations: INS = insulin, INS-OST = insulin after an oral sugar test, GLU = glucose, GLU-OST, glucose after an oral sugar test, NEFA = non-esterified fatty acids, TG = triglycerides, ADIPON = Adiponectin, QH = Quarter horses, TWH= Tennessee Walking Horses.

Breed	n	G/G (WT)	G/A (HET)	A/A (MUT)	A Allele Frequency	G Allele Frequency
Welsh ponies	294	30	80	184	0.76	0.24
Section A	78			78	1.00	0.00
Section B	150	8	62	80	0.74	0.26
Section C	3		1	2	0.83	0.17
Section D	15	14	1		0.03	0.97
Section H	37	8	11	18	0.64	0.37
Unregistered	11		5	6	0.77	0.23
All large breed horses	530	525	5		0.005	0.995
Morgan horses	293	292	1		0.002	0.998
Quarter horses	59	59				1.00
Tennessee Walking horses	48	46	2		0.021	0.98
Arabians	64	64				1.00
Other large breed horses	66	64	2		0.015	0.985

Table 5.2: Genotyping results for the *HMGA2* c.83G>A variant in Welsh ponies and large breed horses. Results are also shown for specific breeds including: sections of Welsh ponies, Morgan horses, Quarter horses, Tennessee Walking horse, and Arabians. Allele frequencies are provided for the G (wild type) and A (mutant) allele. Abbreviations: WT: Homozygous for the wild type allele, HET=Heterozygote, MUT=Homozygous for the mutant allele.

Chapter 6: Conclusions and Future Directions

The term equine metabolic syndrome (EMS) was coined in 2002 when the parallels between human metabolic syndrome (MetS) and what was being observed clinically in hyperinsulinemic, obese, laminitic horses was recognized [7]. Over the past few decades, the working understanding of EMS has become more refined and the clinical overlap between species more pronounced. In both species, metabolic syndrome can manifest as baseline hyperinsulinemia, an exaggerated or prolonged insulin or glucose response post carbohydrate challenge, tissue insulin resistance, dyslipidemia, and alterations in adipokines and inflammatory cytokines. Further, metabolic syndrome can lead to serious medical issues that have a major economic impact [8,21], with EMS being the leading cause of laminitis [21] and humans with MetS being 2 and 4 times more likely to develop cardiovascular disease and diabetes mellitus, respectively [8].

Breed predilections and familial incidence have provided the initial evidence that EMS is a complex trait with a strong genetic basis, but after nearly two decades the genetic risk factors contributing to EMS have remained undiscovered. The identification of risk alleles and gene pathways underlying EMS will allow for a better understanding of the fundamental pathogenesis of the syndrome. Moreover, the promise of a genetic test that can be used to identify high risk horses before they develop clinical signs and laminitis has instigated a drive toward identifying the genetic risk factors of EMS. Further, the similarities between metabolic syndrome in humans and horses make horses an excellent naturally-occurring model of MetS and an ideal candidate for further exploring the genetic contribution to metabolic syndrome across species.

In order to move toward the identification of the specific genes or alleles contributing to EMS, the objectives of this thesis were to: (i) estimate the genetic contribution to EMS metabolic traits, (ii) identify regions of the genome harboring EMS risk alleles, and (iii) identify the candidate genes and putative functional alleles contributing to EMS.

Chapter Summaries and Conclusions

In chapter 2, we provided the first concrete evidence of a genetic contribution to EMS, quantifying the genetic contribution to nine traits that comprise the EMS phenotype. We

used high-density SNP genotype data to estimate the heritability (h^2_{SNP}) of nine biochemical traits in a cohort of 264 Welsh ponies and 286 Morgan horses with a restricted estimated maximum likelihood statistic. In Welsh ponies, seven of the nine traits had statistically significant h^2_{SNP} estimates that were considered moderately to highly heritable (h^2_{SNP} > 0.20) including: triglycerides (0.313; SE=0.146), glucose (0.408; SE=0.135), NEFA (0.434; SE=0.136), INS-OST (0.440; SE=0.148), adiponectin (0.488; SE=0.143), leptin (0.554; SE=0.132) and insulin (0.808; SE=0.108). In Morgan horses, six of the nine traits had statistically significant h^2_{SNP} estimates that were also determined to be moderately to highly heritable including: INS-OST (0.359; SE=0.185), leptin (0.486; SE=0.177), GLU-OST (0.566 SE=0.175), insulin (0.592; SE=0.195), NEFA (0.684; SE=0.164), and adiponectin (0.913; SE=0.181). These results are the first to indicate that EMS biochemical traits are moderately to highly heritable. We hypothesize that differences in h^2_{SNP} estimates in several traits between these two breeds is likely due to differences in risk alleles or the frequency of risk alleles that are contributing to previously identified breed variability in metabolic traits.

In chapter 3, we provided strong evidence that EMS is a complex, polygenic syndrome with dozens of risk alleles contributing to the phenotype. Using high-density SNP genotype data, genome-association analyses (GWA) was performed for twelve EMS relevant traits using a custom code for a linear mixed model in a cohort of 264 Welsh ponies and 286 Morgan horses. Regions were defined as consecutive SNPs within 500kb of each other on the same chromosome. GWA identified up to 139 associated regions in the Welsh ponies and 142 associated regions in the Morgan horses. The boundaries of GWA regions were defined based on a fixed-size (500kb 5' to the minimum SNP and 3' to the maximum SNP) or based on the breakdown of linkage disequilibrium (LD). Approximately 60% of the fixed-size boundaries were found to be larger than the LD boundaries and likely indicates that our fixed-size boundaries were overestimating the region size and including candidate genes that were not in LD with the marker SNP. For the remaining regions, the LD boundaries were on average >1Mb longer then the fixed-size region, likely indicating regions in which the fixed-size boundaries were underestimating the region size and excluding candidate genes which could include the

causal variant. These data indicate that identification of breed and locus-specific LD is imperative to precisely identifying positional candidate genes.

GWA meta-analysis using a random effects model was performed in order to identify GWA regions shared between breeds. Meta-analysis uses GWA summary statistics to effectively combine GWA studies, increasing the number of individuals within the study and improving the power to find unique associations, variants of low effect, and additional shared regions across populations Meta-analysis demonstrated that 65 of the 272 regions were shared across breeds. These data support that EMS risk alleles are shared across breed as well as breed-specific.

In order to reduce false positives, GWA regions were prioritized as regions of interest (ROI) if they contained a minimum of five SNPs that exceeded the suggestive threshold and at least one SNP that exceeded the threshold for genome wide significance. Regions shared across breeds (based on meta-analysis) were given high priority, as these regions were not breed specific and likely to be found in other high-risk breeds. Regions shared across traits with at least one ROI were also assigned high priority, as a variant affecting multiple traits would be expected to have a larger biological effect then a variant affecting a single trait. An ROI identified in one GWA cohort was assigned medium priority as these regions were likely breed or section (Welsh pony) specific and, based on the power of our study, variants of moderate to high effect. Finally, regions that were not ROI but shared across traits were assigned low priority. Within breed prioritization of the LD-defined regions resulted in 56 high priority, 26 medium priority, and 7 low priority regions, for a total of 1,853 candidate genes in the Welsh ponies; and 39 high priority, 8 medium priority, and 9 low priority regions, for a total of 1,167 candidate genes in the Morgan horses. These data clearly support the hypothesis that EMS is a polygenic trait.

In chapter 4, we provided intriguing biological evidence for the role of several coding genes in the pathogenesis of EMS. We utilized imputed whole-genome sequencing (WGS) and linear regression analysis in order to fine-map selected high priority LD-ROI in both the Morgan horses and Welsh ponies. LD-ROI were fine-mapped if they contained at least 5 SNPs with one SNP exceeding the threshold for genome-wide significance. Five fine-mapped regions from each breed were further interrogated for predicted impact using

variant annotation. First, all variants which exceeded the threshold for genome-wide significance mapped to non-coding regions of the genome, with 66.7% of the significantly associated SNPs being intronic, 17.0% intergenic and 10.3% within lncRNA. Second, in order to capture potential causal variants within protein-coding genes, we further evaluated positional candidate genes with exonic variants in our fine-mapped region with a p-value <0.05 (i.e. "sub-threshold"). Protein-coding genes containing non-coding or coding variants within the fine-mapping region were then further prioritized based on known function and biological evidence in other species utilizing the PubMed search engine. A total of 19 positional candidate genes were identified as having biological evidence for a role in EMS including: *SSTR1, SEC23A, FBXO33, MIA2, EIF3D, CSF2B, IFT27, ACE, TACO1, ABCA13, NKAIN2, BBX, XXYLT1, BDH1, NCKAP5L, GPD1, LIAA1, METTL7A, SCL11A2.* These data provide support for the process of fine-mapping GWA ROI by increasing marker density and using biological evidence across species to further prioritize candidate genes.

In chapter 5, we provided the first report of a gene with a pleotropic effect for EMS and provided evidence for the underlying cause of the unique metabolic profiles and increased EMS susceptibility in ponies. Pearson's correlation coefficient identified an inverse relationship between height and baseline insulin in a cohort of 264 Welsh ponies. Genome-wide association analyses of height and insulin revealed the same ~1.3 Mb region on chromosome 6, which was also identified using a di statistic for genomic signatures of selection. Haplotype analysis confirmed that there was a shared ancestral haplotype between height and insulin. The high mobility group AT-hook (*HMGA2)* was identified as a candidate gene, and sequencing identified a single a c.83G>A variant (p.G28E) in *HMGA2,* previously described in other small stature horse breeds [574]. In the EMS cohort of ponies, the A allele had a frequency of .76, was strongly correlated with height (-.75) and was low to moderately correlated with metabolic traits including: insulin (.32), insulin after an oral sugar test (.25), non-esterified fatty acids (.19) and triglyceride (.22) concentrations. For this allele, model analysis suggested an additive mode of inheritance with height and a recessive mode of inheritance with the metabolic traits. In humans, a relationship between individuals of short stature and an increased risk of metabolic

syndrome has been well described. Thus, these data also provide that first phase of using EMS as a model for translational genomics for MetS.

In conclusion, we have provided strong evidence supporting our hypotheses that many traits that comprise the EMS phenotype are moderately to highly heritable, that major genetic risk factors leading to EMS and laminitis susceptibility are shared across breeds of horses, and that differences exist in the severity and secondary features of the EMS phenotype between breeds, or between individuals within a breed. We further hypothesize that such breed differences are the result of modifying genetic risk alleles with variable frequencies between breeds.

Future Directions

The results described in this thesis are major steps forward in understanding the genetic contributions of EMS in two high risk breeds, but still only represents a small piece of a very large puzzle. Thus, additional work is required to continue to unravel the risk alleles contributing to EMS.

In chapter 4, we provided intriguing biological evidence for the role of multiple coding genes in the pathogenesis of EMS but did not conclusively identify the causal variants. Several methods could be utilized to further interrogate our regions for both protein-coding and non-coding causal variants. First, interrogation of the allele frequency of the variants identified in chapter 4 in a large database of mixed-breed horses would allow for the assessment of the frequency of these variants in healthy horses. Given that EMS can manifest at an older age, of particular interest would be assessment in breeds considered low-risk for EMS such as the Quarter Horse or Thoroughbred. Identification of variants at low frequency in these breeds would allow for the prioritization of specific biological candidate genes for interrogation through Sanger sequencing. Second, haplotype analysis can be utilized to find shared ancestral haplotypes to further fine-map the LD-ROI, prioritize variants, and identify additional horses for whole-genome or Sanger sequencing. Third, development of a custom genotyping assay would allow for the validation of imputed genotypes as well as assessment of the statistically significant coding and non-coding variants in an independent population of Welsh ponies and Morgan horses phenotyped for EMS.

In chapter 3, we identified hundreds of regions of the genome associated with EMS relevant traits in both Welsh ponies and Morgan horses; however, this likely only explains a small fraction of the genetic variance for EMS. Several factors have been proposed to contribute to the missing heritability in GWA studies, including study power, stringent thresholds for genome-wide significance, exclusion of rare variants, and the omnigenic hypothesis (see **chapter 1** for further discussion about the missing heritability in GWA), which could all contribute to underestimation of associated loci in our analysis. In addition, we utilized a prioritization method in order to select GWA region for analysis, which could have further excluded regions of the genome containing EMS risk alleles. This method also gave a lower ranking to several interesting GWA associated regions. For example, the region on chromosome 6 harboring the *HMGA2* EMS risk allele in Welsh ponies (see **chapter 5**) was given a medium priority and was not in the initial fine-mapping of high priority LD-ROI as described. Further, the genetic loci identified in our GWA does not explain the genetic variation across all breeds. As noted throughout this thesis, there are several breeds considered high risk for EMS, of which most will likely have breed-specific risk alleles. Therefore, future directions include (i) assessment of the medium and low priority regions, (ii) increasing population size in both breeds with a more equal representation of Welsh pony sections in order to improve the power for variant detection, and (iii) expanding these analyses into other high-risk breeds.

However, given the complexity of the EMS phenotype and the high percentage of associated SNPs located in non-coding regions of the genome, an integrated, trait-relevant multi-omics interrogation of the EMS phenotype will be required to fully investigate the genetic contribution to this syndrome. The rational for this approach lies in the fact that alleles can exert their effects through alterations in gene (transcriptome) or protein (proteome) expression, regulation, or function, which manifest through alterations in metabolic pathways and functions (metabolomics). Thus, each type of omics data provides a list of differences associated with a disease, but only represents one layer of the genetic effect and is typically limited to correlations versus causation [399]. By interrogating a trait with multiple levels of the genome as intermediate phenotypes, or molecular traits that are precursors for the "endpoint of interest", such as trait characteristics or clinical measurements, multi-omics datasets can be used as powerful tools to unravel the causative

changes and therapeutic targets of complex traits [375,399]. For example, transcriptomic intermediate phenotypes are the closest to the genome. Risk alleles within a gene's regulatory regions (promoter, untranslated regions) or regulatory elements (transcription factors, microRNA) can alter a gene's expression. Therefore, alterations in the transcriptome reflect the downstream effect of genome variants on gene regulatory mechanisms and can be used to uncover pathways targeting the phenome. However, risk alleles that alter protein function or regulation may not be directly reflected by changes in gene expression but can manifest by alterations in the metabolome, which lie in the closest proximity to the phenome. Therefore, the metabolome provides insight into the metabolic effect of trait-associated genetic variants [600] and has been used to co-map and refine regions of the genome [601] based on their closer proximity to the phenome.

Therefore, bridging the connections across multiple omics datasets constitutes a powerful approach to explaining the relationship between genotype and phenotype, and leveraging these data has the potential to the lead to the identification of dozens of genetic risk alleles for EMS and provide valuable insight into its molecular pathophysiology. The following plan proposes a workflow for using multi-omics data from the cohort presented in this thesis. (i) *Relationships between the genome, metabolome and phenome:* The serum metabolome of the 286 Morgans and 264 Welsh ponies can be quantified using liquid or gas chromatography coupled to mass spectrometry. These data can then be used to determine which metabolites are statistically associated with EMS clinical phenotypes. Due to the complexity of the EMS phenotype, analyses would need to be performed with EMS traits as quantitative response variables, relative abundance of the metabolites as predictor variables, age and sex as fixed effects, and environment as a random effect. Genetic risk loci for metabolites can also be identified by within-breed genome-wide association, and statistically significant regions on GWA will be assessed to determine if they co-map with within-breed GWA regions identified previously for EMS traits. Co-mapped regions could be interrogated using imputed whole genome sequencing and haplotype analysis for genetic variant and candidate gene discovery. (ii) *Relationships between the transcriptome, phenome, and genome:* Muscle and fat biopsies from selected Morgans and Welsh ponies phenotype for EMS could be collected in order to assess differential gene expression in these tissues using RNA-seq. Regression-based mediation

analysis could be used to assess the effect of the SNP GWA regions of interest and gene expression data. (iii) *Relationship between the genome, transcriptome and metabolome:* Evaluation of metabolomic and transcriptomic co-expression networks could also be analyzed for network interactions. Assessment of network locality could identify positional candidate genes that are proportionally more connected to genes locally (candidate genes from GWA) than they are globally (all genes in the network).

In conclusion, the continued investigation into the genetic risk factors contributing to EMS is necessary in order to gain a better understanding of the pathophysiology of this syndrome as well identify enough risk alleles to create a validated genetic test. Given that this thesis has proven that EMS is a polygenic disease, a genetic test assessing a single risk allele or locus will provide limited information regarding overall disease risk. Instead, genetic testing would require a genotyping array containing several risk loci of high, moderate and low impact. In addition, the array would need to be representative of risk loci that are shared across high risk breeds as well as those that are breed-specific Ideally, variants present on the genotyping array would explain at least half of the genetic variation of EMS across breeds but the number of variants on the array is ultimately a function of the predictive power of the variants.

Bibliography

1. Katz, L.M.; Bailey, S.R. A review of recent advances and current hypotheses on the pathogenesis of acute laminitis. *Equine Vet J* 2012, *44*, 752-761, doi:10.1111/j.2042-3306.2012.00664.x.
2. Driessen, B.; Bauquier, S.H.; Zarucco, L. Neuropathic pain management in chronic laminitis. *Vet Clin North Am Equine Pract* 2010, *26*, 315-337, doi:10.1016/j.cveq.2010.04.002.
3. Frank, N.; Geor, R.J.; Bailey, S.R.; Durham, A.E.; Johnson, P.J. Equine metabolic syndrome. *J Vet Intern Med* 2010, *24*, 467-475, doi:10.1111/j.1939-1676.2010.0503.x.
4. Karikoski, N.P.; Horn, I.; McGowan, T.W.; McGowan, C.M. The prevalence of endocrinopathic laminitis among horses presented for laminitis at a first-opinion/referral equine hospital. *Domest Anim Endocrinol* 2011, *41*, 111-117, doi:10.1016/j.domaniend.2011.05.004.
5. Jeffcott, L.B.; Field, J.R.; McLean, J.G.; O'Dea, K. Glucose tolerance and insulin sensitivity in ponies and Standardbred horses. *Equine Vet J* 1986, *18*, 97-101.
6. Coffman, J.R.; Colles, C.M. Insulin tolerance in laminitic ponies. *Can J Comp Med* 1983, *47*, 347-351.
7. Johnson, P.J. The equine metabolic syndrome peripheral Cushing's syndrome. *Vet Clin North Am. Equine Pract* 2002, *18*, 271-293.
8. Alberti, K.G.; Eckel, R.H.; Grundy, S.M.; Zimmet, P.Z.; Cleeman, J.I.; Donato, K.A.; Fruchart, J.C.; James, W.P.; Loria, C.M.; Smith, S.C., Jr. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009, *120*, 1640-1645, doi:10.1161/circulationaha.109.192644.
9. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults--The Evidence Report. National Institutes of Health. *Obes Res* 1998, *6 Suppl 2*, 51S-209S.
10. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 2001, *285*, 2486-2497.
11. Alberti, K.G.; Zimmet, P.; Shaw, J. The metabolic syndrome--a new worldwide definition. *Lancet* 2005, *366*, 1059-1062, doi:10.1016/s0140-6736(05)67402-8.
12. Alberti, K.G.; Zimmet, P.Z. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998, *15*, 539-553, doi:10.1002/(sici)1096-9136(199807)15:7<539::aid-dia668>3.0.co;2-s.
13. Treiber, K.H.; Kronfeld, D.S.; Hess, T.M.; Byrd, B.M.; Splan, R.K.; Staniar, W.B. Evaluation of genetic and metabolic predispositions and nutritional risk factors for pasture-associated laminitis in ponies. *J Am Vet Med Assoc* 2006, *228*, 1538-1545, doi:10.2460/javma.228.10.1538.

14. Bailey, S.R.; Habershon-Butcher, J.L.; Ransom, K.J.; Elliott, J.; Menzies-Gow, N.J. Hypertension and insulin resistance in a mixed-breed population of ponies predisposed to laminitis. *Am J Vet Res* 2008, *69*, 122-129, doi:10.2460/ajvr.69.1.122.
15. Frank, N.; Elliott, S.B.; Brandt, L.E.; Keisler, D.H. Physical characteristics, blood hormone concentrations, and plasma lipid concentrations in obese horses with insulin resistance. *J Am Vet Med Assoc* 2006, *228*, 1383-1390, doi:10.2460/javma.228.9.1383.
16. Carter, R.A.; Treiber, K.H.; Geor, R.J.; Douglass, L.; Harris, P.A. Prediction of incipient pasture-associated laminitis from hyperinsulinaemia, hyperleptinaemia and generalised and localised obesity in a cohort of ponies. *Equine Vet J* 2009, *41*, 171-178.
17. Hess, T.M.; Kronfeld, D.S.; Treiber, K.H.; Byrd, B.M.; Staniar, W.B.; Splan, R.K. Laminitic metabolic profile in genetically predisposed ponies involves exaggerated compensated insulin resistance. *J Anim Physiol Anim Nutr* 2005, *89*, 431-431, doi:10.1111/j.1439-0396.2005.00611_9.x.
18. Suagee, J.K.; Corl, B.A.; Crisman, M.V.; Pleasant, R.S.; Thatcher, C.D.; Geor, R.J. Relationships between body condition score and plasma inflammatory cytokines, insulin, and lipids in a mixed population of light-breed horses. *J Vet Intern Med* 2013, *27*, 157-163, doi:10.1111/jvim.12021.
19. McCue, M.E.; Geor, R.J.; Schultz, N. Equine Metabolic Syndrome: A Complex Disease Influenced by Genetics and the Environment. *J Equine Vet Sci* 2015, *35*, 367-375, doi:https://doi.org/10.1016/j.jevs.2015.03.004.
20. Schultz, N. Characterization of equine metabolic syndrome and mapping of candidate genetic loci. PhD Dissertation, University of Minnesota, Saint Paul, MN, 2016.
21. Durham, A.E.; Frank, N. ECEIM consensus statement on equine metabolic syndrome. 2019, *33*, 335-349, doi:10.1111/jvim.15423.
22. Morgan, R.; Keen, J.; McGowan, C. Equine metabolic syndrome. *Vet Rec* 2015, *177*, 173-179, doi:10.1136/vr.103226.
23. Muno, J.; Gallatin L.; Geor, R.J. Prevalence and risk factors for hyperinsulinemia in clinically normal horses in central Ohio (Abstract). *J Vet Intern Med* 2009, *23*, 721.
24. Morgan, R.A.; McGowan, T.W.; McGowan, C.M. Prevalence and risk factors for hyperinsulinaemia in ponies in Queensland, Australia. *Aust Vet J* 2014, *92*, 101-106, doi:10.1111/avj.12159.
25. Sillence, M.N. 'Supersize me': on equine obesity. *Veterinary Journal (London, England : 1997)* 2012, *194*, 137-138, doi:10.1016/j.tvjl.2012.08.004.
26. Giles, S.L.; Rands, S.A.; Nicol, C.J.; Harris, P.A. Obesity prevalence and associated risk factors in outdoor living domestic horses and ponies. *PeerJ* 2014, *2*, e299, doi:10.7717/peerj.299.
27. Kosolofski, H.R.; Gow, S.P.; Robinson, K.A. Prevalence of obesity in the equine population of Saskatoon and surrounding area. *Can Vet J* 2017, *58*, 967-970.
28. Potter, S.J.; Bamford, N.J.; Harris, P.A.; Bailey, S.R. Prevalence of obesity and owners' perceptions of body condition in pleasure horses and ponies in south-eastern Australia. *Aust Vet J* 2016, *94*, 427-432, doi:10.1111/avj.12506.

29. Robin, C.A.; Ireland, J.L.; Wylie, C.E.; Collins, S.N.; Verheyen, K.L.; Newton, J.R. Prevalence of and risk factors for equine obesity in Great Britain based on owner-reported body condition scores. *Equine Vet J* 2015, *47*, 196-201, doi:10.1111/evj.12275.
30. Thatcher, C.D.; Pleasant, R.S.; Geor, R.J.; Elvinger, F. Prevalence of overconditioning in mature horses in southwest Virginia during the summer. *J Vet Intern Med* 2012, *26*, 1413-1418, doi:10.1111/j.1939-1676.2012.00995.x.
31. Wyse, C.A.; McNie, K.A.; Tannahill, V.J.; Murray, J.K.; Love, S. Prevalence of obesity in riding horses in Scotland. *Vet Rec* 2008, *162*, 590-591.
32. Carter, R.A.; Geor, R.J.; Burton Staniar, W.; Cubitt, T.A.; Harris, P.A. Apparent adiposity assessed by standardised scoring systems and morphometric measurements in horses and ponies. *Veterinary Journal (London, England : 1997)* 2009, *179*, 204-210, doi:10.1016/j.tvjl.2008.02.029.
33. Sanchez, M.J.; Azor, P.J.; Molina, A.; Parkin, T.; Rivero, J.L.; Valera, M. Prevalence, risk factors and genetic parameters of cresty neck in Pura Raza Espanol horses. *Equine Vet J* 2017, *49*, 196-200, doi:10.1111/evj.12569.
34. Giles, S.L.; Nicol, C.J.; Rands, S.A.; Harris, P.A. Assessing the seasonal prevalence and risk factors for nuchal crest adiposity in domestic horses and ponies using the Cresty Neck Score. *BMC Vet Res* 2015, *11*, 13, doi:10.1186/s12917-015-0327-7.
35. Frank, N. Equine metabolic syndrome. *Vet Clin North Am. Equine Pract* 2011, *27*, 73-92, doi:10.1016/j.cveq.2010.12.004.
36. Jeffcott, L.B.; Field, J.R. Current concepts of hyperlipaemia in horses and ponies. *Vet Rec* 1985, *116*, 461-466.
37. Robie, S.M.; Janson, C.H.; Smith, S.C.; O'Connor, J.T., Jr. Equine serum lipids: serum lipids and glucose in Morgan and Thoroughbred horses and Shetland ponies. *Am J Vet Res* 1975, *36*, 1705-1708.
38. Bamford, N.J.; Potter, S.J.; Harris, P.A.; Bailey, S.R. Breed differences in insulin sensitivity and insulinemic responses to oral glucose in horses and ponies of moderate body condition score. *Domest Anim Endocrinol* 2014, *47*, 101-107, doi:10.1016/j.domaniend.2013.11.001.
39. Prentice, A.M. Early influences on human energy regulation: thrifty genotypes and thrifty phenotypes. *Physiol Behav* 2005, *86*, 640-645, doi:10.1016/j.physbeh.2005.08.055.
40. Neel, J.V. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? *Am J Hum Genet* 1962, *14*, 353-362.
41. Blaue, D.; Schedlbauer, C.; Starzonek, J.; Gittel, C.; Brehm, W.; Einspanier, A.; Vervuert, I. Effects of body weight gain on insulin and lipid metabolism in equines. *Domest Anim Endocrinol* 2019, *68*, 111-118, doi:10.1016/j.domaniend.2019.01.003.
42. Carter, R.A.; McCutcheon, L.J.; George, L.A.; Smith, T.L.; Frank, N.; Geor, R.J. Effects of diet-induced weight gain on insulin sensitivity and plasma hormone and lipid concentrations in horses. *Am J Vet Res*2009, *70*, 1250-1258, doi:10.2460/ajvr.70.10.1250.
43. Lindase, S.S.; Nostell, K.E.; Muller, C.E.; Jensen-Waern, M.; Brojer, J.T. Effects of diet-induced weight gain and turnout to pasture on insulin sensitivity in

moderately insulin resistant horses. *Am J Vet Res*2016, *77*, 300-309, doi:10.2460/ajvr.77.3.300.
44. Quinn, R.W.; Burk, A.O.; Hartsock, T.G.; Petersen, E.D.; Whitley, N.C.; Treiber, K.H.; Boston, R.C. Insulin Sensitivity in Thoroughbred Geldings: Effect of Weight Gain, Diet, and Exercise on Insulin Sensitivity in Thoroughbred Geldings. *J Equine Vet Sci* 2008, *28*, 728-738, doi:https://doi.org/10.1016/j.jevs.2008.10.020.
45. Hoffman, R.M.; Boston, R.C.; Stefanovski, D.; Kronfeld, D.S.; Harris, P.A. Obesity and diet affect glucose dynamics and insulin sensitivity in Thoroughbred geldings. *J Anim Sci* 2003, *81*, 2333-2342, doi:10.2527/2003.8192333x.
46. Treiber, K.H.; Boston, R.C.; Kronfeld, D.S.; Staniar, W.B.; Harris, P.A. Insulin resistance and compensation in Thoroughbred weanlings adapted to high-glycemic meals. *J Anim Sci* 2005, *83*, 2357-2364, doi:10.2527/2005.83102357x.
47. Jacob, S.I.; Geor, R.J.; Weber, P.S.D.; Harris, P.A.; McCue, M.E. Effect of age and dietary carbohydrate profiles on glucose and insulin dynamics in horses. *Equine Vet J* 2018, *50*, 249-254, doi:10.1111/evj.12745.
48. Bamford, N.J.; Potter, S.J.; Harris, P.A.; Bailey, S.R. Effect of increased adiposity on insulin sensitivity and adipokine concentrations in horses and ponies fed a high fat diet, with or without a once daily high glycaemic meal. *Equine Vet J* 2016, *48*, 368-373, doi:10.1111/evj.12434.
49. Bluher, M. The distinction of metabolically 'healthy' from 'unhealthy' obese individuals. *Curr Opin Lipidol* 2010, *21*, 38-43, doi:10.1097/MOL.0b013e3283346ccc.
50. Manfredi, J. Identifying breed differences in insulin dynamics, skeletal muscle and adipose tissue histology and gene expression. PhD Dissertation, Michigan State University, 2016.
51. Bamford, N.J.; Potter, S.J.; Baskerville, C.L.; Harris, P.A.; Bailey, S.R. Effect of increased adiposity on insulin sensitivity and adipokine concentrations in different equine breeds adapted to cereal-rich or fat-rich meals. *Veterinary Journal (London, England : 1997)* 2016, *214*, 14-20, doi:10.1016/j.tvjl.2016.02.002.
52. Selim, S.; Elo, K.; Jaakkola, S.; Karikoski, N.; Boston, R.; Reilas, T.; Sarkijarvi, S.; Saastamoinen, M.; Kokkonen, T. Relationships among Body Condition, Insulin Resistance and Subcutaneous Adipose Tissue Gene Expression during the Grazing Season in Mares. *PloS One* 2015, *10*, e0125968, doi:10.1371/journal.pone.0125968.
53. Coleman, M.C.; Belknap, J.K.; Eades, S.C.; Galantino-Homer, H.L.; Hunt, R.J.; Geor, R.J.; McCue, M.E.; McIlwraith, C.W.; Moore, R.M.; Peroni, J.F., et al. Case-control study of risk factors for pasture-and endocrinopathy-associated laminitis in North American horses. *J Am Vet Med Assoc*2018, *253*, 470-478, doi:10.2460/javma.253.4.470.
54. Arnold, W.; Ruf, T.; Kuntz, R. Seasonal adjustment of energy budget in a large wild mammal, the Przewalski horse (Equus ferus przewalskii) II. Energy expenditure. *J Exp Biol* 2006, *209*, 4566-4573, doi:10.1242/jeb.02536.
55. Brinkmann, L.; Gerken, M.; Riek, A. Adaptation strategies to seasonal changes in environmental conditions of a domesticated horse breed, the Shetland pony (Equus ferus caballus). *J Exp Biol* 2012, *215*, 1061-1068, doi:10.1242/jeb.064832.

56. Dugdale, A.H.; Curtis, G.C.; Cripps, P.J.; Harris, P.A.; Argo, C.M. Effects of season and body condition on appetite, body mass and body composition in ad libitum fed pony mares. *Veterinary Journal (London, England : 1997)* 2011, *190*, 329-337, doi:10.1016/j.tvjl.2010.11.009.
57. Buff, P.R.; Messer, N.T.t.; Cogswell, A.M.; Johnson, P.J.; Keisler, D.H.; Ganjam, V.K. Seasonal and pulsatile dynamics of thyrotropin and leptin in mares maintained under a constant energy balance. *Domest Anim Endocrinol* 2007, *33*, 430-436, doi:10.1016/j.domaniend.2006.08.007.
58. Wray, H.; Elliott, J.; Bailey, S.R.; Harris, P.A.; Menzies-Gow, N.J. Plasma concentrations of inflammatory markers in previously laminitic ponies. *Equine Vet J* 2013, *45*, 546-551, doi:10.1111/evj.12031.
59. Donaldson, M.T.; McDonnell, S.M.; Schanbacher, B.J.; Lamb, S.V.; McFarlane, D.; Beech, J. Variation in plasma adrenocorticotropic hormone concentration and dexamethasone suppression test results with season, age, and sex in healthy ponies and horses. *J Vet Intern Med*2005, *19*, 217-222, doi:10.1892/0891-6640(2005)19<217:vipahc>2.0.co;2.
60. Place, N.J.; McGowan, C.M.; Lamb, S.V.; Schanbacher, B.J.; McGowan, T.; Walsh, D.M. Seasonal variation in serum concentrations of selected metabolic hormones in horses. *J Vet Intern Med*2010, *24*, 650-654, doi:10.1111/j.1939-1676.2010.0500.x.
61. Schreiber, C.M.; Stewart, A.J.; Kwessi, E.; Behrend, E.N.; Wright, J.C.; Kemppainen, R.J.; Busch, K.A. Seasonal variation in results of diagnostic tests for pituitary pars intermedia dysfunction in older, clinically normal geldings. *J Am Vet Med Assoc*2012, *241*, 241-248, doi:10.2460/javma.241.2.241.
62. Frank, N.; Elliott, S.B.; Chameroy, K.A.; Toth, F.; Chumbler, N.S.; McClamroch, R. Association of season and pasture grazing with blood hormone and metabolite concentrations in horses with presumed pituitary pars intermedia dysfunction. *J Vet Intern Med*2010, *24*, 1167-1175, doi:10.1111/j.1939-1676.2010.0547.x.
63. Copas, V.E.; Durham, A.E. Circannual variation in plasma adrenocorticotropic hormone concentrations in the UK in normal horses and ponies, and those with pituitary pars intermedia dysfunction. *Equine Vet J* 2012, *44*, 440-443, doi:10.1111/j.2042-3306.2011.00444.x.
64. Secombe, C.J.; Tan, R.H.H.; Perara, D.I.; Byrne, D.P.; Watts, S.P.; Wearn, J.G. The Effect of Geographic Location on Circannual Adrenocorticotropic Hormone Plasma Concentrations in Horses in Australia. *J Vet Intern Med* 2017, *31*, 1533-1540, doi:10.1111/jvim.14782.
65. Hart, K.A.; Wochele, D.M.; Norton, N.A.; McFarlane, D.; Wooldridge, A.A.; Frank, N. Effect of Age, Season, Body Condition, and Endocrine Status on Serum Free Cortisol Fraction and Insulin Concentration in Horses. *J Vet Intern Med* 2016, *30*, 653-663, doi:10.1111/jvim.13839.
66. Borer-Weir, K.E.; Menzies-Gow, N.J.; Bailey, S.R.; Harris, P.A.; Elliott, J. Seasonal and annual influence on insulin and cortisol results from overnight dexamethasone suppression tests in normal ponies and ponies predisposed to laminitis. *Equine Vet J* 2013, *45*, 688-693, doi:10.1111/evj.12053.
67. Funk, R.A.; Wooldridge, A.A.; Stewart, A.J.; Behrend, E.N.; Kemppainen, R.J.; Zhong, Q.; Johnson, A.K. Seasonal changes in the combined glucose-insulin

tolerance test in normal aged horses. *J Vet Intern Med* 2012, *26*, 1035-1041, doi:10.1111/j.1939-1676.2012.00939.x.
68. Nielsen, B.D.; O'Connor-Robison, C.I.; Spooner, H.S.; Shelton, J. Glycemic and Insulinemic Responses Are Affected by Age of Horse and Method of Feed Processing. *J Equine Vet Sci* 2010, *30*, 249-258, doi:https://doi.org/10.1016/j.jevs.2010.03.008.
69. Rapson, J.L.; Schott II, H.C.; Nielsen, B.D.; McCutcheon, L.J.; Harris, P.A.; Geor, R.J. Effects of age and diet on glucose and insulin dynamics in the horse. *Equine Vet J* 2018, *50*, 690-696, doi:10.1111/evj.12812.
70. Kawasumi, K.; Yamamoto, M.; Koide, M.; Okada, Y.; Mori, N.; Yamamoto, I.; Arai, T. Aging effect on plasma metabolites and hormones concentrations in riding horses. *Open Vet J* 2015, *5*, 154-157.
71. Adams, A.A.; Katepalli, M.P.; Kohler, K.; Reedy, S.E.; Stilz, J.P.; Vick, M.M.; Fitzgerald, B.P.; Lawrence, L.M.; Horohov, D.W. Effect of body condition, body weight and adiposity on inflammatory cytokine responses in old horses. *Vet Immunol Immunopathol* 2009, *127*, 286-294, doi:10.1016/j.vetimm.2008.10.323.
72. Cartmill, J.A.; Thompson, D.L., Jr.; Del Vecchio, R.P.; Storer, W.A.; Crowley, J.C. Leptin secretion in horses: effects of dexamethasone, gender, and testosterone. *Domest Anim Endocrinol* 2006, *31*, 197-210, doi:10.1016/j.domaniend.2005.10.006.
73. de Laat, M.A.; Reiche, D.B.; Sillence, M.N.; McGree, J.M. Incidence and risk factors for recurrence of endocrinopathic laminitis in horses. *J Vet Intern Med* 2019, *33*, 1473-1482, doi:10.1111/jvim.15497.
74. Durward-Akhurst, S.A.; Schultz, N.E.; Norton, E.M.; Rendahl, A.K.; Besselink, H.; Behnisch, P.A.; Brouwer, A.; Geor, R.J.; Mickelson, J.R.; McCue, M.E. Associations between endocrine disrupting chemicals and equine metabolic syndrome phenotypes. *Chemosphere* 2019, *218*, 652-661, doi:10.1016/j.chemosphere.2018.11.136.
75. Orsini, J.A.; Parsons, C.S.; Capewell, L.; Smith, G. Prognostic indicators of poor outcome in horses with laminitis at a tertiary care hospital. *Can Vet J* 2010, *51*, 623-628.
76. Karikoski, N.P.; McGowan, C.M.; Singer, E.R.; Asplin, K.E.; Tulamo, R.M.; Patterson-Kane, J.C. Pathology of Natural Cases of Equine Endocrinopathic Laminitis Associated With Hyperinsulinemia. *Vet Pathol* 2015, *52*, 945-956, doi:10.1177/0300985814549212.
77. USDA. Lameness and laminitis in U.S horses #N318.0400. APHIS, V., CEAH, National Animal Health Monitoring System, Ed. Fort Collins, CO, 2000.
78. McGowan, C.M. Endocrinopathic laminitis. *Vet Clin North Am Equine Pract* 2010, *26*, 233-237, doi:10.1016/j.cveq.2010.04.009.
79. Patterson-Kane, J.C.; Karikoski, N.P.; McGowan, C.M. Paradigm shifts in understanding equine laminitis. *Veterinary Journal (London, England : 1997)* 2018, *231*, 33-40, doi:10.1016/j.tvjl.2017.11.011.
80. Asplin, K.E.; Patterson-Kane, J.C.; Sillence, M.N.; Pollitt, C.C.; Mc Gowan, C.M. Histopathology of insulin-induced laminitis in ponies. *Equine Vet J* 2010, *42*, 700-706, doi:10.1111/j.2042-3306.2010.00111.x.

81. de Laat, M.A.; Patterson-Kane, J.C.; Pollitt, C.C.; Sillence, M.N.; McGowan, C.M. Histological and morphometric lesions in the pre-clinical, developmental phase of insulin-induced laminitis in Standardbred horses. *Veterinary Journal (London, England : 1997)* 2013, *195*, 305-312, doi:10.1016/j.tvjl.2012.07.003.
82. Watts, M.R.; Hegedus, O.C.; Eades, S.C.; Belknap, J.K. Association of sustained supraphysiologic hyperinsulinemia and inflammatory signaling within the digital lamellae in light-breed horses. *J Vet Intern Med* 2019, *33*, 1483-1492, doi:10.1111/jvim.15480.
83. Burns, T.A.; Watts, M.R.; Weber, P.S.; McCutcheon, L.J.; Geor, R.J.; Belknap, J.K. Laminar inflammatory events in lean and obese ponies subjected to high carbohydrate feeding: Implications for pasture-associated laminitis. *Equine Vet J* 2015, *47*, 489-493, doi:10.1111/evj.12314.
84. de Laat, M.A.; Clement, C.K.; McGowan, C.M.; Sillence, M.N.; Pollitt, C.C.; Lacombe, V.A. Toll-like receptor and pro-inflammatory cytokine expression during prolonged hyperinsulinaemia in horses: implications for laminitis. *Vet Immunol Immunopathol* 2014, *157*, 78-86, doi:10.1016/j.vetimm.2013.10.010.
85. de Laat, M.A.; McGowan, C.M.; Sillence, M.N.; Pollitt, C.C. Equine laminitis: induced by 48 h hyperinsulinaemia in Standardbred horses. *Equine Vet J* 2010, *42*, 129-135, doi:10.2746/042516409x475779.
86. de Laat, M.A.; Sillence, M.N.; McGowan, C.M.; Pollitt, C.C. Continuous intravenous infusion of glucose induces endogenous hyperinsulinaemia and lamellar histopathology in Standardbred horses. *Veterinary Journal (London, England : 1997)* 2012, *191*, 317-322, doi:10.1016/j.tvjl.2011.07.007.
87. Asplin, K.E.; Sillence, M.N.; Pollitt, C.C.; McGowan, C.M. Induction of laminitis by prolonged hyperinsulinaemia in clinically normal ponies. *Veterinary Journal (London, England : 1997)* 2007, *174*, 530-535, doi:10.1016/j.tvjl.2007.07.003.
88. de Laat, M.A.; Clement, C.K.; Sillence, M.N.; McGowan, C.M.; Pollitt, C.C.; Lacombe, V.A. The impact of prolonged hyperinsulinaemia on glucose transport in equine skeletal muscle and digital lamellae. *Equine Vet J* 2014, 10.1111/evj.12320, doi:10.1111/evj.12320.
89. de Laat, M.A.; Sillence, M.N.; Reiche, D.B. Phenotypic, hormonal, and clinical characteristics of equine endocrinopathic laminitis. *J Vet Intern Med* 2019, *33*, 1456-1463, doi:10.1111/jvim.15419.
90. Walsh, D.M.; McGowan, C.M.; McGowan, T.; Lamb, S.V.; Schanbacher, B.J.; Place, N.J. Correlation of Plasma Insulin Concentration with Laminitis Score in a Field Study of Equine Cushing's Disease and Equine Metabolic Syndrome. *J Equine Vet Sci* 2009, *29*, 87-94, doi:https://doi.org/10.1016/j.jevs.2008.12.006.
91. Burns, T.A.; Watts, M.R.; Weber, P.S.; Mccutcheon, L.J.; Geor, R.J.; Belknap, J.K. Distribution of insulin receptor and insulin-like growth factor-1 receptor in the digital laminae of mixed-breed ponies: An immunohistochemical study. *Equine Vet J* 2013, *45*, 326-332, doi:10.1111/j.2042-3306.2012.00631.x.
92. Ciaraldi, T.P.; Sasaoka, T. Review on the in vitro interaction of insulin glargine with the insulin/insulin-like growth factor system: potential implications for metabolic and mitogenic activities. *Horm Metab Res* 2011, *43*, 1-10, doi:10.1055/s-0030-1267203.

93. Doyle, S.L.; Donohoe, C.L.; Finn, S.P.; Howard, J.M.; Lithander, F.E.; Reynolds, J.V.; Pidgeon, G.P.; Lysaght, J. IGF-1 and its receptor in esophageal cancer: association with adenocarcinoma and visceral obesity. *Am J Gastroenterol* 2012, *107*, 196-204, doi:10.1038/ajg.2011.417.
94. Brady, G.; Crean, S.J.; Naik, P.; Kapas, S. Upregulation of IGF-2 and IGF-1 receptor expression in oral cancer cell lines. *Int J Oncol* 2007, *31*, 875-881.
95. de Laat, M.A.; Pollitt, C.C.; Kyaw-Tanner, M.T.; McGowan, C.M.; Sillence, M.N. A potential role for lamellar insulin-like growth factor-1 receptor in the pathogenesis of hyperinsulinaemic laminitis. *Veterinary Journal (London, England : 1997)* 2013, *197*, 302-306, doi:10.1016/j.tvjl.2012.12.026.
96. Kullmann, A.; Weber, P.S.; Bishop, J.B.; Roux, T.M.; Norby, B.; Burns, T.A.; McCutcheon, L.J.; Belknap, J.K.; Geor, R.J. Equine insulin receptor and insulin-like growth factor-1 receptor expression in digital lamellar tissue and insulin target tissues. *Equine Vet J* 2016, *48*, 626-632, doi:10.1111/evj.12474.
97. Muniyappa, R.; Montagnani, M.; Koh, K.K.; Quon, M.J. Cardiovascular Actions of Insulin. *Endocrine Reviews* 2007, *28*, 463-491, doi:10.1210/er.2007-0006.
98. Kim, J.A.; Montagnani, M.; Koh, K.K.; Quon, M.J. Reciprocal relationships between insulin resistance and endothelial dysfunction: molecular and pathophysiological mechanisms. *Circulation* 2006, *113*, 1888-1904, doi:10.1161/circulationaha.105.563213.
99. Cusi, K.; Maezono, K.; Osman, A.; Pendergrass, M.; Patti, M.E.; Pratipanawatr, T.; DeFronzo, R.A.; Kahn, C.R.; Mandarino, L.J. Insulin resistance differentially affects the PI 3-kinase- and MAP kinase-mediated signaling in human muscle. *J Clin Invest* 2000, *105*, 311-320, doi:10.1172/jci7535.
100. Potenza, M.A.; Marasciulo, F.L.; Chieppa, D.M.; Brigiani, G.S.; Formoso, G.; Quon, M.J.; Montagnani, M. Insulin resistance in spontaneously hypertensive rats is associated with endothelial dysfunction characterized by imbalance between NO and ET-1 production. *Am J Physiol Heart Circ Physiol* 2005, *289*, H813-H822, doi:10.1152/ajpheart.00092.2005.
101. Gauff, F.; Patan-Zugaj, B.; Licka, T.F. Hyperinsulinaemia increases vascular resistance and endothelin-1 expression in the equine digit. *Equine Vet J* 2013, *45*, 613-618, doi:10.1111/evj.12040.
102. Venugopal, C.S.; Eades, S.; Holmes, E.P.; Beadle, R.E. Insulin resistance in equine digital vessel rings: an in vitro model to study vascular dysfunction in equine laminitis. *Equine Vet J* 2011, *43*, 744-749, doi:10.1111/j.2042-3306.2010.00351.x.
103. Wooldridge, A.A.; Waguespack, R.W.; Schwartz, D.D.; Venugopal, C.S.; Eades, S.C.; Beadle, R.E. Vasorelaxation responses to insulin in laminar vessel rings from healthy, lean horses. *Veterinary Journal (London, England : 1997)* 2014, *202*, 83-88, doi:10.1016/j.tvjl.2014.07.021.
104. Morgan, R.A.; Keen, J.A.; Walker, B.R.; Hadoke, P.W. Vascular Dysfunction in Horses with Endocrinopathic Laminitis. *PloS One* 2016, *11*, e0163815, doi:10.1371/journal.pone.0163815.
105. Heliczer, N.; Gerber, V.; Bruckmaier, R.; van der Kolk, J.H.; de Solis, C.N. Cardiovascular findings in ponies with equine metabolic syndrome. *J Am Vet Med Assoc* 2017, *250*, 1027-1035, doi:10.2460/javma.250.9.1027.

106. Nostell, K.; Lindase, S. The effect of insulin infusion on heart rate and systemic blood pressure in horses with equine metabolic syndrome. *Equine Vet J* 2019, 10.1111/evj.13110, doi:10.1111/evj.13110.
107. Nostell, K.E.; Lindase, S.S.; Brojer, J.T. Blood pressure in Warmblood horses before and during a euglycemic-hyperinsulinemic clamp. *Acta Vet Scand* 2016, *58*, 65, doi:10.1186/s13028-016-0247-y.
108. Elzinga, S.; Reedy, S.; Barker, V.D.; Chambers, T.M.; Adams, A.A. Humoral and cell-mediated immune responses to influenza vaccination in equine metabolic syndrome (EMS) horses. *Vet Immunol Immunopathol* 2018, *199*, 32-38, doi:10.1016/j.vetimm.2018.03.009.
109. Sessions, D.R.; Reedy, S.E.; Vick, M.M.; Murphy, B.A.; Fitzgerald, B.P. Development of a model for inducing transient insulin resistance in the mare: preliminary implications regarding the estrous cycle. *J Anim Sci* 2004, *82*, 2321-2328, doi:10.2527/2004.8282321x.
110. Vick, M.M.; Sessions, D.R.; Murphy, B.A.; Kennedy, E.L.; Reedy, S.E.; Fitzgerald, B.P. Obesity is associated with altered metabolic and reproductive activity in the mare: effects of metformin on insulin sensitivity and reproductive cyclicity. *Reprod Fertil Dev* 2006, *18*, 609-617.
111. Sessions-Bresnahan, D.R.; Carnevale, E.M. The effect of equine metabolic syndrome on the ovarian follicular environment. *J Anim Sci* 2014, *92*, 1485-1494, doi:10.2527/jas.2013-7275.
112. Van Weyenberg, S.; Hesta, M.; Buyse, J.; Janssens, G.P. The effect of weight loss by energy restriction on metabolic profile and glucose tolerance in ponies. *J Anim Physiol Anim Nutr (Berl)* 2008, *92*, 538-545, doi:10.1111/j.1439-0396.2007.00744.x.
113. Argo, C.M.; Curtis, G.C.; Grove-White, D.; Dugdale, A.H.; Barfoot, C.F.; Harris, P.A. Weight loss resistance: a further consideration for the nutritional management of obese Equidae. *Veterinary Journal (London, England : 1997)* 2012, *194*, 179-188, doi:10.1016/j.tvjl.2012.09.020.
114. Borgia, L.; Valberg, S.; McCue, M.; Watts, K.; Pagan, J. Glycaemic and insulinaemic responses to feeding hay with different non-structural carbohydrate content in control and polysaccharide storage myopathy-affected horses. *J Anim Physiol Anim Nutr (Berl)* 2011, *95*, 798-807, doi:10.1111/j.1439-0396.2010.01116.x.
115. Geor, R.J. Nutrition and Exercise in the Management of Horses and Ponies at High Risk for Laminitis. *J Equine Vet Sci* 2010, *30*, 463-470, doi:https://doi.org/10.1016/j.jevs.2010.07.011.
116. Longland, A.C.; Barfoot, C.; Harris, P.A. Effects of soaking on the water-soluble carbohydrate and crude protein content of hay. *Vet Rec* 2011, *168*, 618, doi:10.1136/vr.d157.
117. Mack, S.J.; Dugdale, A.H.; Argo, C.M.; Morgan, R.A.; McGowan, C.M. Impact of water-soaking on the nutrient composition of UK hays. *Vet Rec* 2014, *174*, 452, doi:10.1136/vr.102074.
118. Argo, C.M.; Dugdale, A.H.; McGowan, C.M. Considerations for the use of restricted, soaked grass hay diets to promote weight loss in the management of

equine metabolic syndrome and obesity. *Veterinary Journal (London, England : 1997)* 2015, *206*, 170-177, doi:10.1016/j.tvjl.2015.07.027.
119. Martinson, K.L.; Hathaway, M.; Jung, H.; Sheaffer, C. The Effect of Soaking on Protein and Mineral Loss in Orchardgrass and Alfalfa Hay. *J Equine Vet Sci* 2012, *32*, 776-782, doi:https://doi.org/10.1016/j.jevs.2012.03.007.
120. Longland, A.C.; Byrd, B.M. Pasture nonstructural carbohydrates and equine laminitis. *J Nutr* 2006, *136*, 2099s-2102s, doi:10.1093/jn/136.7.2099S.
121. Bailey, S.R.; Menzies-Gow, N.J.; Harris, P.A.; Habershon-Butcher, J.L.; Crawford, C.; Berhane, Y.; Boston, R.C.; Elliott, J. Effect of dietary fructans and dexamethasone administration on the insulin response of ponies predisposed to laminitis. *J Am Vet Med Assoc*2007, *231*, 1365-1373, doi:10.2460/javma.231.9.1365.
122. Fitzgerald, D.M.; Walsh, D.M.; Sillence, M.N.; Pollitt, C.C.; de Laat, M.A. Insulin and incretin responses to grazing in insulin-dysregulated and healthy ponies. *J Vet Intern Med* 2019, *33*, 225-232, doi:10.1111/jvim.15363.
123. Marlow, C.H.; van Tonder, E.M.; Hayward, F.C.; van der Merwe, S.S.; Price, L.E. A report on the consumption, composition and nutritional adequacy of a mixture of lush green perennial ryegrass (Lolium perenne) and cocksfoot (Dactylis glomerata) fed ad libitum to Thoroughbred mares. *J S Afr Vet Assoc* 1983, *54*, 155-157.
124. McMeniman, N.P. Nutrition of grazing broodmares, their foals and young horses : a report for the Rural Industries Research and Development Corporation / by N.P. McMeniman; Rural Industries Research & Development Corporation: Barton, A.C.T, 2000.
125. Ince, J.; Longland, A.C.; Newbold, J.; Harris, P.A. Changes in proportions of dry matter intakes by ponies with access to pasture and haylage for 3 and 20 hours per day respectively, for six weeks. *J Equine Vet Sci* 2011, *31*, 283.
126. Powell, D.M.; Reedy, S.E.; Sessions, D.R.; Fitzgerald, B.P. Effect of short-term exercise training on insulin sensitivity in obese and lean mares. *Equine Vet J Suppl* 2002, 10.1111/j.2042-3306.2002.tb05396.x, 81-84, doi:10.1111/j.2042-3306.2002.tb05396.x.
127. Stewart-Hunt, L.; Geor, R.J.; McCutcheon, L.J. Effects of short-term training on insulin sensitivity and skeletal muscle glucose metabolism in standardbred horses. *Equine Vet J Suppl* 2006, 10.1111/j.2042-3306.2006.tb05544.x, 226-232, doi:10.1111/j.2042-3306.2006.tb05544.x.
128. Gordon, M.E.; McKeever, K.H.; Betros, C.L.; Manso Filho, H.C. Exercise-induced alterations in plasma concentrations of ghrelin, adiponectin, leptin, glucose, insulin, and cortisol in horses. *Veterinary Journal (London, England : 1997)* 2007, *173*, 532-540, doi:10.1016/j.tvjl.2006.01.003.
129. Freestone, J.F.; Beadle, R.; Shoemaker, K.; Bessin, R.T.; Wolfsheimer, K.J.; Church, C. Improved insulin sensitivity in hyperinsulinaemic ponies through physical conditioning and controlled feed intake. *Equine Vet J* 1992, *24*, 187-190.
130. de Laat, M.A.; Hampson, B.A.; Sillence, M.N.; Pollitt, C.C. Sustained, Low-Intensity Exercise Achieved by a Dynamic Feeding System Decreases Body Fat in Ponies. *J Vet Intern Med* 2016, *30*, 1732-1738, doi:10.1111/jvim.14577.
131. Carter, R.A.; McCutcheon, L.J.; Valle, E.; Meilahn, E.N.; Geor, R.J. Effects of exercise training on adiposity, insulin sensitivity, and plasma hormone and lipid

concentrations in overweight or obese, insulin-resistant horses. *Am J Vet Res*2010, *71*, 314-321, doi:10.2460/ajvr.71.3.314.
132. Bamford, N.J.; Potter, S.J.; Baskerville, C.L.; Harris, P.A.; Bailey, S.R. Influence of dietary restriction and low-intensity exercise on weight loss and insulin sensitivity in obese equids. *J Vet Intern Med* 2019, *33*, 280-286, doi:10.1111/jvim.15374.
133. Menzies-Gow, N.J.; Wray, H.; Bailey, S.R.; Harris, P.A.; Elliott, J. The effect of exercise on plasma concentrations of inflammatory markers in normal and previously laminitic ponies. *Equine Vet J* 2014, *46*, 317-321, doi:10.1111/evj.12132.
134. Morgan, R.A.; Keen, J.A.; McGowan, C.M. Treatment of equine metabolic syndrome: A clinical case series. *Equine Vet J* 2016, *48*, 422-426, doi:10.1111/evj.12445.
135. Sakar, Y.; Meddah, B.; Faouzi, M.A.; Cherrah, Y.; Bado, A.; Ducroc, R. Metformin-induced regulation of the intestinal D-glucose transporters. *J Physiol Pharmacol* 2010, *61*, 301-307.
136. Hustace, J.L.; Firshman, A.M.; Mata, J.E. Pharmacokinetics and bioavailability of metformin in horses. *Am J Vet Res*2009, *70*, 665-668, doi:10.2460/ajvr.70.5.665.
137. Tinworth, K.D.; Edwards, S.; Noble, G.K.; Harris, P.A.; Sillence, M.N.; Hackett, L.P. Pharmacokinetics of metformin after enteral administration in insulin-resistant ponies. *Am J Vet Res*2010, *71*, 1201-1206, doi:10.2460/ajvr.71.10.1201.
138. Tinworth, K.D.; Boston, R.C.; Harris, P.A.; Sillence, M.N.; Raidal, S.L.; Noble, G.K. The effect of oral metformin on insulin sensitivity in insulin-resistant ponies. *Veterinary Journal (London, England : 1997)* 2012, *191*, 79-84, doi:10.1016/j.tvjl.2011.01.015.
139. Durham, A.E.; Rendle, D.I.; Newton, J.R. The effect of metformin on measurements of insulin sensitivity and β cell response in 18 horses and ponies with insulin resistance. *Equine Vet J* 2008, *40*, 493-500, doi:10.2746/042516408x273648.
140. Rendle, D.I.; Rutledge, F.; Hughes, K.J.; Heller, J.; Durham, A.E. Effects of metformin hydrochloride on blood glucose and insulin responses to oral dextrose in horses. *Equine Vet J* 2013, *45*, 751-754, doi:10.1111/evj.12068.
141. Sommardahl, C.S.; Frank, N.; Elliott, S.B.; Webb, L.L.; Refsal, K.R.; Denhart, J.W.; Thompson, D.L., Jr. Effects of oral administration of levothyroxine sodium on serum concentrations of thyroid gland hormones and responses to injections of thyrotropin-releasing hormone in healthy adult mares. *Am J Vet Res*2005, *66*, 1025-1031.
142. Frank, N.; Buchanan, B.R.; Elliott, S.B. Effects of long-term oral administration of levothyroxine sodium on serum thyroid hormone concentrations, clinicopathologic variables, and echocardiographic measurements in healthy adult horses. *Am J Vet Res*2008, *69*, 68-75, doi:10.2460/ajvr.69.1.68.
143. Bertin, F.R.; Eichstadt Forsythe, L.; Kritchevsky, J.E. Effects of high doses of levothyroxine sodium on serum concentrations of triiodothyronine and thyroxine in horses. *Am J Vet Res*2019, *80*, 565-571, doi:10.2460/ajvr.80.6.565.
144. Frank, N.; Sommardahl, C.S.; Eiler, H.; Webb, L.L.; Denhart, J.W.; Boston, R.C. Effects of oral administration of levothyroxine sodium on concentrations of plasma

lipids, concentration and composition of very-low-density lipoproteins, and glucose dynamics in healthy adult mares. *Am J Vet Res*2005, *66*, 1032-1038.
145. Frank, N.; Elliott, S.B.; Boston, R.C. Effects of long-term oral administration of levothyroxine sodium on glucose dynamics in healthy adult horses. *Am J Vet Res* 2008, *69*, 76-81, doi:10.2460/ajvr.69.1.76.
146. Winter, J.C.; Sponder, G.; Merle, R.; Aschenbach, J.R.; Gehlen, H. Intracellular free magnesium concentration in healthy horses. *J Anim Physiol Anim Nutr (Berl)* 2018, *102*, 1351-1356, doi:10.1111/jpn.12921.
147. McGowan, C.M.; Dugdale, A.H.; Pinchbeck, G.L.; Argo, C.M. Dietary restriction in combination with a nutraceutical supplement for the management of equine metabolic syndrome in horses. *Veterinary Journal (London, England : 1997)* 2013, *196*, 153-159, doi:10.1016/j.tvjl.2012.10.007.
148. Kheder, M.H.; Sillence, M.N.; Bryant, L.M.; de Laat, M.A. The equine glucose-dependent insulinotropic polypeptide receptor: A potential therapeutic target for insulin dysregulation. *J Anim Sci* 2017, *95*, 2509-2516, doi:10.2527/jas.2017.1468.
149. Meier, A.; de Laat, M.; Reiche, D.; Fitzgerald, D.; Sillence, M. The efficacy and safety of velagliflozin over 16 weeks as a treatment for insulin dysregulation in ponies. *BMC Vet Res* 2019, *15*, 65, doi:10.1186/s12917-019-1811-2.
150. Meier, A.; Reiche, D.; de Laat, M. The sodium-glucose co-transporter 2 inhibitor velagliflozin reduces hyperinsulinemia and prevents laminitis in insulin-dysregulated ponies. *PLoS One* 2018, *13*, e0203655, doi:10.1371/journal.pone.0203655.
151. Marycz, K.; Michalak, I.; Kornicka, K. Advanced nutritional and stem cells approaches to prevent equine metabolic syndrome. *Res Vet Sci* 2018, *118*, 115-125, doi:10.1016/j.rvsc.2018.01.015.
152. Wilcox, G. Insulin and insulin resistance. *Clin Biochem Rev* 2005, *26*, 19-39.
153. MacDonald, P.E.; Joseph, J.W.; Rorsman, P. Glucose-sensing mechanisms in pancreatic beta-cells. *Philos Trans R Soc Lond B Biol Sci* 2005, *360*, 2211-2225, doi:10.1098/rstb.2005.1762.
154. Thorens, B. GLUT2, glucose sensing and glucose homeostasis. *Diabetologia* 2015, *58*, 221-232, doi:10.1007/s00125-014-3451-1.
155. McCulloch, L.J.; van de Bunt, M.; Braun, M.; Frayn, K.N.; Clark, A.; Gloyn, A.L. GLUT2 (SLC2A2) is not the principal glucose transporter in human pancreatic beta cells: implications for understanding genetic association signals at this locus. *Mol Genet Metab* 2011, *104*, 648-653, doi:10.1016/j.ymgme.2011.08.026.
156. Bratanova-Tochkova, T.K.; Cheng, H.; Daniel, S.; Gunawardana, S.; Liu, Y.J.; Mulvaney-Musa, J.; Schermerhorn, T.; Straub, S.G.; Yajima, H.; Sharp, G.W. Triggering and augmentation mechanisms, granule pools, and biphasic insulin secretion. *Diabetes* 2002, *51 Suppl 1*, S83-90, doi:10.2337/diabetes.51.2007.s83.
157. Duhlmeier, R.; Deegen, E.; Fuhrmann, H.; Widdel, A.; Sallmann, H.P. Glucose-dependent insulinotropic polypeptide (GIP) and the enteroinsular axis in equines (Equus caballus). *Comp Biochem Physiol A Mol Integr Physiol* 2001, *129*, 563-575.
158. Toth, F.; Frank, N.; Elliott, S.B.; Perdue, K.; Geor, R.J.; Boston, R.C. Optimisation of the frequently sampled intravenous glucose tolerance test to reduce urinary glucose spilling in horses. *Equine Vet J* 2009, *41*, 844-851.

159. Pratt, S.E.; Geor, R.J.; McCutcheon, L.J. Repeatability of 2 methods for assessment of insulin sensitivity and glucose dynamics in horses. *J Vet Intern Med*2005, *19*, 883-888, doi:10.1892/0891-6640(2005)19[883:romfao]2.0.co;2.
160. Burks, D.J.; White, M.F. IRS proteins and beta-cell function. *Diabetes* 2001, *50 Suppl 1*, S140-145, doi:10.2337/diabetes.50.2007.s140.
161. Czech, M.P.; Buxton, J.M. Insulin action on the internalization of the GLUT4 glucose transporter in isolated rat adipocytes. *J Biol Chem* 1993, *268*, 9187-9190.
162. Mueckler, M. Insulin resistance and the disruption of Glut4 trafficking in skeletal muscle. *J Clin Invest* 2001, *107*, 1211-1213, doi:10.1172/JCI13020.
163. Ruud, J.; Steculorum, S.M.; Brüning, J.C. Neuronal control of peripheral insulin sensitivity and glucose metabolism. *Nat Commun* 2017, *8*, 15259, doi:10.1038/ncomms15259.
164. Waller, A.P.; Burns, T.A.; Mudge, M.C.; Belknap, J.K.; Lacombe, V.A. Insulin resistance selectively alters cell-surface glucose transporters but not their total protein expression in equine skeletal muscle. *J Vet Intern Med* 2011, *25*, 315-321, doi:10.1111/j.1939-1676.2010.0674.x.
165. Suagee, J.K.; Corl, B.A.; Hulver, M.W.; McCutcheon, L.J.; Geor, R.J. Effects of hyperinsulinemia on glucose and lipid transporter expression in insulin-sensitive horses. *Domest Anim Endocrinol* 2011, *40*, 173-181, doi:https://doi.org/10.1016/j.domaniend.2010.11.002.
166. Cavaghan, M.K.; Ehrmann, D.A.; Polonsky, K.S. Interactions between insulin resistance and insulin secretion in the development of glucose intolerance. *J Clin Invest* 2000, *106*, 329-333, doi:10.1172/JCI10761.
167. de Graaf-Roelfsema, E. Glucose homeostasis and the enteroinsular axis in the horse: a possible role in equine metabolic syndrome. *Veterinary Journal (London, England : 1997)* 2014, *199*, 11-18, doi:10.1016/j.tvjl.2013.09.064.
168. Polonsky, K.S.; Given, B.D.; Hirsch, L.; Shapiro, E.T.; Tillil, H.; Beebe, C.; Galloway, J.A.; Frank, B.H.; Karrison, T.; Van Cauter, E. Quantitative study of insulin secretion and clearance in normal and obese subjects. *J Clin Invest* 1988, *81*, 435-441, doi:10.1172/JCI113338.
169. Toth, F.; Frank, N.; Martin-Jimenez, T.; Elliott, S.B.; Geor, R.J.; Boston, R.C. Measurement of C-peptide concentrations and responses to somatostatin, glucose infusion, and insulin resistance in horses. *Equine Vet J* 2010, *42*, 149-155, doi:10.2746/042516409x478497.
170. Faber, O.K.; Christensen, K.; Kehlet, H.; Madsbad, S.; Binder, C. Decreased insulin removal contributes to hyperinsulinemia in obesity. *J Clin Endocrinol Metab* 1981, *53*, 618-621, doi:10.1210/jcem-53-3-618.
171. Kahn, S.E. The importance of the beta-cell in the pathogenesis of type 2 diabetes mellitus. *Am J Med* 2000, *108 Suppl 6a*, 2s-8s.
172. Ward, W.K.; LaCava, E.C.; Paquette, T.L.; Beard, J.C.; Wallum, B.J.; Porte, D., Jr. Disproportionate elevation of immunoreactive proinsulin in type 2 (non-insulin-dependent) diabetes mellitus and in experimental insulin resistance. *Diabetologia* 1987, *30*, 698-702.
173. Cerasi, E.; Kaiser, N.; Leibowitz, G. [Type 2 diabetes and beta cell apoptosis]. *Diabetes Metab* 2000, *26 Suppl 3*, 13-16.

174. Cerf, M.E. High fat programming of beta cell compensation, exhaustion, death and dysfunction. *Pediatr Diabetes* 2015, *16*, 71-78, doi:10.1111/pedi.12137.
175. Ruoff, W.W.; Baker, D.C.; Morgan, S.J.; Abbitt, B. Type II diabetes mellitus in a horse. *Equine Vet J* 1986, *18*, 143-144.
176. Johnson, P.J.; Scotty, N.C.; Wiedmeyer, C.; Messer, N.T.; Kreeger, J.M. Diabetes mellitus in a domesticated Spanish mustang. *J Am Vet Med Assoc* 2005, *226*, 584-588, 542.
177. Durham, A.E.; Hughes, K.J.; Cottle, H.J.; Rendle, D.I.; Boston, R.C. Type 2 diabetes mellitus with pancreatic beta cell dysfunction in 3 horses confirmed with minimal model analysis. *Equine Vet J* 2009, *41*, 924-929.
178. Kazafeos, K. Incretin effect: GLP-1, GIP, DPP4. *Diabetes Res Clin Pract* 2011, *93 Suppl 1*, S32-36, doi:10.1016/s0168-8227(11)70011-0.
179. Kim, W.; Egan, J.M. The role of incretins in glucose homeostasis and diabetes treatment. *Pharmacol Rev* 2008, *60*, 470-512, doi:10.1124/pr.108.000604.
180. Laat, M.A.d.; McGree, J.M.; Sillence, M.N. Equine hyperinsulinemia: investigation of the enteroinsular axis during insulin dysregulation. *Am J Physiol Endocrinol Metab* 2016, *310*, E61-E72, doi:10.1152/ajpendo.00362.2015.
181. Drucker, D.J.; Philippe, J.; Mojsov, S.; Chick, W.L.; Habener, J.F. Glucagon-like peptide I stimulates insulin gene expression and increases cyclic AMP levels in a rat islet cell line. *Proc Natl Acad Sci USA* 1987, *84*, 3434-3438, doi:10.1073/pnas.84.10.3434.
182. Terry, N.A.; Walp, E.R.; Lee, R.A.; Kaestner, K.H.; May, C.L. Impaired enteroendocrine development in intestinal-specific Islet1 mouse mutants causes impaired glucose homeostasis. *Am J Physiol Gastrointest Liver Physiol* 2014, *307*, G979-991, doi:10.1152/ajpgi.00390.2013.
183. Muscelli, E.; Mari, A.; Casolaro, A.; Camastra, S.; Seghieri, G.; Gastaldelli, A.; Holst, J.J.; Ferrannini, E. Separate impact of obesity and glucose tolerance on the incretin effect in normal subjects and type 2 diabetic patients. *Diabetes* 2008, *57*, 1340-1348, doi:10.2337/db07-1315.
184. Toft-Nielsen, M.B.; Damholt, M.B.; Madsbad, S.; Hilsted, L.M.; Hughes, T.E.; Michelsen, B.K.; Holst, J.J. Determinants of the impaired secretion of glucagon-like peptide-1 in type 2 diabetic patients. *J Clin Endocrinol Metab* 2001, *86*, 3717-3723, doi:10.1210/jcem.86.8.7750.
185. Faerch, K.; Torekov, S.S.; Vistisen, D.; Johansen, N.B.; Witte, D.R.; Jonsson, A.; Pedersen, O.; Hansen, T.; Lauritzen, T.; Sandbaek, A., et al. GLP-1 Response to Oral Glucose Is Reduced in Prediabetes, Screen-Detected Type 2 Diabetes, and Obesity and Influenced by Sex: The ADDITION-PRO Study. *Diabetes* 2015, *64*, 2513-2525, doi:10.2337/db14-1751.
186. Ranganath, L.R.; Beety, J.M.; Morgan, L.M.; Wright, J.W.; Howland, R.; Marks, V. Attenuated GLP-1 secretion in obesity: cause or consequence? *Gut* 1996, *38*, 916-919, doi:10.1136/gut.38.6.916.
187. Verdich, C.; Toubro, S.; Buemann, B.; Lysgard Madsen, J.; Juul Holst, J.; Astrup, A. The role of postprandial releases of insulin and incretin hormones in meal-induced satiety--effect of obesity and weight reduction. *Int J Obes Relat Metab Disord* 2001, *25*, 1206-1214, doi:10.1038/sj.ijo.0801655.

188. Frank, N.; Walsh, D.M. Repeatability of Oral Sugar Test Results, Glucagon-Like Peptide-1 Measurements, and Serum High-Molecular-Weight Adiponectin Concentrations in Horses. *J Vet Intern Med*2 017, *31*, 1178-1187, doi:10.1111/jvim.14725.
189. Yabe, D.; Kuroe, A.; Watanabe, K.; Iwasaki, M.; Hamasaki, A.; Hamamoto, Y.; Harada, N.; Yamane, S.; Lee, S.; Murotani, K., et al. Early phase glucagon and insulin secretory abnormalities, but not incretin secretion, are similarly responsible for hyperglycemia after ingestion of nutrients. *J Diabetes Complications* 2015, *29*, 413-421, doi:10.1016/j.jdiacomp.2014.12.010.
190. Knop, F.K.; Aaboe, K.; Vilsboll, T.; Volund, A.; Holst, J.J.; Krarup, T.; Madsbad, S. Impaired incretin effect and fasting hyperglucagonaemia characterizing type 2 diabetic subjects are early signs of dysmetabolism in obesity. *Diabetes Obes Metab* 2012, *14*, 500-510, doi:10.1111/j.1463-1326.2011.01549.x.
191. Smushkin, G.; Sathananthan, A.; Man, C.D.; Zinsmeister, A.R.; Camilleri, M.; Cobelli, C.; Rizza, R.A.; Vella, A. Defects in GLP-1 response to an oral challenge do not play a significant role in the pathogenesis of prediabetes. *J Clin Endocrinol Metab* 2012, *97*, 589-598, doi:10.1210/jc.2011-2561.
192. Nauck, M.A.; Vardarli, I.; Deacon, C.F.; Holst, J.J.; Meier, J.J. Secretion of glucagon-like peptide-1 (GLP-1) in type 2 diabetes: what is up, what is down? *Diabetologia* 2011, *54*, 10-18, doi:10.1007/s00125-010-1896-4.
193. Bamford, N.J.; Baskerville, C.L.; Harris, P.A.; Bailey, S.R. Postprandial glucose, insulin, and glucagon-like peptide-1 responses of different equine breeds adapted to meals containing micronized maize. *J Anim Sci* 2015, *93*, 3377-3383, doi:10.2527/jas.2014-8736.
194. Chameroy, K.A.; Frank, N.; Elliott, S.B.; Boston, R.C. Comparison of Plasma Active Glucagon-Like Peptide 1 Concentrations in Normal Horses and Those With Equine Metabolic Syndrome and in Horses Placed on a High-Grain Diet. *J Equine Vet Sci* 2016, *40*, 16-25, doi:https://doi.org/10.1016/j.jevs.2016.01.009.
195. Meier, J.J.; Gallwitz, B.; Siepmann, N.; Holst, J.J.; Deacon, C.F.; Schmidt, W.E.; Nauck, M.A. Gastric inhibitory polypeptide (GIP) dose-dependently stimulates glucagon secretion in healthy human subjects at euglycaemia. *Diabetologia* 2003, *46*, 798-801, doi:10.1007/s00125-003-1103-y.
196. Gault, V.A.; Porter, D.W.; Irwin, N.; Flatt, P.R. Comparison of sub-chronic metabolic effects of stable forms of naturally occurring GIP(1-30) and GIP(1-42) in high-fat fed mice. *J Endocrinol* 2011, *208*, 265-271, doi:10.1530/joe-10-0419.
197. Skrha, J.; Hilgertova, J.; Jarolimkova, M.; Kunesova, M.; Hill, M. Meal test for glucose-dependent insulinotropic peptide (GIP) in obese and type 2 diabetic patients. *Physiol Res* 2010, *59*, 749-755.
198. Vilsboll, T.; Krarup, T.; Madsbad, S.; Holst, J.J. Defective amplification of the late phase insulin response to glucose by GIP in obese Type II diabetic patients. *Diabetologia* 2002, *45*, 1111-1119, doi:10.1007/s00125-002-0878-6.
199. Nauck, M.A.; Meier, J.J. The incretin effect in healthy individuals and those with type 2 diabetes: physiology, pathophysiology, and response to therapeutic interventions. *Lancet Diabetes Endocrinol* 2016, *4*, 525-536, doi:10.1016/s2213-8587(15)00482-9.

200. Kreymann, B.; Williams, G.; Ghatei, M.A.; Bloom, S.R. Glucagon-like peptide-1 7-36: a physiological incretin in man. *Lancet* 1987, *2*, 1300-1304.
201. Christensen, M.; Calanna, S.; Sparre-Ulrich, A.H.; Kristensen, P.L.; Rosenkilde, M.M.; Faber, J.; Purrello, F.; van Hall, G.; Holst, J.J.; Vilsboll, T., et al. Glucose-dependent insulinotropic polypeptide augments glucagon responses to hypoglycemia in type 1 diabetes. *Diabetes* 2015, *64*, 72-78, doi:10.2337/db14-0440.
202. Chia, C.W.; Carlson, O.D.; Kim, W.; Shin, Y.K.; Charles, C.P.; Kim, H.S.; Melvin, D.L.; Egan, J.M. Exogenous glucose-dependent insulinotropic polypeptide worsens post prandial hyperglycemia in type 2 diabetes. *Diabetes* 2009, *58*, 1342-1349, doi:10.2337/db08-0958.
203. Gniuli, D.; Calcagno, A.; Dalla Libera, L.; Calvani, R.; Leccesi, L.; Caristo, M.E.; Vettor, R.; Castagneto, M.; Ghirlanda, G.; Mingrone, G. High-fat feeding stimulates endocrine, glucose-dependent insulinotropic polypeptide (GIP)-expressing cell hyperplasia in the duodenum of Wistar rats. *Diabetologia* 2010, *53*, 2233-2240, doi:10.1007/s00125-010-1830-9.
204. Vilsboll, T.; Agerso, H.; Lauritsen, T.; Deacon, C.F.; Aaboe, K.; Madsbad, S.; Krarup, T.; Holst, J.J. The elimination rates of intact GIP as well as its primary metabolite, GIP 3-42, are similar in type 2 diabetic patients and healthy subjects. *Regul Pept* 2006, *137*, 168-172, doi:10.1016/j.regpep.2006.07.007.
205. Bagger, J.I.; Knop, F.K.; Lund, A.; Vestergaard, H.; Holst, J.J.; Vilsboll, T. Impaired regulation of the incretin effect in patients with type 2 diabetes. *J Clin Endocrinol Metab* 2011, *96*, 737-745, doi:10.1210/jc.2010-2435.
206. Xu, G.; Kaneto, H.; Laybutt, D.R.; Duvivier-Kali, V.F.; Trivedi, N.; Suzuma, K.; King, G.L.; Weir, G.C.; Bonner-Weir, S. Downregulation of GLP-1 and GIP receptor expression by hyperglycemia: possible contribution to impaired incretin effects in diabetes. *Diabetes* 2007, *56*, 1551-1558, doi:10.2337/db06-1033.
207. Piteau, S.; Olver, A.; Kim, S.J.; Winter, K.; Pospisilik, J.A.; Lynn, F.; Manhart, S.; Demuth, H.U.; Speck, M.; Pederson, R.A., et al. Reversal of islet GIP receptor down-regulation and resistance to GIP by reducing hyperglycemia in the Zucker rat. *Biochem Biophys Res Commun* 2007, *362*, 1007-1012, doi:10.1016/j.bbrc.2007.08.115.
208. Rajan, S.; Dickson, L.M.; Mathew, E.; Orr, C.M.; Ellenbroek, J.H.; Philipson, L.H.; Wicksteed, B. Chronic hyperglycemia downregulates GLP-1 receptor signaling in pancreatic beta-cells via protein kinase A. *Mol Metab* 2015, *4*, 265-276, doi:10.1016/j.molmet.2015.01.010.
209. Mulvihill, E.E. Dipeptidyl peptidase inhibitor therapy in type 2 diabetes: Control of the incretin axis and regulation of postprandial glucose and lipid metabolism. *Peptides* 2018, *100*, 158-164, doi:10.1016/j.peptides.2017.11.023.
210. Min, S.H.; Yoon, J.H.; Hahn, S.; Cho, Y.M. Comparison between SGLT2 inhibitors and DPP4 inhibitors added to insulin therapy in type 2 diabetes: a systematic review with indirect comparison meta-analysis. *Diabetes Metab Res Rev* 2017, *33*, doi:10.1002/dmrr.2818.
211. Goldberg, I.J.; Eckel, R.H.; Abumrad, N.A. Regulation of fatty acid uptake into tissues: lipoprotein lipase- and CD36-mediated pathways. *J Lipid Res* 2009, *50 Suppl*, S86-90, doi:10.1194/jlr.R800085-JLR200.

212. Feingold, K.R.; Grunfeld, C. Introduction to Lipids and Lipoproteins. In *Endotext*, Feingold, K.R., Anawalt, B., Boyce, A., Chrousos, G., Dungan, K., Grossman, A., Hershman, J.M., Kaltsas, G., Koch, C., Kopp, P., et al., Eds. MDText.com, Inc.: South Dartmouth (MA), 2000.
213. Le Goff, D.; Pastier, D.; Hannan, Y.; Petit, E.; Ayrault-Jarrier, M.; Nouvelot, A. Lipid and apolipoprotein distribution as a function of density in equine plasma lipoprotein. *Comp Biochem Physiol B* 1989, *93*, 371-377.
214. Watson, T.D.; Burns, L.; Love, S.; Packard, C.J.; Shepherd, J. The isolation, characterisation and quantification of the equine plasma lipoproteins. *Equine Vet J* 1991, *23*, 353-359.
215. Coleman, M.C.; Walzem, R.L.; Kieffer, A.J.; Minamoto, T.; Suchodolski, J.; Cohen, N.D. Novel lipoprotein density profiling in laminitic, obese, and healthy horses. *Domest Anim Endocrinol* 2019, *68*, 92-99, doi:https://doi.org/10.1016/j.domaniend.2018.11.003.
216. Asadi, F.; Mohri, M.; Adibmoradi, M.; Pourkabir, M. Serum lipid and lipoprotein parameters of Turkman horses. *Vet Clin Pathol* 2006, *35*, 332-334.
217. Kersten, S. Mechanisms of nutritional and hormonal regulation of lipogenesis. *EMBO Rep* 2001, *2*, 282-286, doi:10.1093/embo-reports/kve071.
218. Vaughan, M.; Berger, J.E.; Steinberg, D. Hormone-Sensitive Lipase and Monoglyceride Lipase Activities in Adipose Tissue. *J Biol Chem* 1964, *239*, 401-409.
219. Nakae, J.; Accili, D. The mechanism of insulin action. *J Pediatr Endocrinol Metab* 1999, *12 Suppl 3*, 721-731.
220. Wang, Y.; Botolin, D.; Xu, J.; Christian, B.; Mitchell, E.; Jayaprakasam, B.; Nair, M.G.; Peters, J.M.; Busik, J.V.; Olson, L.K., et al. Regulation of hepatic fatty acid elongase and desaturase expression in diabetes and obesity. *J Lipid Res* 2006, *47*, 2028-2041, doi:10.1194/jlr.M600177-JLR200.
221. Postic, C.; Girard, J. Contribution of de novo fatty acid synthesis to hepatic steatosis and insulin resistance: lessons from genetically engineered mice. *J Clin Invest* 2008, *118*, 829-838, doi:10.1172/jci34275.
222. Wang, H.; Eckel, R.H. Lipoprotein lipase: from gene to obesity. *Am J Physiol Endocrinol Metab* 2009, *297*, E271-288, doi:10.1152/ajpendo.90920.2008.
223. Olivecrona, G. Role of lipoprotein lipase in lipid metabolism. *Curr Opin Lipidol* 2016, *27*, 233-241, doi:10.1097/mol.0000000000000297.
224. Eisenberg, S.; Sehayek, E.; Olivecrona, T.; Vlodavsky, I. Lipoprotein lipase enhances binding of lipoproteins to heparan sulfate on cell surfaces and extracellular matrix. *J Clin Invest*1992, *90*, 2013-2021, doi:10.1172/JCI116081.
225. Goldberg, I.J.; Eckel, R.H.; Abumrad, N.A. Regulation of fatty acid uptake into tissues: lipoprotein lipase- and CD36-mediated pathways. *Journal of lipid research* 2009, *50 Suppl*, S86-S90, doi:10.1194/jlr.R800085-JLR200.
226. Wu, Q.; Ortegon, A.M.; Tsang, B.; Doege, H.; Feingold, K.R.; Stahl, A. FATP1 is an insulin-sensitive fatty acid transporter involved in diet-induced obesity. *Mol Cell Biol* 2006, *26*, 3455-3467, doi:10.1128/mcb.26.9.3455-3467.2006.
227. Sparks, J.D.; Sparks, C.E.; Adeli, K. Selective hepatic insulin resistance, VLDL overproduction, and hypertriglyceridemia. *Arterioscler Thromb Vasc Biol* 2012, *32*, 2104-2112, doi:10.1161/atvbaha.111.241463.

228. Sukonina, V.; Lookene, A.; Olivecrona, T.; Olivecrona, G. Angiopoietin-like protein 4 converts lipoprotein lipase to inactive monomers and modulates lipase activity in adipose tissue. *Proc Natl Acad Sci USA* 2006, *103*, 17450-17455, doi:10.1073/pnas.0604026103.
229. Beigneux, A.P.; Davies, B.S.; Gin, P.; Weinstein, M.M.; Farber, E.; Qiao, X.; Peale, F.; Bunting, S.; Walzem, R.L.; Wong, J.S., et al. Glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1 plays a critical role in the lipolytic processing of chylomicrons. *Cell Metab* 2007, *5*, 279-291, doi:10.1016/j.cmet.2007.02.002.
230. Medh, J.D.; Bowen, S.L.; Fry, G.L.; Ruben, S.; Andracki, M.; Inoue, I.; Lalouel, J.M.; Strickland, D.K.; Chappell, D.A. Lipoprotein lipase binds to low density lipoprotein receptors and induces receptor-mediated catabolism of very low density lipoproteins in vitro. *J Biol Chem* 1996, *271*, 17073-17080, doi:10.1074/jbc.271.29.17073.
231. Loeffler, B.; Heeren, J.; Blaeser, M.; Radner, H.; Kayser, D.; Aydin, B.; Merkel, M. Lipoprotein lipase-facilitated uptake of LDL is mediated by the LDL receptor. *J Lipid Res* 2007, *48*, 288-298, doi:10.1194/jlr.M600292-JLR200.
232. Arner, P. Insulin resistance in type 2 diabetes: role of fatty acids. *Diabetes Metab Res Rev* 2002, *18 Suppl 2*, S5-9.
233. Bergman, R.N.; Mittelman, S.D. Central role of the adipocyte in insulin resistance. *J Basic Clin Physiol Pharmacol* 1998, *9*, 205-221.
234. Randle, P.J. Regulatory interactions between lipids and carbohydrates: the glucose fatty acid cycle after 35 years. *Diabetes Metab Rev* 1998, *14*, 263-283.
235. DeFronzo, R.A. Insulin resistance: a multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidaemia and atherosclerosis. *Neth J Med* 1997, *50*, 191-197.
236. Parrizas, M.; Maestro, M.A.; Boj, S.F.; Paniagua, A.; Casamitjana, R.; Gomis, R.; Rivera, F.; Ferrer, J. Hepatic nuclear factor 1-alpha directs nucleosomal hyperacetylation to its tissue-specific transcriptional targets. *Mol Cell Biol* 2001, *21*, 3234-3243, doi:10.1128/mcb.21.9.3234-3243.2001.
237. Petersen, M.C.; Shulman, G.I. Mechanisms of Insulin Action and Insulin Resistance. *Physiol Rev* 2018, *98*, 2133-2223, doi:10.1152/physrev.00063.2017.
238. Rytka, J.M.; Wueest, S.; Schoenle, E.J.; Konrad, D. The portal theory supported by venous drainage-selective fat transplantation. *Diabetes* 2011, *60*, 56-63, doi:10.2337/db10-0697.
239. Soli, A.H.; Kahn, C.R.; Neville, D.M., Jr.; Roth, J. Insulin receptor deficiency in genetic and acquired obesity. *J Clin Invest* 1975, *56*, 769-780, doi:10.1172/jci108155.
240. Freidenberg, G.R.; Reichart, D.; Olefsky, J.M.; Henry, R.R. Reversibility of defective adipocyte insulin receptor kinase activity in non-insulin-dependent diabetes mellitus. Effect of weight loss. *J Clin Invest* 1988, *82*, 1398-1406, doi:10.1172/jci113744.
241. Graham, T.E.; Yang, Q.; Bluher, M.; Hammarstedt, A.; Ciaraldi, T.P.; Henry, R.R.; Wason, C.J.; Oberbach, A.; Jansson, P.A.; Smith, U., et al. Retinol-binding protein 4 and insulin resistance in lean, obese, and diabetic subjects. *N Engl J Med* 2006, *354*, 2552-2563, doi:10.1056/NEJMoa054862.

242. Adams-Huet, B.; Devaraj, S.; Siegel, D.; Jialal, I. Increased adipose tissue insulin resistance in metabolic syndrome: relationship to circulating adipokines. *Metab Syndr Relat Disord* 2014, *12*, 503-507, doi:10.1089/met.2014.0092.
243. Reina, M.; Brunzell, J.D.; Deeb, S.S. Molecular basis of familial chylomicronemia: mutations in the lipoprotein lipase and apolipoprotein C-II genes. *J Lipid Res* 1992, *33*, 1823-1832.
244. Garge, N.R.; Bobashev, G.; Eggleston, B. Random forest methodology for model-based recursive partitioning: the mobForest package for R. *BMC Bioinformatics* 2013, *14*, 125, doi:10.1186/1471-2105-14-125.
245. Wang, H.; Knaub, L.A.; Jensen, D.R.; Young Jung, D.; Hong, E.G.; Ko, H.J.; Coates, A.M.; Goldberg, I.J.; de la Houssaye, B.A.; Janssen, R.C., et al. Skeletal muscle-specific deletion of lipoprotein lipase enhances insulin signaling in skeletal muscle but causes insulin resistance in liver and other tissues. *Diabetes* 2009, *58*, 116-124, doi:10.2337/db07-1839.
246. Kim, J.K.; Fillmore, J.J.; Chen, Y.; Yu, C.; Moore, I.K.; Pypaert, M.; Lutz, E.P.; Kako, Y.; Velez-Carrasco, W.; Goldberg, I.J., et al. Tissue-specific overexpression of lipoprotein lipase causes tissue-specific insulin resistance. *Proc Natl Acad Sci USA* 2001, *98*, 7522-7527, doi:10.1073/pnas.121164498.
247. Vishram, J.K.; Hansen, T.W.; Torp-Pedersen, C.; Madsbad, S.; Jorgensen, T.; Fenger, M.; Lyngbaek, S.; Jeppesen, J. Relationship Between Two Common Lipoprotein Lipase Variants and the Metabolic Syndrome and Its Individual Components. *Metab Syndr Relat Disord* 2016, *14*, 442-448, doi:10.1089/met.2016.0030.
248. Coburn, C.T.; Knapp, F.F., Jr.; Febbraio, M.; Beets, A.L.; Silverstein, R.L.; Abumrad, N.A. Defective uptake and utilization of long chain fatty acids in muscle and adipose tissues of CD36 knockout mice. *J Biol Chem* 2000, *275*, 32523-32529, doi:10.1074/jbc.M003826200.
249. Pravenec, M.; Landa, V.; Zidek, V.; Musilova, A.; Kren, V.; Kazdova, L.; Aitman, T.J.; Glazier, A.M.; Ibrahimi, A.; Abumrad, N.A., et al. Transgenic rescue of defective Cd36 ameliorates insulin resistance in spontaneously hypertensive rats. *Nature Genetics* 2001, *27*, 156-158, doi:10.1038/84777.
250. Bonen, A.; Parolin, M.L.; Steinberg, G.R.; Calles-Escandon, J.; Tandon, N.N.; Glatz, J.F.; Luiken, J.J.; Heigenhauser, G.J.; Dyck, D.J. Triacylglycerol accumulation in human obesity and type 2 diabetes is associated with increased rates of skeletal muscle fatty acid transport and increased sarcolemmal FAT/CD36. *Faseb j* 2004, *18*, 1144-1146, doi:10.1096/fj.03-1065fje.
251. Sethi, J.K.; Vidal-Puig, A.J. Thematic review series: adipocyte biology. Adipose tissue function and plasticity orchestrate nutritional adaptation. *J Lipid Res* 2007, *48*, 1253-1262, doi:10.1194/jlr.R700005-JLR200.
252. Cinti, S.; Mitchell, G.; Barbatelli, G.; Murano, I.; Ceresi, E.; Faloia, E.; Wang, S.; Fortier, M.; Greenberg, A.S.; Obin, M.S. Adipocyte death defines macrophage localization and function in adipose tissue of obese mice and humans. *J Lipid Res* 2005, *46*, 2347-2355, doi:10.1194/jlr.M500294-JLR200.
253. Barchetta, I.; Cimini, F.A.; Ciccarelli, G.; Baroni, M.G. Sick fat: the good and the bad of old and new circulating markers of adipose tissue inflammation. *J*

Endocrinol Invest 2019, 10.1007/s40618-019-01052-3, doi:10.1007/s40618-019-01052-3.
254. Grandl, G.; Wolfrum, C. Hemostasis, endothelial stress, inflammation, and the metabolic syndrome. *Semin Immunopathol* 2018, *40*, 215-224, doi:10.1007/s00281-017-0666-5.
255. Hotamisligil, G.S.; Shargill, N.S.; Spiegelman, B.M. Adipose expression of tumor necrosis factor-alpha: direct role in obesity-linked insulin resistance. *Science (New York, N.Y.)* 1993, *259*, 87-91, doi:10.1126/science.7678183.
256. Borst, S.E. The role of TNF-alpha in insulin resistance. *Endocrine* 2004, *23*, 177-182, doi:10.1385/endo:23:2-3:177.
257. Aguirre, V.; Uchida, T.; Yenush, L.; Davis, R.; White, M.F. The c-Jun NH2-terminal Kinase Promotes Insulin Resistance during Association with Insulin Receptor Substrate-1 and Phosphorylation of Ser307. *J Biol Chem* 2000, *275*, 9047-9054.
258. Eldor, R.; Yeffet, A.; Baum, K.; Doviner, V.; Amar, D.; Ben-Neriah, Y.; Christofori, G.; Peled, A.; Carel, J.C.; Boitard, C., et al. Conditional and specific NF-kappaB blockade protects pancreatic beta cells from diabetogenic agents. *Proc Natl Acad Sci USA* 2006, *103*, 5072-5077, doi:10.1073/pnas.0508166103.
259. Zabolotny, J.M.; Kim, Y.B.; Welsh, L.A.; Kershaw, E.E.; Neel, B.G.; Kahn, B.B. Protein-tyrosine phosphatase 1B expression is induced by inflammation in vivo. *J Biol Chem* 2008, *283*, 14230-14241, doi:10.1074/jbc.M800061200.
260. Zhang, X.; Zhang, G.; Zhang, H.; Karin, M.; Bai, H.; Cai, D. Hypothalamic IKKbeta/NF-kappaB and ER stress link overnutrition to energy imbalance and obesity. *Cell* 2008, *135*, 61-73, doi:10.1016/j.cell.2008.07.043.
261. Tang, T.; Zhang, J.; Yin, J.; Staszkiewicz, J.; Gawronska-Kozak, B.; Jung, D.Y.; Ko, H.J.; Ong, H.; Kim, J.K.; Mynatt, R., et al. Uncoupling of inflammation and insulin resistance by NF-kappaB in transgenic mice through elevated energy expenditure. *J Biol Chem* 2010, *285*, 4637-4644, doi:10.1074/jbc.M109.068007.
262. Kern, P.A.; Ranganathan, S.; Li, C.; Wood, L.; Ranganathan, G. Adipose tissue tumor necrosis factor and interleukin-6 expression in human obesity and insulin resistance. *Am J Physiol Endocrinol Metab* 2001, *280*, E745-751, doi:10.1152/ajpendo.2001.280.5.E745.
263. Nagareddy, P.R.; Kraakman, M.; Masters, S.L.; Stirzaker, R.A.; Gorman, D.J.; Grant, R.W.; Dragoljevic, D.; Hong, E.S.; Abdel-Latif, A.; Smyth, S.S., et al. Adipose tissue macrophages promote myelopoiesis and monocytosis in obesity. *Cell Metab* 2014, *19*, 821-835, doi:10.1016/j.cmet.2014.03.029.
264. Bendtzen, K.; Mandrup-Poulsen, T.; Nerup, J.; Nielsen, J.H.; Dinarello, C.A.; Svenson, M. Cytotoxicity of human pI 7 interleukin-1 for pancreatic islets of Langerhans. *Science (New York, N.Y.)* 1986, *232*, 1545-1547, doi:10.1126/science.3086977.
265. Bristulf, J.; Gatti, S.; Malinowsky, D.; Bjork, L.; Sundgren, A.K.; Bartfai, T. Interleukin-1 stimulates the expression of type I and type II interleukin-1 receptors in the rat insulinoma cell line Rinm5F; sequencing a rat type II interleukin-1 receptor cDNA. *Eur Cytokine Netw* 1994, *5*, 319-330.
266. Kraakman, M.J.; Kammoun, H.L.; Allen, T.L.; Deswaerte, V.; Henstridge, D.C.; Estevez, E.; Matthews, V.B.; Neill, B.; White, D.A.; Murphy, A.J., et al. Blocking

IL-6 trans-signaling prevents high-fat diet-induced adipose tissue macrophage recruitment but does not improve insulin resistance. *Cell Metab* 2015, *21*, 403-416, doi:10.1016/j.cmet.2015.02.006.
267. Kishimoto, T. Interleukin-6: discovery of a pleiotropic cytokine. *Arthritis Res Ther* 2006, *8 Suppl 2*, S2, doi:10.1186/ar1916.
268. Weigert, C.; Hennige, A.M.; Lehmann, R.; Brodbeck, K.; Baumgartner, F.; Schauble, M.; Haring, H.U.; Schleicher, E.D. Direct cross-talk of interleukin-6 and insulin signal transduction via insulin receptor substrate-1 in skeletal muscle cells. *J Biol Chem* 2006, *281*, 7060-7067, doi:10.1074/jbc.M509782200.
269. Wallenius, V.; Wallenius, K.; Ahren, B.; Rudling, M.; Carlsten, H.; Dickson, S.L.; Ohlsson, C.; Jansson, J.O. Interleukin-6-deficient mice develop mature-onset obesity. *Nat Med* 2002, *8*, 75-79, doi:10.1038/nm0102-75.
270. Carey, A.L.; Steinberg, G.R.; Macaulay, S.L.; Thomas, W.G.; Holmes, A.G.; Ramm, G.; Prelovsek, O.; Hohnen-Behrens, C.; Watt, M.J.; James, D.E., et al. Interleukin-6 increases insulin-stimulated glucose disposal in humans and glucose uptake and fatty acid oxidation in vitro via AMP-activated protein kinase. *Diabetes* 2006, *55*, 2688-2697, doi:10.2337/db05-1404.
271. Basinska, K.; Marycz, K.; Sieszek, A.; Nicpon, J. The production and distribution of IL-6 and TNF-a in subcutaneous adipose tissue and their correlation with serum concentrations in Welsh ponies with equine metabolic syndrome. *J Vet Sci* 2015, *16*, 113-120, doi:10.4142/jvs.2015.16.1.113.
272. Reynolds, A.; Keen, J.A. Adipose tissue dysfunction in obese horses with equine metabolic syndrome. *Equine Vet J* 2019, 10.1111/evj.13097, doi:10.1111/evj.13097.
273. Bruynsteen, L.; Erkens, T.; Peelman, L.J.; Ducatelle, R.; Janssens, G.P.; Harris, P.A.; Hesta, M. Expression of inflammation-related genes is associated with adipose tissue location in horses. *BMC Vet Res* 2013, *9*, 240, doi:10.1186/1746-6148-9-240.
274. Burns, T.A.; Geor, R.J.; Mudge, M.C.; McCutcheon, L.J.; Hinchcliff, K.W.; Belknap, J.K. Proinflammatory cytokine and chemokine gene expression profiles in subcutaneous and visceral adipose tissue depots of insulin-resistant and insulin-sensitive light breed horses. *J Vet Intern Med* 2010, *24*, 932-939, doi:10.1111/j.1939-1676.2010.0551.x.
275. Waller, A.P.; Huettner, L.; Kohler, K.; Lacombe, V.A. Novel link between inflammation and impaired glucose transport during equine insulin resistance. *Vet Immunol Immunopathol* 2012, *149*, 208-215, doi:10.1016/j.vetimm.2012.07.003.
276. Holbrook, T.C.; Tipton, T.; McFarlane, D. Neutrophil and cytokine dysregulation in hyperinsulinemic obese horses. *Vet Immunol Immunopathol* 2012, *145*, 283-289, doi:10.1016/j.vetimm.2011.11.013.
277. Treiber, K.; Carter, R.; Gay, L.; Williams, C.; Geor, R. Inflammatory and redox status of ponies with a history of pasture-associated laminitis. *Vet Immunol Immunopathol* 2009, *129*, 216-220, doi:10.1016/j.vetimm.2008.11.004.
278. Vick, M.M.; Adams, A.A.; Murphy, B.A.; Sessions, D.R.; Horohov, D.W.; Cook, R.F.; Shelton, B.J.; Fitzgerald, B.P. Relationships among inflammatory cytokines, obesity, and insulin sensitivity in the horse. *J Anim Sci* 2007, *85*, 1144-1155, doi:10.2527/jas.2006-673.

279. Radin, M.J.; Sharkey, L.C.; Holycross, B.J. Adipokines: a review of biological and analytical principles and an update in dogs, cats, and horses. *Vet Clin Pathol* 2009, *38*, 136-156, doi:10.1111/j.1939-165X.2009.00133.x.
280. Lam, Q.L.; Lu, L. Role of leptin in immunity. *Cell Mol Immunol* 2007, *4*, 1-13.
281. Moore, J.L.; Siciliano, P.D.; Pratt-Phillips, S.E. Effects of Diet Versus Exercise on Morphometric Measurements, Blood Hormone Concentrations, and Oral Sugar Test Response in Obese Horses. *J Equine Vet Sci* 2019, *78*, 38-45, doi:10.1016/j.jevs.2019.03.214.
282. Glunk, E.C.; Hathaway, M.R.; Grev, A.M.; Lamprecht, E.D.; Maher, M.C.; Martinson, K.L. The effect of a limit-fed diet and slow-feed hay nets on morphometric measurements and postprandial metabolite and hormone patterns in adult horses. *J Anim Sci* 2015, *93*, 4144-4152, doi:10.2527/jas.2015-9150.
283. Kearns, C.F.; McKeever, K.H.; Roegner, V.; Brady, S.M.; Malinowski, K. Adiponectin and leptin are related to fat mass in horses. *Veterinary Journal (London, England : 1997)* 2006, *172*, 460-465, doi:10.1016/j.tvjl.2005.05.002.
284. Buff, P.R.; Dodds, A.C.; Morrison, C.D.; Whitley, N.C.; McFadin, E.L.; Daniel, J.A.; Djiane, J.; Keisler, D.H. Leptin in horses: tissue localization and relationship between peripheral concentrations of leptin and body condition. *J Anim Sci* 2002, *80*, 2942-2948, doi:10.2527/2002.80112942x.
285. Martin-Gimenez, T.; de Blas, I.; Aguilera-Tejero, E.; Diez de Castro, E.; Aguirre-Pascasio, C.N. Endocrine, morphometric, and ultrasonographic characterization of neck adiposity in Andalusian horses. *Domest Anim Endocrinol* 2016, *56*, 57-62, doi:10.1016/j.domaniend.2016.02.003.
286. Hosoi, T.; Kawagishi, T.; Okuma, Y.; Tanaka, J.; Nomura, Y. Brain stem is a direct target for leptin's action in the central nervous system. *Endocrinology* 2002, *143*, 3498-3504, doi:10.1210/en.2002-220077.
287. Farr, O.M.; Gavrieli, A.; Mantzoros, C.S. Leptin applications in 2015: what have we learned about leptin and obesity? *Curr Opin Endocrinol Diabetes Obes* 2015, *22*, 353-359, doi:10.1097/med.0000000000000184.
288. D'Souza A, M.; Neumann, U.H.; Glavas, M.M.; Kieffer, T.J. The glucoregulatory actions of leptin. *Mol Metab* 2017, *6*, 1052-1065, doi:10.1016/j.molmet.2017.04.011.
289. Van Weyenberg, S.; Buyse, J.; Kalmar, I.D.; Swennen, Q.; Janssens, G.P. Voluntary feed intake and leptin sensitivity in ad libitum fed obese ponies following a period of restricted feeding: a pilot study. *J Anim Physiol Anim Nutr (Berl)* 2013, *97*, 624-631, doi:10.1111/j.1439-0396.2012.01300.x.
290. Banks, W.A. Role of the blood-brain barrier in the evolution of feeding and cognition. *Ann N Y Acad Sci* 2012, *1264*, 13-19, doi:10.1111/j.1749-6632.2012.06568.x.
291. Schwartz, M.W.; Peskind, E.; Raskind, M.; Boyko, E.J.; Porte, D., Jr. Cerebrospinal fluid leptin levels: relationship to plasma levels and to adiposity in humans. *Nat Med* 1996, *2*, 589-593.
292. Oh, I.S.; Shimizu, H.; Sato, T.; Uehara, Y.; Okada, S.; Mori, M. Molecular mechanisms associated with leptin resistance: n-3 polyunsaturated fatty acids induce alterations in the tight junction of the brain. *Cell Metab* 2005, *1*, 331-341, doi:10.1016/j.cmet.2005.04.004.

293. Hosoi, T.; Sasaki, M.; Miyahara, T.; Hashimoto, C.; Matsuo, S.; Yoshii, M.; Ozawa, K. Endoplasmic reticulum stress induces leptin resistance. *Mol Pharmacol* 2008, *74*, 1610-1619, doi:10.1124/mol.108.050070.
294. Jung, C.H.; Kim, M.S. Molecular mechanisms of central leptin resistance in obesity. *Arch Pharm Res* 2013, *36*, 201-207, doi:10.1007/s12272-013-0020-y.
295. Goncalves, G.H.; Li, W.; Garcia, A.V.; Figueiredo, M.S.; Bjorbaek, C. Hypothalamic agouti-related peptide neurons and the central melanocortin system are crucial mediators of leptin's antidiabetic actions. *Cell Rep* 2014, *7*, 1093-1103, doi:10.1016/j.celrep.2014.04.010.
296. Kieffer, T.J.; Habener, J.F. The adipoinsular axis: effects of leptin on pancreatic beta-cells. *Am J Physiol Endocrinol Metab* 2000, *278*, E1-e14, doi:10.1152/ajpendo.2000.278.1.E1.
297. Liang, Y.; Osborne, M.C.; Monia, B.P.; Bhanot, S.; Gaarde, W.A.; Reed, C.; She, P.; Jetton, T.L.; Demarest, K.T. Reduction in glucagon receptor expression by an antisense oligonucleotide ameliorates diabetic syndrome in db/db mice. *Diabetes* 2004, *53*, 410-417, doi:10.2337/diabetes.53.2.410.
298. Toda, C.; Shiuchi, T.; Lee, S.; Yamato-Esaki, M.; Fujino, Y.; Suzuki, A.; Okamoto, S.; Minokoshi, Y. Distinct effects of leptin and a melanocortin receptor agonist injected into medial hypothalamic nuclei on glucose uptake in peripheral tissues. *Diabetes* 2009, *58*, 2757-2765, doi:10.2337/db09-0638.
299. Santos-Alvarez, J.; Goberna, R.; Sanchez-Margalet, V. Human leptin stimulates proliferation and activation of human circulating monocytes. *Cell Immunol* 1999, *194*, 6-11, doi:10.1006/cimm.1999.1490.
300. Shen, J.; Sakaida, I.; Uchida, K.; Terai, S.; Okita, K. Leptin enhances TNF-alpha production via p38 and JNK MAPK in LPS-stimulated Kupffer cells. *Life Sci* 2005, *77*, 1502-1515, doi:10.1016/j.lfs.2005.04.004.
301. Agrawal, S.; Gollapudi, S.; Su, H.; Gupta, S. Leptin activates human B cells to secrete TNF-alpha, IL-6, and IL-10 via JAK2/STAT3 and p38MAPK/ERK1/2 signaling pathway. *J Clin Immunol* 2011, *31*, 472-478, doi:10.1007/s10875-010-9507-1.
302. van Dielen, F.M.; van't Veer, C.; Schols, A.M.; Soeters, P.B.; Buurman, W.A.; Greve, J.W. Increased leptin concentrations correlate with increased concentrations of inflammatory markers in morbidly obese individuals. *Int J Obes Relat Metab Disord* 2001, *25*, 1759-1766, doi:10.1038/sj.ijo.0801825.
303. Belhayara, M.I.; Mellouk, Z.; Hamdaoui, M.S.; Bachaoui, M.; Kheroua, O.; Malaisse, W.J. Relationship between the insulin resistance and circulating predictive biochemical markers in metabolic syndrome among young adults in western Algeria. *Diabetes Metab Syndr* 2019, *13*, 504-509, doi:10.1016/j.dsx.2018.11.019.
304. Adela, R.; Reddy, P.N.C.; Ghosh, T.S.; Aggarwal, S.; Yadav, A.K.; Das, B.; Banerjee, S.K. Serum protein signature of coronary artery disease in type 2 diabetes mellitus. *J Transl Med* 2019, *17*, 17, doi:10.1186/s12967-018-1755-5.
305. Adejumo, E.N.; Adejumo, O.A.; Azenabor, A.; Ekun, A.O.; Enitan, S.S.; Adebola, O.K.; Ogundahunsi, O.A. Leptin: Adiponectin ratio discriminated the risk of metabolic syndrome better than adiponectin and leptin in Southwest Nigeria. *Diabetes Metab Syndr* 2019, *13*, 1845-1849, doi:10.1016/j.dsx.2019.04.008.

306. Ungru, J.; Bluher, M.; Coenen, M.; Raila, J.; Boston, R.; Vervuert, I. Effects of body weight reduction on blood adipokines and subcutaneous adipose tissue adipokine mRNA expression profiles in obese ponies. *Vet Rec* 2012, *171*, 528, doi:10.1136/vr.100911.
307. Saito, K.; Tobe, T.; Yoda, M.; Nakano, Y.; Choi-Miura, N.H.; Tomita, M. Regulation of gelatin-binding protein 28 (GBP28) gene expression by C/EBP. *Biol Pharm Bull* 1999, *22*, 1158-1162, doi:10.1248/bpb.22.1158.
308. Seo, J.B.; Moon, H.M.; Noh, M.J.; Lee, Y.S.; Jeong, H.W.; Yoo, E.J.; Kim, W.S.; Park, J.; Youn, B.S.; Kim, J.W., et al. Adipocyte determination- and differentiation-dependent factor 1/sterol regulatory element-binding protein 1c regulates mouse adiponectin expression. *J Biol Chem* 2004, *279*, 22108-22117, doi:10.1074/jbc.M400238200.
309. Maeda, N.; Takahashi, M.; Funahashi, T.; Kihara, S.; Nishizawa, H.; Kishida, K.; Nagaretani, H.; Matsuda, M.; Komuro, R.; Ouchi, N., et al. PPARgamma ligands increase expression and plasma concentrations of adiponectin, an adipose-derived protein. *Diabetes* 2001, *50*, 2094-2099, doi:10.2337/diabetes.50.9.2094.
310. Cnop, M.; Havel, P.J.; Utzschneider, K.M.; Carr, D.B.; Sinha, M.K.; Boyko, E.J.; Retzlaff, B.M.; Knopp, R.H.; Brunzell, J.D.; Kahn, S.E. Relationship of adiponectin to body fat distribution, insulin sensitivity and plasma lipoproteins: evidence for independent roles of age and sex. *Diabetologia* 2003, *46*, 459-469, doi:10.1007/s00125-003-1074-z.
311. Ouchi, N.; Walsh, K. Adiponectin as an anti-inflammatory factor. *Clin Chim Acta* 2007, *380*, 24-30, doi:10.1016/j.cca.2007.01.026.
312. Ouchi, N.; Kihara, S.; Arita, Y.; Maeda, K.; Kuriyama, H.; Okamoto, Y.; Hotta, K.; Nishida, M.; Takahashi, M.; Nakamura, T., et al. Novel modulator for endothelial adhesion molecules: adipocyte-derived plasma protein adiponectin. *Circulation* 1999, *100*, 2473-2476, doi:10.1161/01.cir.100.25.2473.
313. Ouchi, N.; Kihara, S.; Arita, Y.; Okamoto, Y.; Maeda, K.; Kuriyama, H.; Hotta, K.; Nishida, M.; Takahashi, M.; Muraguchi, M., et al. Adiponectin, an adipocyte-derived plasma protein, inhibits endothelial NF-kappaB signaling through a cAMP-dependent pathway. *Circulation* 2000, *102*, 1296-1301, doi:10.1161/01.cir.102.11.1296.
314. Ouedraogo, R.; Wu, X.; Xu, S.Q.; Fuchsel, L.; Motoshima, H.; Mahadev, K.; Hough, K.; Scalia, R.; Goldstein, B.J. Adiponectin suppression of high-glucose-induced reactive oxygen species in vascular endothelial cells: evidence for involvement of a cAMP signaling pathway. *Diabetes* 2006, *55*, 1840-1846, doi:10.2337/db05-1174.
315. Wulster-Radcliffe, M.C.; Ajuwon, K.M.; Wang, J.; Christian, J.A.; Spurlock, M.E. Adiponectin differentially regulates cytokines in porcine macrophages. *Biochem Biophys Res Commun* 2004, *316*, 924-929, doi:10.1016/j.bbrc.2004.02.130.
316. Kumada, M.; Kihara, S.; Ouchi, N.; Kobayashi, H.; Okamoto, Y.; Ohashi, K.; Maeda, K.; Nagaretani, H.; Kishida, K.; Maeda, N., et al. Adiponectin specifically increased tissue inhibitor of metalloproteinase-1 through interleukin-10 expression in human macrophages. *Circulation* 2004, *109*, 2046-2049, doi:10.1161/01.cir.0000127953.98131.ed.

317. Kubota, N.; Terauchi, Y.; Yamauchi, T.; Kubota, T.; Moroi, M.; Matsui, J.; Eto, K.; Yamashita, T.; Kamon, J.; Satoh, H., et al. Disruption of adiponectin causes insulin resistance and neointimal formation. *J Biol Chem* 2002, *277*, 25863-25866, doi:10.1074/jbc.C200251200.
318. Yamauchi, T.; Kamon, J.; Waki, H.; Terauchi, Y.; Kubota, N.; Hara, K.; Mori, Y.; Ide, T.; Murakami, K.; Tsuboyama-Kasaoka, N., et al. The fat-derived hormone adiponectin reverses insulin resistance associated with both lipoatrophy and obesity. *Nat Med* 2001, *7*, 941-946, doi:10.1038/90984.
319. Hotta, K.; Funahashi, T.; Bodkin, N.L.; Ortmeyer, H.K.; Arita, Y.; Hansen, B.C.; Matsuzawa, Y. Circulating concentrations of the adipocyte protein adiponectin are decreased in parallel with reduced insulin sensitivity during the progression to type 2 diabetes in rhesus monkeys. *Diabetes* 2001, *50*, 1126-1133, doi:10.2337/diabetes.50.5.1126.
320. Yamauchi, T.; Kadowaki, T. Adiponectin receptor as a key player in healthy longevity and obesity-related diseases. *Cell Metab* 2013, *17*, 185-196, doi:10.1016/j.cmet.2013.01.001.
321. Yamauchi, T.; Kamon, J.; Ito, Y.; Tsuchida, A.; Yokomizo, T.; Kita, S.; Sugiyama, T.; Miyagishi, M.; Hara, K.; Tsunoda, M., et al. Cloning of adiponectin receptors that mediate antidiabetic metabolic effects. *Nature* 2003, *423*, 762-769, doi:10.1038/nature01705.
322. Fordham, T.; Morgan, R.A.
323. Hara, K.; Horikoshi, M.; Yamauchi, T.; Yago, H.; Miyazaki, O.; Ebinuma, H.; Imai, Y.; Nagai, R.; Kadowaki, T. Measurement of the high-molecular weight form of adiponectin in plasma is useful for the prediction of insulin resistance and metabolic syndrome. *Diabetes Care* 2006, *29*, 1357-1362, doi:10.2337/dc05-1801.
324. Pajvani, U.B.; Du, X.; Combs, T.P.; Berg, A.H.; Rajala, M.W.; Schulthess, T.; Engel, J.; Brownlee, M.; Scherer, P.E. Structure-function studies of the adipocyte-secreted hormone Acrp30/adiponectin. Implications fpr metabolic regulation and bioactivity. *J Biol Chem* 2003, *278*, 9073-9085, doi:10.1074/jbc.M207198200.
325. Tsuchida, A.; Yamauchi, T.; Ito, Y.; Hada, Y.; Maki, T.; Takekawa, S.; Kamon, J.; Kobayashi, M.; Suzuki, R.; Hara, K., et al. Insulin/Foxo1 pathway regulates expression levels of adiponectin receptors and adiponectin sensitivity. *J Biol Chem* 2004, *279*, 30817-30822, doi:10.1074/jbc.M402367200.
326. Semple, R.K.; Cochran, E.K.; Soos, M.A.; Burling, K.A.; Savage, D.B.; Gorden, P.; O'Rahilly, S. Plasma adiponectin as a marker of insulin receptor dysfunction: clinical utility in severe insulin resistance. *Diabetes Care* 2008, *31*, 977-979, doi:10.2337/dc07-2194.
327. Yamamoto, Y.; Hirose, H.; Saito, I.; Tomita, M.; Taniyama, M.; Matsubara, K.; Okazaki, Y.; Ishii, T.; Nishikai, K.; Saruta, T. Correlation of the adipocyte-derived protein adiponectin with insulin resistance index and serum high-density lipoprotein-cholesterol, independent of body mass index, in the Japanese population. *Clin Sci (Lond)* 2002, *103*, 137-142, doi:10.1042/.
328. Menzies-Gow, N.J.; Harris, P.A.; Elliott, J. Prospective cohort study evaluating risk factors for the development of pasture-associated laminitis in the United Kingdom. *Equine Vet J* 2017, *49*, 300-306, doi:10.1111/evj.12606.

329. Kue Young, T.; Chateau, D.; Zhang, M. Factor analysis of ethnic variation in the multiple metabolic (insulin resistance) syndrome in three Canadian populations. *Am J Hum Biol* 2002, *14*, 649-658, doi:10.1002/ajhb.10083.
330. Lee, K.E.; Klein, B.E.; Klein, R. Familial aggregation of components of the multiple metabolic syndrome in the Framingham Heart and Offspring Cohorts: Genetic Analysis Workshop Problem 1. *BMC Genetics* 2003, *4 Suppl 1*, S94, doi:10.1186/1471-2156-4-s1-s94.
331. Stoger, R. The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? *Bioessays* 2008, *30*, 156-166, doi:10.1002/bies.20700.
332. Speakman, J.R. Thrifty genes for obesity, an attractive but flawed idea, and an alternative perspective: the 'drifty gene' hypothesis. *Int J Obes (Lond)* 2008, *32*, 1611-1617, doi:10.1038/ijo.2008.161.
333. Qasim, A.; Turcotte, M.; de Souza, R.J.; Samaan, M.C.; Champredon, D.; Dushoff, J.; Speakman, J.R.; Meyre, D. On the origin of obesity: identifying the biological, environmental and cultural drivers of genetic risk among human populations. *Obes Rev* 2018, *19*, 121-149, doi:10.1111/obr.12625.
334. Li, Y.I.; van de Geijn, B.; Raj, A.; Knowles, D.A.; Petti, A.A.; Golan, D.; Gilad, Y.; Pritchard, J.K. RNA splicing is a primary link between genetic variation and disease. *Science (New York, N.Y.)* 2016, *352*, 600-604, doi:10.1126/science.aad9417.
335. Pickrell, J.K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am J Hum Genet* 2014, *94*, 559-573, doi:10.1016/j.ajhg.2014.03.004.
336. Welter, D.; MacArthur, J.; Morales, J.; Burdett, T.; Hall, P.; Junkins, H.; Klemm, A.; Flicek, P.; Manolio, T.; Hindorff, L., et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 2014, *42*, D1001-1006, doi:10.1093/nar/gkt1229.
337. Visscher, P.M.; Hill, W.G.; Wray, N.R. Heritability in the genomics era--concepts and misconceptions. *Nature Reviews. Genetics* 2008, *9*, 255-266, doi:10.1038/nrg2322.
338. Bailey, E. Heritability and the equine clinician. *Equine Vet J* 2014, *46*, 12-14, doi:10.1111/evj.12196.
339. Matthews, D.R.; Hosker, J.P.; Rudenski, A.S.; Naylor, B.A.; Treacher, D.F.; Turner, R.C. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 1985, *28*, 412-419.
340. Graziano, F.; Grassi, M.; Sacco, S.; Concas, M.P.; Vaccargiu, S.; Pirastu, M.; Biino, G. Probing the factor structure of metabolic syndrome in Sardinian genetic isolates. *Nutr Metab Cardiovasc Dis* 2015, *25*, 548-555, doi:10.1016/j.numecd.2015.02.004.
341. Bayoumi, R.A.; Al-Yahyaee, S.A.; Albarwani, S.A.; Rizvi, S.G.; Al-Hadabi, S.; Al-Ubaidi, F.F.; Al-Hinai, A.T.; Al-Kindi, M.N.; Adnan, H.T.; Al-Barwany, H.S., et al. Heritability of determinants of the metabolic syndrome among healthy Arabs of the Oman family study. *Obesity (Silver Spring)* 2007, *15*, 551-556, doi:10.1038/oby.2007.555.

342. Henneman, P.; Aulchenko, Y.S.; Frants, R.R.; van Dijk, K.W.; Oostra, B.A.; van Duijn, C.M. Prevalence and heritability of the metabolic syndrome and its individual components in a Dutch isolate: the Erasmus Rucphen Family study. *J Med Genet* 2008, *45*, 572-577, doi:10.1136/jmg.2008.058388.
343. Khan, R.J.; Gebreab, S.Y.; Sims, M.; Riestra, P.; Xu, R.; Davis, S.K. Prevalence, associated factors and heritabilities of metabolic syndrome and its individual components in African Americans: the Jackson Heart Study. *BMJ Open* 2015, *5*, e008675, doi:10.1136/bmjopen-2015-008675.
344. Lin, H.F.; Boden-Albala, B.; Juo, S.H.; Park, N.; Rundek, T.; Sacco, R.L. Heritabilities of the metabolic syndrome and its components in the Northern Manhattan Family Study. *Diabetologia* 2005, *48*, 2006-2012, doi:10.1007/s00125-005-1892-2.
345. Austin, M.A.; Edwards, K.L.; McNeely, M.J.; Chandler, W.L.; Leonetti, D.L.; Talmud, P.J.; Humphries, S.E.; Fujimoto, W.Y. Heritability of multivariate factors of the metabolic syndrome in nondiabetic Japanese americans. *Diabetes* 2004, *53*, 1166-1169, doi:10.2337/diabetes.53.4.1166.
346. Bellia, A.; Giardina, E.; Lauro, D.; Tesauro, M.; Di Fede, G.; Cusumano, G.; Federici, M.; Rini, G.B.; Novelli, G.; Lauro, R., et al. "The Linosa Study": epidemiological and heritability data of the metabolic syndrome in a Caucasian genetic isolate. *Nutr Metab Cardiovasc Dis* 2009, *19*, 455-461, doi:10.1016/j.numecd.2008.11.002.
347. Graziano, F.; Biino, G.; Bonati, M.T.; Neale, B.M.; Do, R.; Concas, M.P.; Vaccargiu, S.; Pirastu, M.; Terradura-Vagnarelli, O.; Cirillo, M., et al. Estimation of metabolic syndrome heritability in three large populations including full pedigree and genomic information. *Human Genetics* 2019, *138*, 739-748, doi:10.1007/s00439-019-02024-6.
348. Henneman, P.; Aulchenko, Y.S.; Frants, R.R.; Zorkoltseva, I.V.; Zillikens, M.C.; Frolich, M.; Oostra, B.A.; van Dijk, K.W.; van Duijn, C.M. Genetic architecture of plasma adiponectin overlaps with the genetics of metabolic syndrome-related traits. *Diabetes Care* 2010, *33*, 908-913, doi:10.2337/dc09-1385.
349. Martin, L.J.; North, K.E.; Dyer, T.; Blangero, J.; Comuzzie, A.G.; Williams, J. Phenotypic, genetic, and genome-wide structure in the metabolic syndrome. *BMC Genetics* 2003, *4 Suppl 1*, S95, doi:10.1186/1471-2156-4-s1-s95.
350. Mills, G.W.; Avery, P.J.; McCarthy, M.I.; Hattersley, A.T.; Levy, J.C.; Hitman, G.A.; Sampson, M.; Walker, M. Heritability estimates for beta cell function and features of the insulin resistance syndrome in UK families with an increased susceptibility to type 2 diabetes. *Diabetologia* 2004, *47*, 732-738, doi:10.1007/s00125-004-1338-2.
351. Wu, K.D.; Hsiao, C.F.; Ho, L.T.; Sheu, W.H.; Pei, D.; Chuang, L.M.; Curb, D.; Chen, Y.D.; Tsai, H.J.; Dzau, V.J., et al. Clustering and heritability of insulin resistance in Chinese and Japanese hypertensive families: a Stanford-Asian Pacific Program in Hypertension and Insulin Resistance sibling study. *Hypertens Res* 2002, *25*, 529-536.
352. Zarkesh, M.; Daneshpour, M.S.; Faam, B.; Fallah, M.S.; Hosseinzadeh, N.; Guity, K.; Hosseinpanah, F.; Momenan, A.A.; Azizi, F. Heritability of the metabolic

syndrome and its components in the Tehran Lipid and Glucose Study (TLGS). *Genet Res (Camb)* 2012, *94*, 331-337, doi:10.1017/s001667231200050x.
353. Musani, S.K.; Martin, L.J.; Woo, J.G.; Olivier, M.; Gurka, M.J.; DeBoer, M.D. Heritability of the Severity of the Metabolic Syndrome in Whites and Blacks in 3 Large Cohorts. *Circ Cardiovasc Genet* 2017, *10*, doi:10.1161/circgenetics.116.001621.
354. Lee, C.; Pollak, E.J. Influence of sire misidentification on sire x year interaction variance and direct-maternal genetic covariance for weaning weight in beef cattle. *J Anim Sci* 1997, *75*, 2858-2863.
355. Yang, J.; Benyamin, B.; McEvoy, B.P.; Gordon, S.; Henders, A.K.; Nyholt, D.R.; Madden, P.A.; Heath, A.C.; Martin, N.G.; Montgomery, G.W., et al. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* 2010, *42*, 565-569, doi:10.1038/ng.608.
356. Lee, S.H.; DeCandia, T.R.; Ripke, S.; Yang, J.; Sullivan, P.F.; Goddard, M.E.; Keller, M.C.; Visscher, P.M.; Wray, N.R. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature Genetics* 2012, *44*, 247-250, doi:10.1038/ng.1108.
357. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011, *88*, 76-82, doi:10.1016/j.ajhg.2010.11.011.
358. Lee, S.H.; Wray, N.R.; Goddard, M.E.; Visscher, P.M. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 2011, *88*, 294-305, doi:10.1016/j.ajhg.2011.02.002.
359. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A., et al. Finding the missing heritability of complex diseases. *Nature* 2009, *461*, 747-753, doi:10.1038/nature08494.
360. Vattikuti, S.; Guo, J.; Chow, C.C. Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. *PLoS Genetics* 2012, *8*, e1002637, doi:10.1371/journal.pgen.1002637.
361. Dawn Teare, M.; Barrett, J.H. Genetic linkage studies. *Lancet* 2005, *366*, 1036-1044, doi:10.1016/s0140-6736(05)67382-5.
362. Ott, J.; Wang, J.; Leal, S.M. Genetic linkage analysis in the age of whole-genome sequencing. *Nature Reviews Genetics* 2015, *16*, 275, doi:10.1038/nrg3908.
363. Kissebah, A.H.; Sonnenberg, G.E.; Myklebust, J.; Goldstein, M.; Broman, K.; James, R.G.; Marks, J.A.; Krakower, G.R.; Jacob, H.J.; Weber, J., et al. Quantitative trait loci on chromosomes 3 and 17 influence phenotypes of the metabolic syndrome. *Proc Natl Acad Sci USA* 2000, *97*, 14478-14483, doi:10.1073/pnas.97.26.14478.
364. Tang, W.; Miller, M.B.; Rich, S.S.; North, K.E.; Pankow, J.S.; Borecki, I.B.; Myers, R.H.; Hopkins, P.N.; Leppert, M.; Arnett, D.K. Linkage analysis of a composite factor for the multiple metabolic syndrome: the National Heart, Lung, and Blood Institute Family Heart Study. *Diabetes* 2003, *52*, 2840-2847, doi:10.2337/diabetes.52.11.2840.
365. Hoffmann, K.; Mattheisen, M.; Dahm, S.; Nurnberg, P.; Roe, C.; Johnson, J.; Cox, N.J.; Wichmann, H.E.; Wienker, T.F.; Schulze, J., et al. A German genome-wide

linkage scan for type 2 diabetes supports the existence of a metabolic syndrome locus on chromosome 1p36.13 and a type 2 diabetes locus on chromosome 16p12.2. *Diabetologia* 2007, *50*, 1418-1422, doi:10.1007/s00125-007-0658-4.

366. Puppala, S.; Dodd, G.D.; Fowler, S.; Arya, R.; Schneider, J.; Farook, V.S.; Granato, R.; Dyer, T.D.; Almasy, L.; Jenkinson, C.P., et al. A genomewide search finds major susceptibility loci for gallbladder disease on chromosome 1 in Mexican Americans. *Am J Hum Genet* 2006, *78*, 377-392, doi:10.1086/500274.
367. Cai, G.; Cole, S.A.; Freeland-Graves, J.H.; MacCluer, J.W.; Blangero, J.; Comuzzie, A.G. Principal component for metabolic syndrome risk maps to chromosome 4p in Mexican Americans: the San Antonio Family Heart Study. *Hum Biol* 2004, *76*, 651-665.
368. Bosse, Y.; Despres, J.P.; Chagnon, Y.C.; Rice, T.; Rao, D.C.; Bouchard, C.; Perusse, L.; Vohl, M.C. Quantitative trait locus on 15q for a metabolic syndrome variable derived from factor analysis. *Obesity (Silver Spring)* 2007, *15*, 544-550, doi:10.1038/oby.2007.577.
369. Edwards, K.L.; Hutter, C.M.; Wan, J.Y.; Kim, H.; Monks, S.A. Genome-wide linkage scan for the metabolic syndrome: the GENNID study. *Obesity (Silver Spring)* 2008, *16*, 1596-1601, doi:10.1038/oby.2008.236.
370. Loos, R.J.; Katzmarzyk, P.T.; Rao, D.C.; Rice, T.; Leon, A.S.; Skinner, J.S.; Wilmore, J.H.; Rankinen, T.; Bouchard, C. Genome-wide linkage scan for the metabolic syndrome in the HERITAGE Family Study. *J Clin Endocrinol Metab* 2003, *88*, 5935-5943, doi:10.1210/jc.2003-030553.
371. Pearson, T.A.; Manolio, T.A. How to interpret a genome-wide association study. *JAMA* 2008, *299*, 1335-1344, doi:10.1001/jama.299.11.1335.
372. Lanktree, M.B.; Johansen, C.T.; Joy, T.R.; Hegele, R.A. A translational view of the genetics of lipodystrophy and ectopic fat deposition. *Prog Mol Biol Transl Sci* 2010, *94*, 159-196, doi:10.1016/s1877-1173(10)94006-4
10.1016/b978-0-12-375003-7.00006-6.
373. Lahiry, P.; Pollex, R.L.; Hegele, R.A. Uncloaking the genetic determinants of metabolic syndrome. *J Nutrigenet Nutrigenomics* 2008, *1*, 118-125, doi:10.1159/000112459.
374. Cargill, M.; Altshuler, D.; Ireland, J.; Sklar, P.; Ardlie, K.; Patil, N.; Shaw, N.; Lane, C.R.; Lim, E.P.; Kalyanaraman, N., et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genetics* 1999, *22*, 231-238, doi:10.1038/10290.
375. Blanco-Gómez, A.; Castillo-Lluva, S.; Del Mar Sáez-Freire, M.; Hontecillas-Prieto, L.; Mao, J.H.; Castellanos-Martín, A.; Pérez-Losada, J. Missing heritability of complex diseases: Enlightenment by genetic variants from intermediate phenotypes. *BioEssays* 2016, *38*, 664-673, doi:10.1002/bies.201600084.
376. Genin, E. Missing heritability of complex diseases: case solved? *Human Genetics* 2019, 10.1007/s00439-019-02034-4, doi:10.1007/s00439-019-02034-4.
377. Young, A.I. Solving the missing heritability problem. *PLoS Genetics* 2019, *15*, e1008222, doi:10.1371/journal.pgen.1008222.
378. Yu, J.; Pressoir, G.; Briggs, W.H.; Vroh Bi, I.; Yamasaki, M.; Doebley, J.F.; McMullen, M.D.; Gaut, B.S.; Nielsen, D.M.; Holland, J.B., et al. A unified mixed-

model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 2006, *38*, 203-208, doi:10.1038/ng1702.
379. Zhou, X.; Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* 2012, *44*, 821-824, doi:10.1038/ng.2310.
380. Chanock, S.J.; Manolio, T.; Boehnke, M.; Boerwinkle, E.; Hunter, D.J.; Thomas, G.; Hirschhorn, J.N.; Abecasis, G.; Altshuler, D.; Bailey-Wilson, J.E., et al. Replicating genotype-phenotype associations. *Nature* 2007, *447*, 655-660, doi:10.1038/447655a.
381. Ioannidis, J.P.; Ntzani, E.E.; Trikalinos, T.A.; Contopoulos-Ioannidis, D.G. Replication validity of genetic association studies. *Nature Genetics* 2001, *29*, 306-309, doi:10.1038/ng749.
382. König, I.R. Validation in Genetic Association Studies. *Briefings in Bioinformatics* 2011, *12*, 253-258, doi:10.1093/bib/bbq074.
383. Kraft, P. Curses--winner's and otherwise--in genetic epidemiology. *Epidemiology* 2008, *19*, 649-651; discussion 657-648, doi:10.1097/EDE.0b013e318181b865.
384. Wood, A.R.; Esko, T.; Yang, J.; Vedantam, S.; Pers, T.H.; Gustafsson, S.; Chu, A.Y.; Estrada, K.; Luan, J.; Kutalik, Z., et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics* 2014, *46*, 1173-1186, doi:10.1038/ng.3097.
385. Goddard, M.E.; Kemper, K.E.; MacLeod, I.M.; Chamberlain, A.J.; Hayes, B.J. Genetics of complex traits: prediction of phenotype, identification of causal polymorphisms and genetic architecture. *Proc Biol Sci* 2016, *283*, doi:10.1098/rspb.2016.0569.
386. Zabaneh, D.; Balding, D.J. A Genome-Wide Association Study of the Metabolic Syndrome in Indian Asian Men. *PloS One* 2010, *5*, e11961, doi:10.1371/journal.pone.0011961.
387. Kristiansson, K.; Perola, M.; Tikkanen, E.; Kettunen, J.; Surakka, I.; Havulinna, A.S.; Stancakova, A.; Barnes, C.; Widen, E.; Kajantie, E., et al. Genome-wide screen for metabolic syndrome susceptibility Loci reveals strong lipid gene contribution but no evidence for common genetic basis for clustering of metabolic syndrome traits. *Circ Cardiovasc Genet* 2012, *5*, 242-249, doi:10.1161/circgenetics.111.961482.
388. Tekola-Ayele, F.; Doumatey, A.P.; Shriner, D.; Bentley, A.R.; Chen, G.; Zhou, J.; Fasanmade, O.; Johnson, T.; Oli, J.; Okafor, G., et al. Genome-wide association study identifies African-ancestry specific variants for metabolic syndrome. *Mol Genet Metab* 2015, *116*, 305-313, doi:https://doi.org/10.1016/j.ymgme.2015.10.008.
389. Avery, C.L.; He, Q.; North, K.E.; Ambite, J.L.; Boerwinkle, E.; Fornage, M.; Hindorff, L.A.; Kooperberg, C.; Meigs, J.B.; Pankow, J.S., et al. A phenomics-based strategy identifies loci on APOC1, BRAP, and PLCG1 associated with metabolic syndrome phenotype domains. *PLoS Genetics* 2011, *7*, e1002322, doi:10.1371/journal.pgen.1002322.
390. Carty, C.L.; Bhattacharjee, S.; Haessler, J.; Cheng, I.; Hindorff, L.A.; Aroda, V.; Carlson, C.S.; Hsu, C.-N.; Wilkens, L.; Liu, S., et al. Analysis of Metabolic Syndrome Components in >15 000 African Americans Identifies Pleiotropic

Variants. Circ Cardiovasc Genet 2014, *7*, 505-513, doi:10.1161/CIRCGENETICS.113.000386.
391. Kraja, A.T.; Vaidya, D.; Pankow, J.S.; Goodarzi, M.O.; Assimes, T.L.; Kullo, I.J.; Sovio, U.; Mathias, R.A.; Sun, Y.V.; Franceschini, N., et al. A bivariate genome-wide approach to metabolic syndrome: STAMPEED consortium. *Diabetes* 2011, *60*, 1329-1339, doi:10.2337/db10-1011.
392. Kong, S.; Cho, Y.S. Identification of female-specific genetic variants for metabolic syndrome and its component traits to improve the prediction of metabolic syndrome in females. *BMC Medical Genetics* 2019, *20*, 99-99, doi:10.1186/s12881-019-0830-y.
393. Fathi Dizaji, B. The investigations of genetic determinants of the metabolic syndrome. *Diabetes Metab Syndr* 2018, *12*, 783-789, doi:10.1016/j.dsx.2018.04.009.
394. Povel, C.M.; Boer, J.M.; Reiling, E.; Feskens, E.J. Genetic variants and the metabolic syndrome: a systematic review. *Obes Rev* 2011, *12*, 952-967, doi:10.1111/j.1467-789X.2011.00907.x.
395. Makvandi-Nejad, S.; Hoffman, G.E.; Allen, J.J.; Chu, E.; Gu, E.; Chandler, A.M.; Loredo, A.I.; Bellone, R.R.; Mezey, J.G.; Brooks, S.A., et al. Four loci explain 83% of size variation in the horse. *PloS One* 2012, *7*, e39929, doi:10.1371/journal.pone.0039929.
396. Schroderus, E.; Ojala, M. Estimates of genetic parameters for conformation measures and scores in Finnhorse and Standardbred foals. *J Anim Breed Genet* 2010, *127*, 395-403, doi:10.1111/j.1439-0388.2010.00856.x.
397. Perola, M. Genome-wide association approaches for identifying loci for human height genes. *Best Pract Res Clin Endocrinol Metab* 2011, *25*, 19-23, doi:10.1016/j.beem.2010.10.013.
398. Institute of Medicine Committee on Assessing Interactions Among Social, B.; Genetic Factors in, H. The National Academies Collection: Reports funded by National Institutes of Health. In *Genes, Behavior, and the Social Environment: Moving Beyond the Nature/Nurture Debate*, Hernandez, L.M., Blazer, D.G., Eds. National Academies Press (US). National Academy of Sciences.: Washington (DC), 2006; 10.17226/11693.
399. Hasin, Y.; Seldin, M.; Lusis, A. Multi-omics approaches to disease. *Genome Biol* 2017, *18*, 83, doi:10.1186/s13059-017-1215-1.
400. Andersson, L.; Georges, M. Domestic-animal genomics: deciphering the genetics of complex traits. *Nature Reviews. Genetics* 2004, *5*, 202-212, doi:10.1038/nrg1294.
401. Lewis, S.L.; Holl, H.M.; Streeter, C.; Posbergh, C.; Schanbacher, B.J.; Place, N.J.; Mallicote, M.F.; Long, M.T.; Brooks, S.A. Genomewide association study reveals a risk locus for equine metabolic syndrome in the Arabian horse. *J Anim Sci* 2017, *95*, 1071-1079, doi:10.2527/jas.2016.1221.
402. Stefaniuk-Szmukier, M.; Ropla-Molik, K.; Bugno-Poniewierska, M. Identyfikacja wariantu w rejonie genu FAM14A potencjalnie związanego z występowaniem syndromu metabolicznego (EMS) u koni czystej krwi arabskiej. *Wiadomosci Aootechniczne* 2017, *5*, 46-48.

403. Stancakova, A.; Laakso, M. Genetics of metabolic syndrome. *Reviews in Endocrine & Metabolic Disorders* 2014, *15*, 243-252, doi:10.1007/s11154-014-9293-9.
404. McCue, M.E.; Geor, R.J.; Schultz, N. Equine Metabolic Syndrome: A Complex Disease Influenced by Genetics and the Environment. *J Equine Vet Sci* 2015, *35*, 367-375, doi:10.1016/j.jevs.2015.03.004.
405. Yang, J.; Zeng, J.; Goddard, M.E.; Wray, N.R.; Visscher, P.M. Concepts, estimation and interpretation of SNP-based heritability. *Nature Genetics* 2017, *49*, 1304-1310, doi:10.1038/ng.3941.
406. Schuver, A.; Frank, N.; Chameroy, K.A.; Elliott, S.B. Assessment of Insulin and Glucose Dynamics by Using an Oral Sugar Test in Horses. *J Equine Vet Sci* 2014, *34*, 465-470, doi:10.1016/j.jevs.2013.09.006.
407. Browning, S.R.; Browning, B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 2007, *81*, 1084-1097, doi:10.1086/521987.
408. McCoy, A.M.; McCue, M.E. Validation of imputation between equine genotyping arrays. *Anim Genet* 2014, *45*, 153, doi:10.1111/age.12093.
409. Schaefer, R.J.; Schubert, M.; Bailey, E.; Bannasch, D.L.; Barrey, E.; Bar-Gal, G.K.; Brem, G.; Brooks, S.A.; Distl, O.; Fries, R., et al. Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds. *BMC Genomics* 2017, *18*, 565, doi:10.1186/s12864-017-3943-8.
410. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; Daly, M.J., et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007, *81*, 559-575, doi:10.1086/519795.
411. Speed, D.; Hemani, G.; Johnson, M.R.; Balding, D.J. Improved heritability estimation from genome-wide SNPs. *Am J Hum Genet* 2012, *91*, 1011-1021, doi:10.1016/j.ajhg.2012.10.010.
412. R Core Team *R: A language and environment for statistical computing*, R Foundation for Statistical Computing: Vienna, Austria, 2015.
413. Speed, D.; Cai, N.; Consortium, U.; Johnson, M.R.; Nejentsev, S.; Balding, D.J. Reevaluation of SNP heritability in complex human traits. *Nature Genetics* 2017, *49*, 986-992, doi:10.1038/ng.3865.
414. Menzies-Gow, N.J. Endocrinopathic laminitis: reducing the risk through diet and exercise. *Vet Clin North Am Equine Pract* 2010, *26*, 371-378, doi:10.1016/j.cveq.2010.04.005.
415. Norton, E.M.; Mickelson, J.R.; Binns, M.M.; Blott, S.C.; Caputo, P.; Isgren, C.M.; McCoy, A.M.; Moore, A.; Piercy, R.J.; Swinburne, J.E., et al. Heritability of Recurrent Exertional Rhabdomyolysis in Standardbred and Thoroughbred Racehorses Derived From SNP Genotyping Data. *J Hered* 2016, *107*, 537-543, doi:10.1093/jhered/esw042.
416. DeMenna, J.; Puppala, S.; Chittoor, G.; Schneider, J.; Kim, J.Y.; Shaibi, G.Q.; Mandarino, L.J.; Duggirala, R.; Coletta, D.K. Association of common genetic variants with diabetes and metabolic syndrome related traits in the Arizona Insulin Resistance registry: a focus on Mexican American families in the Southwest. *Hum Hered* 2014, *78*, 47-58, doi:10.1159/000363411.

417. Herbeth, B.; Samara, A.; Ndiaye, C.; Marteau, J.B.; Berrahmoune, H.; Siest, G.; Visvikis-Siest, S. Metabolic syndrome-related composite factors over 5 years in the STANISLAS family study: genetic heritability and common environmental influences. *Clinica Chim Acta* 2010, *411*, 833-839, doi:10.1016/j.cca.2010.02.070.
418. Naccache, F.; Metzger, J.; Distl, O. Genetic risk factors for osteochondrosis in various horse breeds. *Equine Vet J* 2018, *50*, 556-563, doi:10.1111/evj.12824.
419. Russell, J.; Matika, O.; Russell, T.; Reardon, R.J. Heritability and prevalence of selected osteochondrosis lesions in yearling Thoroughbred horses. *Equine Vet J* 2017, *49*, 282-287, doi:10.1111/evj.12613.
420. Velie, B.D.; Hamilton, N.A.; Wade, C.M. Heritability of racing performance in the Australian Thoroughbred racing population. *Anim Genet* 2015, *46*, 23-29, doi:10.1111/age.12234.
421. Kamezaki, F.; Sonoda, S.; Tomotsune, Y.; Yunaka, H.; Otsuji, Y. Seasonal variation in metabolic syndrome prevalence. *Hypertens Res* 2010, *33*, 568-572, doi:10.1038/hr.2010.32.
422. Lutsey, P.L.; Steffen, L.M.; Stevens, J. Dietary intake and the development of the metabolic syndrome: the Atherosclerosis Risk in Communities study. *Circulation* 2008, *117*, 754-761, doi:10.1161/circulationaha.107.716159.
423. Park, Y.W.; Zhu, S.; Palaniappan, L.; Heshka, S.; Carnethon, M.R.; Heymsfield, S.B. The metabolic syndrome: prevalence and associated risk factor findings in the US population from the Third National Health and Nutrition Examination Survey, 1988-1994. *Arch Intern Med* 2003, *163*, 427-436.
424. Shi, H.; Kichaev, G.; Pasaniuc, B. Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *Am J Hum Genet* 2016, *99*, 139-153, doi:10.1016/j.ajhg.2016.05.013.
425. Kissebah, A.H.; Sonnenberg, G.E.; Myklebust, J.; Goldstein, M.; Broman, K.; James, R.G.; Marks, J.A.; Krakower, G.R.; Jacob, H.J.; Weber, J., et al. Quantitative trait loci on chromosomes 3 and 17 influence phenotypes of the metabolic syndrome. *Proc Natl Acad Sci USA* 2000, *97*, 14478-14483, doi:10.1073/pnas.97.26.14478.
426. Fall, T.; Ingelsson, E. Genome-wide association studies of obesity and metabolic syndrome. *Mol Cell Endocrinol* 2014, *382*, 740-757, doi:10.1016/j.mce.2012.08.018.
427. Beeson, S.K.; Schaefer, R.J.; Mason, V.C.; McCue, M.E. Robust remapping of equine SNP array coordinates to EquCab3. *Anim Genet* 2019, *50*, 114-115, doi:10.1111/age.12745.
428. Freedman, M.L.; Reich, D.; Penney, K.L.; McDonald, G.J.; Mignault, A.A.; Patterson, N.; Gabriel, S.B.; Topol, E.J.; Smoller, J.W.; Pato, C.N., et al. Assessing the impact of population stratification on genetic association studies. *Nature Genetics* 2004, *36*, 388-393, doi:10.1038/ng1333.
429. Pinheiro J; Bates D; DebRoy S; Sarkar D; Team, R.C. nlme: Linear and Nonlinear Mixed Effects Model. *R Package Version 3.1-137* 2018.
430. Zhou, X.; Carbonetto, P.; Stephens, M. Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genetics* 2013, *9*, e1003264, doi:10.1371/journal.pgen.1003264.

431. Lippert, C.; Listgarten, J.; Liu, Y.; Kadie, C.M.; Davidson, R.I.; Heckerman, D. FaST linear mixed models for genome-wide association studies. *Nature Methods* 2011, *8*, 833-835, doi:10.1038/nmeth.1681.
432. O., T. GNU Parallel 2016, May 2016. *ISBN 9781387509881, DOI https://doi.org/10.5281/zenodo.1146014* 2016.
433. Speidel, S.E.; Buckley, B.A.; Boldt, R.J.; Enns, R.M.; Lee, J.; Spangler, M.L.; Thomas, M.G. Genome-wide association study of Stayability and Heifer Pregnancy in Red Angus cattle. *J Anim Sci* 2018, *96*, 846-853, doi:10.1093/jas/sky041.
434. Li, M.X.; Yeung, J.M.; Cherny, S.S.; Sham, P.C. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics* 2012, *131*, 747-756, doi:10.1007/s00439-011-1118-2.
435. Schroder, W.; Klostermann, A.; Stock, K.F.; Distl, O. A genome-wide association study for quantitative trait loci of show-jumping in Hanoverian warmblood horses. *Anim Genet* 2012, *43*, 392-400, doi:10.1111/j.1365-2052.2011.02265.x.
436. Lykkjen, S.; Dolvik, N.I.; McCue, M.E.; Rendahl, A.K.; Mickelson, J.R.; Roed, K.H. Genome-wide association analysis of osteochondrosis of the tibiotarsal joint in Norwegian Standardbred trotters. *Anim Genet* 2010, *41 Suppl 2*, 111-120, doi:10.1111/j.1365-2052.2010.02117.x.
437. Han, B.; Eskin, E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet* 2011, *88*, 586-598, doi:10.1016/j.ajhg.2011.04.014.
438. McCue, M.E.; Bannasch, D.L.; Petersen, J.L.; Gurr, J.; Bailey, E.; Binns, M.M.; Distl, O.; Guerin, G.; Hasegawa, T.; Hill, E.W., et al. A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genetics* 2012, *8*, e1002451, doi:10.1371/journal.pgen.1002451.
439. Petersen, J.L.; Mickelson, J.R.; Rendahl, A.K.; Valberg, S.J.; Andersson, L.S.; Axelsson, J.; Bailey, E.; Bannasch, D.; Binns, M.M.; Borges, A.S., et al. Genome-wide analysis reveals selection for important traits in domestic horse breeds. *PLoS Genetics* 2013, *9*, e1003211, doi:10.1371/journal.pgen.1003211.
440. Corbin, L.J.; Blott, S.C.; Swinburne, J.E.; Vaudin, M.; Bishop, S.C.; Woolliams, J.A. Linkage disequilibrium and historical effective population size in the Thoroughbred horse. *Anim Genet* 2010, *41 Suppl 2*, 8-15, doi:10.1111/j.1365-2052.2010.02092.x.
441. Petersen, J.L.; Mickelson, J.R.; Cothran, E.G.; Andersson, L.S.; Axelsson, J.; Bailey, E.; Bannasch, D.; Binns, M.M.; Borges, A.S.; Brama, P., et al. Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PloS One* 2013, *8*, e54997, doi:10.1371/journal.pone.0054997.
442. Schaefer, R.J.; Schubert, M.; Bailey, E.; Bannasch, D.L.; Barrey, E.; Bar-Gal, G.K.; Brem, G.; Brooks, S.A.; Distl, O.; Fries, R., et al. Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds. *BMC Genomics* 2017, *18*, 565, doi:10.1186/s12864-017-3943-8.
443. Wade, C.M.; Giulotto, E.; Sigurdsson, S.; Zoli, M.; Gnerre, S.; Imsland, F.; Lear, T.L.; Adelson, D.L.; Bailey, E.; Bellone, R.R., et al. Genome sequence,

comparative analysis, and population genetics of the domestic horse. *Science (New York, N.Y.)* 2009, *326*, 865-867, doi:10.1126/science.1178158.
444. Durinck, S.; Spellman, P.T.; Birney, E.; Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 2009, *4*, 1184-1191, doi:10.1038/nprot.2009.97.
445. Kalbfleisch, T.S.; Rice, E.S.; DePriest, M.S., Jr.; Walenz, B.P.; Hestand, M.S.; Vermeesch, J.R.; BL, O.C.; Fiddes, I.T.; Vershinina, A.O.; Saremi, N.F., et al. Improved reference genome for the domestic horse increases assembly contiguity and composition. *Commun Biol* 2018, *1*, 197, doi:10.1038/s42003-018-0199-z.
446. Petersen, J.L.; Mickelson, J.R.; Cleary, K.D.; McCue, M.E. The American Quarter Horse: population structure and relationship to the thoroughbred. *J Hered* 2014, *105*, 148-162, doi:10.1093/jhered/est079.
447. Peters, U.; Bien, S.; Zubair, N. Genetic architecture of colorectal cancer. *Gut* 2015, *64*, 1623-1636, doi:10.1136/gutjnl-2013-306705.
448. Ardlie, K.G.; Kruglyak, L.; Seielstad, M. Patterns of linkage disequilibrium in the human genome. *Nature Reviews. Genetics* 2002, *3*, 299-309, doi:10.1038/nrg777.
449. Slatkin, M.; Theunert, C. Distinguishing Recent Admixture from Ancestral Population Structure. *Genome Biol Evol* 2017, *9*, 427-437, doi:10.1093/gbe/evx018.
450. Marchini, J.; Cardon, L.R.; Phillips, M.S.; Donnelly, P. The effects of human population structure on large genetic association studies. *Nature Genetics* 2004, *36*, 512, doi:10.1038/ng1337https://www.nature.com/articles/ng1337#supplementary-information.
451. Begum, F.; Ghosh, D.; Tseng, G.C.; Feingold, E. Comprehensive literature review and statistical considerations for GWAS meta-analysis. *Nucleic Acids Res* 2012, *40*, 3777-3784, doi:10.1093/nar/gkr1255.
452. Evangelou, E.; Ioannidis, J.P. Meta-analysis methods for genome-wide association studies and beyond. *Nature Reviews. Genetics* 2013, *14*, 379-389, doi:10.1038/nrg3472.
453. Ioannidis, J.P.; Patsopoulos, N.A.; Evangelou, E. Heterogeneity in meta-analyses of genome-wide association investigations. *PloS One* 2007, *2*, e841, doi:10.1371/journal.pone.0000841.
454. McMahon, F.J.; Akula, N.; Schulze, T.G.; Muglia, P.; Tozzi, F.; Detera-Wadleigh, S.D.; Steele, C.J.; Breuer, R.; Strohmaier, J.; Wendland, J.R., et al. Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1. *Nature Genetics* 2010, *42*, 128-131, doi:10.1038/ng.523.
455. Williams, S.M.; Haines, J.L. Correcting away the hidden heritability. *Ann Hum Genet* 2011, *75*, 348-350, doi:10.1111/j.1469-1809.2011.00640.x.
456. Hirschhorn, J.N.; Daly, M.J. Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics* 2005, *6*, 95, doi:10.1038/nrg1521.
457. Schaid, D.J.; Chen, W.; Larson, N.B. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews. Genetics* 2018, *19*, 491-504, doi:10.1038/s41576-018-0016-z.
458. Spain, S.L.; Barrett, J.C. Strategies for fine-mapping complex traits. *Hum Mol Genet* 2015, *24*, R111-119, doi:10.1093/hmg/ddv260.

459. Schubert, M.; Ermini, L.; Der Sarkissian, C.; Jonsson, H.; Ginolhac, A.; Schaefer, R.; Martin, M.D.; Fernandez, R.; Kircher, M. Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat Protoc* 2014, *9*, 1056-1082, doi:10.1038/nprot.2014.063.
460. Rimmer, A.; Phan, H.; Mathieson, I.; Iqbal, Z.; Twigg, S.R.F.; Wilkie, A.O.M.; McVean, G.; Lunter, G. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nature Genetics* 2014, *46*, 912-918, doi:10.1038/ng.3036.
461. Poplin, R.; Ruano-Rubio, V.; DePristo, M.A.; Fennell, T.J.; Carneiro, M.O.; Van der Auwera, G.A.; Kling, D.E.; Gauthier, L.D.; Levy-Moonshine, A.; Roazen, D., et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* 2017, 10.1101/201178, 201178, doi:10.1101/201178.
462. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* 2009, *25*, 2078-2079, doi:10.1093/bioinformatics/btp352.
463. Cingolani, P.; Platts, A.; Wang le, L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* 2012, *6*, 80-92, doi:10.4161/fly.19695.
464. Seboek, D.; Linscheid, P.; Zulewski, H.; Langer, I.; Christ-Crain, M.; Keller, U.; Muller, B. Somatostatin is expressed and secreted by human adipose tissue upon infection and inflammation. *J Clin Endocrinol Metab* 2004, *89*, 4833-4839, doi:10.1210/jc.2004-0271.
465. Faraj, M.; Beauregard, G.; Tardif, A.; Loizon, E.; Godbout, A.; Cianflone, K.; Vidal, H.; Rabasa-Lhoret, R. Regulation of leptin, adiponectin and acylation-stimulating protein by hyperinsulinaemia and hyperglycaemia in vivo in healthy lean young men. *Diabetes Metab* 2008, *34*, 334-342, doi:10.1016/j.diabet.2008.01.014.
466. Rigamonti, A.E.; Sartorio, A.; Bonomo, S.M.; Giunta, M.; Grassi, G.; Perotti, M.; Cella, S.G.; Muller, E.E.; Pincelli, A.I. Effect of a somatostatin infusion on circulating levels of adipokines in obese women. *Metabolism* 2012, *61*, 1797-1802, doi:10.1016/j.metabol.2012.05.019.
467. Jiang, X.; O'Reilly, P.F.; Aschard, H. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat Commun* 2018, *9*, 260, doi:10.1038/s41467-017-02662-2.
468. Alkhatatbeh, M.J.; Abdul-Razzak, K.K.; Khasawneh, L.Q.; Saadeh, N.A. High Prevalence of Vitamin D Deficiency and Correlation of Serum Vitamin D with Cardiovascular Risk in Patients with Metabolic Syndrome. *Metab Syndr Relat Disord* 2017, *15*, 213-219, doi:10.1089/met.2017.0003.
469. Schmitt, E.B.; Nahas-Neto, J.; Bueloni-Dias, F.; Poloni, P.F.; Orsatti, C.L.; Petri Nahas, E.A. Vitamin D deficiency is associated with metabolic syndrome in postmenopausal women. *Maturitas* 2018, *107*, 97-102, doi:10.1016/j.maturitas.2017.10.011.

470. Kramkowska, M.; Grzelak, T.; Walczak, M.; Bogdanski, P.; Pupek-Musialik, D.; Czyzewska, K. Relationship between deficiency of vitamin D and exponents of metabolic syndrome. *Eur Rev Med Pharmacol Sci* 2015, *19*, 2180-2187.
471. Wimalawansa, S.J. Associations of vitamin D with insulin resistance, obesity, type 2 diabetes, and metabolic syndrome. *J Steroid Biochem Mol Biol* 2018, *175*, 177-189, doi:10.1016/j.jsbmb.2016.09.017.
472. Karras, S.N.; Polyzos, S.A.; Newton, D.A.; Wagner, C.L.; Hollis, B.W.; Ouweland, J.v.d.; Dursun, E.; Gezen-Ak, D.; Kotsa, K.; Annweiler, C., et al. Adiponectin and vitamin D-binding protein are independently associated at birth in both mothers and neonates. *Endocrine* 2018, *59*, 164-174, doi:10.1007/s12020-017-1475-2.
473. Adams, J.N.; Raffield, L.M.; Martelle, S.E.; Freedman, B.I.; Langefeld, C.D.; Carr, J.J.; Cox, A.J.; Bowden, D.W. Genetic analysis of advanced glycation end products in the DHS MIND study. *Gene* 2016, *584*, 173-179, doi:10.1016/j.gene.2016.02.029.
474. Santos, A.J.; Nogueira, C.; Ortega-Bellido, M.; Malhotra, V. TANGO1 and Mia2/cTAGE5 (TALI) cooperate to export bulky pre-chylomicrons/VLDLs from the endoplasmic reticulum. *J Cell Biol* 2016, *213*, 343-354, doi:10.1083/jcb.201603072.
475. Hor, H.; Francescatto, L.; Bartesaghi, L.; Ortega-Cubero, S.; Kousi, M.; Lorenzo-Betancor, O.; Jimenez-Jimenez, F.J.; Gironell, A.; Clarimon, J.; Drechsel, O., et al. Missense mutations in TENM4, a regulator of axon guidance and central myelination, cause essential tremor. *Hum Mol Genet* 2015, *24*, 5677-5686, doi:10.1093/hmg/ddv281.
476. Xue, C.B.; Xu, Z.H.; Zhu, J.; Wu, Y.; Zhuang, X.H.; Chen, Q.L.; Wu, C.R.; Hu, J.T.; Zhou, H.S.; Xie, W.H., et al. Exome Sequencing Identifies TENM4 as a Novel Candidate Gene for Schizophrenia in the SCZD2 Locus at 11q14-21. *Front Genet* 2018, *9*, 725, doi:10.3389/fgene.2018.00725.
477. Ngollo, M.; Lebert, A.; Daures, M.; Judes, G.; Rifai, K.; Dubois, L.; Kemeny, J.L.; Penault-Llorca, F.; Bignon, Y.J.; Guy, L., et al. Global analysis of H3K27me3 as an epigenetic marker in prostate cancer progression. *BMC Cancer* 2017, *17*, 261, doi:10.1186/s12885-017-3256-y.
478. Zhang, F.; Xiang, S.; Cao, Y.; Li, M.; Ma, Q.; Liang, H.; Li, H.; Ye, Y.; Zhang, Y.; Jiang, L., et al. EIF3D promotes gallbladder cancer development by stabilizing GRK2 kinase and activating PI3K-AKT signaling pathway. *Cell Death Dis* 2017, *8*, e2868, doi:10.1038/cddis.2017.263.
479. Shah, M.; Su, D.; Scheliga, J.S.; Pluskal, T.; Boronat, S.; Motamedchaboki, K.; Campos, A.R.; Qi, F.; Hidalgo, E.; Yanagida, M., et al. A Transcript-Specific eIF3 Complex Mediates Global Translational Control of Energy Metabolism. *Cell Rep* 2016, *16*, 1891-1902, doi:10.1016/j.celrep.2016.07.006.
480. Huang, X.; Liu, G.; Guo, J.; Su, Z. The PI3K/AKT pathway in obesity and type 2 diabetes. *Int J Biol Sci* 2018, *14*, 1483-1496, doi:10.7150/ijbs.27173.
481. Mayor, F.; Lucas, E.; Jurado-Pueyo, M.; Garcia-Guerra, L.; Nieto-Vazquez, I.; Vila-Bedmar, R.; Fernández-Veledo, S.; Murga, C. G Protein-coupled receptor kinase 2 (GRK2): A novel modulator of insulin resistance. *Arch Physiol Biochem* 2011, *117*, 125-130, doi:10.3109/13813455.2011.584693.

482. Henegar, C.; Tordjman, J.; Achard, V.; Lacasa, D.; Cremer, I.; Guerre-Millo, M.; Poitou, C.; Basdevant, A.; Stich, V.; Viguerie, N., et al. Adipose tissue transcriptomic signature highlights the pathological relevance of extracellular matrix in human obesity. *Genome Biology* 2008, *9*, R14-R14, doi:10.1186/gb-2008-9-1-r14.
483. Lee, H.; Song, J.; Jung, J.H.; Ko, H.W. Primary cilia in energy balance signaling and metabolic disorder. *BMB Reports* 2015, *48*, 647-654, doi:10.5483/bmbrep.2015.48.12.229.
484. Zhou, D.; Ruiter, R.; Zhang, J.; Zhou, M.; Liu, H.; Liu, W.; Wang, S. Angiotensin-converting enzyme I/D polymorphism is not associated with type 2 diabetes in a Chinese population. *J Renin Angiotensin Aldosterone Syst* 2012, *13*, 372-378, doi:10.1177/1470320311435535.
485. Caro-Gomez, M.A.; Naranjo-Gonzalez, C.A.; Gallego-Lopera, N.; Parra-Marin, M.V.; Valencia, D.M.; Arcos, E.G.; Villegas-Perrasse, A.; Bedoya-Berrio, G. Association of Native American ancestry and common variants in ACE, ADIPOR2, MTNR1B, GCK, TCF7L2 and FTO genes with glycemic traits in Colombian population. *Gene* 2018, *677*, 198-210, doi:10.1016/j.gene.2018.07.066.
486. Al-Saikhan, F.I.; Abd-Elaziz, M.A.; Ashour, R.H. Association between risk of type 2 diabetes mellitus and angiotensin-converting enzyme insertion/deletion gene polymorphisms in a Saudi Arabian population. *Biomed Rep* 2017, *7*, 56-60, doi:10.3892/br.2017.920.
487. Ramachandran, V.; Ismail, P.; Stanslas, J.; Shamsudin, N.; Moin, S.; Mohd Jas, R. Association of insertion/deletion polymorphism of angiotensin-converting enzyme gene with essential hypertension and type 2 diabetes mellitus in Malaysian subjects. *J Renin Angiotensin Aldosterone Syst* 2008, *9*, 208-214, doi:10.1177/1470320308097499.
488. Xi, B.; Ruiter, R.; Chen, J.; Pan, H.; Wang, Y.; Mi, J. The ACE insertion/deletion polymorphism and its association with metabolic syndrome. *Metabolism* 2012, *61*, 891-897, doi:10.1016/j.metabol.2011.10.022.
489. Pan, Y.H.; Wang, M.; Huang, Y.M.; Wang, Y.H.; Chen, Y.L.; Geng, L.J.; Zhang, X.X. ACE Gene I/D Polymorphism and Obesity in 1,574 Patients with Type 2 Diabetes Mellitus. *Dis Markers* 2016, *2016*, 7420540, doi:10.1155/2016/7420540.
490. Chmaisse, H.N.; Jammal, M.; Fakhoury, H.; Fakhoury, R. A study on the association between angiotensin-I converting enzyme I/D dimorphism and type-2 diabetes mellitus. *Saudi J Kidney Dis Transpl* 2009, *20*, 1038-1046.
491. Pirozzi, F.F.; Belini Junior, E.; Okumura, J.V.; Salvarani, M.; Bonini-Domingos, C.R.; Ruiz, M.A. The relationship between of ACE I/D and the MTHFR C677T polymorphisms in the pathophysiology of type 2 diabetes mellitus in a population of Brazilian obese patients. *Arch Endocrinol Metab* 2018, *62*, 21-26, doi:10.20945/2359-3997000000005.
492. Lee, D.E.; Brown, J.L.; Rosa, M.E.; Brown, L.A.; Perry, R.A.; Washington, T.A.; Greene, N.P. Translational machinery of mitochondrial mRNA is promoted by physical activity in Western diet-induced obese mice. *Acta Physiol (Oxf)* 2016, *218*, 167-177, doi:10.1111/apha.12687.

493. Albrecht, C.; Viturro, E. The ABCA subfamily--gene and protein structures, functions and associated hereditary diseases. *Pflugers Arch* 2007, *453*, 581-589, doi:10.1007/s00424-006-0047-8.
494. Babashamsi, M.M.; Koukhaloo, S.Z.; Halalkhor, S.; Salimi, A.; Babashamsi, M. ABCA1 and metabolic syndrome; a review of the ABCA1 role in HDL-VLDL production, insulin-glucose homeostasis, inflammation and obesity. *Diabetes Metab Syndr* 2019, *13*, 1529-1534, doi:10.1016/j.dsx.2019.03.004.
495. Brunham, L.R.; Kruit, J.K.; Pape, T.D.; Timmins, J.M.; Reuwer, A.Q.; Vasanji, Z.; Marsh, B.J.; Rodrigues, B.; Johnson, J.D.; Parks, J.S., et al. Beta-cell ABCA1 influences insulin secretion, glucose homeostasis and response to thiazolidinedione treatment. *Nat Med* 2007, *13*, 340-347, doi:10.1038/nm1546.
496. de Haan, W.; Bhattacharjee, A.; Ruddle, P.; Kang, M.H.; Hayden, M.R. ABCA1 in adipocytes regulates adipose tissue lipid content, glucose tolerance, and insulin sensitivity. *J Lipid Res* 2014, *55*, 516-523, doi:10.1194/jlr.M045294.
497. Sanchez-Aguilera, P.; Diaz-Vegas, A.; Campos, C.; Quinteros-Waltemath, O.; Cerda-Kohler, H.; Barrientos, G.; Contreras-Ferrat, A.; Llanos, P. Role of ABCA1 on membrane cholesterol content, insulin-dependent Akt phosphorylation and glucose uptake in adult skeletal muscle fibers from mice. *Biochim Biophys Acta Mol Cell Biol Lipids* 2018, *1863*, 1469-1477, doi:10.1016/j.bbalip.2018.09.005.
498. Iritani, S.; Torii, Y.; Habuchi, C.; Sekiguchi, H.; Fujishiro, H.; Yoshida, M.; Go, Y.; Iriki, A.; Isoda, M.; Ozaki, N. The neuropathological investigation of the brain in a monkey model of autism spectrum disorder with ABCA13 deletion. *Int J Dev Neurosci* 2018, *71*, 130-139, doi:10.1016/j.ijdevneu.2018.09.002.
499. Prades, C.; Arnould, I.; Annilo, T.; Shulenin, S.; Chen, Z.Q.; Orosco, L.; Triunfol, M.; Devaud, C.; Maintoux-Larois, C.; Lafargue, C., et al. The human ATP binding cassette gene ABCA13, located on chromosome 7p12.3, encodes a 5058 amino acid protein with an extracellular domain encoded in part by a 4.8-kb conserved exon. *Cytogenet Genome Res* 2002, *98*, 160-168, doi:10.1159/000069852.
500. Gorokhova, S.; Bibert, S.; Geering, K.; Heintz, N. A novel family of transmembrane proteins interacting with beta subunits of the Na,K-ATPase. *Hum Mol Genet* 2007, *16*, 2394-2410, doi:10.1093/hmg/ddm167.
501. Edwards, A.C.; Bigdeli, T.B.; Docherty, A.R.; Bacanu, S.; Lee, D.; de Candia, T.R.; Moscati, A.; Thiselton, D.L.; Maher, B.S.; Wormley, B.K., et al. Meta-analysis of Positive and Negative Symptoms Reveals Schizophrenia Modifier Genes. *Schizophr Bull* 2016, *42*, 279-287, doi:10.1093/schbul/sbv119.
502. Calboli, F.C.; Tozzi, F.; Galwey, N.W.; Antoniades, A.; Mooser, V.; Preisig, M.; Vollenweider, P.; Waterworth, D.; Waeber, G.; Johnson, M.R., et al. A genome-wide association study of neuroticism in a population-based sample. *PloS One* 2010, *5*, e11504, doi:10.1371/journal.pone.0011504.
503. Zhao, S.C.; Zhou, B.W.; Luo, F.; Mao, X.; Lu, Y.J. The structure and function of NKAIN2-a candidate tumor suppressor. *Int J Clin Exp Med* 2015, *8*, 17072-17079.
504. Mao, X.; Luo, F.; Boyd, L.K.; Zhou, B.; Zhang, Y.; Stankiewicz, E.; Marzec, J.; Vasiljevic, N.; Yu, Y.; Feng, N., et al. NKAIN2 functions as a novel tumor suppressor in prostate cancer. *Oncotarget* 2016, *7*, 63793-63803, doi:10.18632/oncotarget.11690.

505. Yasukochi, Y.; Sakuma, J.; Takeuchi, I.; Kato, K.; Oguri, M.; Fujimaki, T.; Horibe, H.; Yamada, Y. Identification of three genetic variants as novel susceptibility loci for body mass index in a Japanese population. *Physiol Genomics* 2018, *50*, 179-189, doi:10.1152/physiolgenomics.00117.2017.
506. Rudkowska, I.; Perusse, L.; Bellis, C.; Blangero, J.; Despres, J.P.; Bouchard, C.; Vohl, M.C. Interaction between Common Genetic Variants and Total Fat Intake on Low-Density Lipoprotein Peak Particle Diameter: A Genome-Wide Association Study. *J Nutrigenet Nutrigenomics* 2015, *8*, 44-53, doi:10.1159/000431151.
507. Chen, T.; Zhou, L.; Yuan, Y.; Fang, Y.; Guo, Y.; Huang, H.; Zhou, Q.; Lv, X. Characterization of Bbx, a member of a novel subfamily of the HMG-box superfamily together with Cic. *Dev Genes Evol* 2014, *224*, 261-268, doi:10.1007/s00427-014-0476-x.
508. Murugesapillai, D.; McCauley, M.J.; Maher, L.J., 3rd; Williams, M.C. Single-molecule studies of high-mobility group B architectural DNA bending proteins. *Biophysical Reviews* 2016, *9*, 17-40, doi:10.1007/s12551-016-0236-4.
509. Yu, H.; Takeuchi, H. Protein O-glucosylation: another essential role of glucose in biology. *Curr Opin Struct Biol* 2019, *56*, 64-71, doi:10.1016/j.sbi.2018.12.001.
510. Hu, Y.; Tan, L.J.; Chen, X.D.; Liu, Z.; Min, S.S.; Zeng, Q.; Shen, H.; Deng, H.W. Identification of Novel Potentially Pleiotropic Variants Associated With Osteoporosis and Obesity Using the cFDR Method. *J Clin Endocrinol Metab* 2018, *103*, 125-138, doi:10.1210/jc.2017-01531.
511. Wang, W.; Ishibashi, J.; Trefely, S.; Shao, M.; Cowan, A.J.; Sakers, A.; Lim, H.W.; O'Connor, S.; Doan, M.T.; Cohen, P., et al. A PRDM16-Driven Metabolic Signal from Adipocytes Regulates Precursor Cell Fate. *Cell Metab* 2019, *30*, 174-189.e175, doi:10.1016/j.cmet.2019.05.005.
512. Sikder, K.; Shukla, S.K.; Patel, N.; Singh, H.; Rafiq, K. High Fat Diet Upregulates Fatty Acid Oxidation and Ketogenesis via Intervention of PPAR-gamma. *Cell Physiol Biochem* 2018, *48*, 1317-1331, doi:10.1159/000492091.
513. Chen, J.; Meng, Y.; Zhou, J.; Zhuo, M.; Ling, F.; Zhang, Y.; Du, H.; Wang, X. Identifying candidate genes for Type 2 Diabetes Mellitus and obesity through gene expression profiling in multiple tissues or cells. *J Diabetes Res* 2013, *2013*, 970435, doi:10.1155/2013/970435.
514. Bogacka, I.; Xie, H.; Bray, G.A.; Smith, S.R. The effect of pioglitazone on peroxisome proliferator-activated receptor-gamma target genes related to lipid storage in vivo. *Diabetes Care* 2004, *27*, 1660-1667, doi:10.2337/diacare.27.7.1660.
515. Poulain-Godefroy, O.; Lecoeur, C.; Pattou, F.; Fruhbeck, G.; Froguel, P. Inflammation is associated with a decrease of lipogenic factors in omental fat in women. *Am J Physiol Regul Integr Comp Physiol* 2008, *295*, R1-7, doi:10.1152/ajpregu.00926.2007.
516. Rogers, C.; Moukdar, F.; McGee, M.A.; Davis, B.; Buehrer, B.M.; Daniel, K.W.; Collins, S.; Barakat, H.; Robidoux, J. EGF receptor (ERBB1) abundance in adipose tissue is reduced in insulin-resistant and type 2 diabetic women. *J Clin Endocrinol Metab* 2012, *97*, E329-340, doi:10.1210/jc.2011-1033.
517. Lim, G.B. LIMA1 variant influences cholesterol absorption. *Nat Rev Cardiol* 2018, *15*, 502, doi:10.1038/s41569-018-0054-4.

518. Yi, X.; Liu, J.; Wu, P.; Gong, Y.; Xu, X.; Li, W. The key microRNA on lipid droplet formation during adipogenesis from human mesenchymal stem cells. *J Cell Physiol* 2019, 10.1002/jcp.28972, doi:10.1002/jcp.28972.
519. Moreno-Navarrete, J.M.; Rodriguez, A.; Becerril, S.; Valenti, V.; Salvador, J.; Fruhbeck, G.; Fernandez-Real, J.M. Increased Small Intestine Expression of Non-Heme Iron Transporters in Morbidly Obese Patients With Newly Diagnosed Type 2 Diabetes. *Mol Nutr Food Res* 2018, *62*, doi:10.1002/mnfr.201700301.
520. Verkman, A.S. Aquaporins in clinical medicine. *Annu Rev Med* 2012, *63*, 303-316, doi:10.1146/annurev-med-043010-193843.
521. Hara-Chikuma, M.; Sohara, E.; Rai, T.; Ikawa, M.; Okabe, M.; Sasaki, S.; Uchida, S.; Verkman, A.S. Progressive adipocyte hypertrophy in aquaporin-7-deficient mice: adipocyte glycerol permeability as a novel regulator of fat accumulation. *J Biol Chem* 2005, *280*, 15493-15496, doi:10.1074/jbc.C500028200.
522. Maeda, N.; Hibuse, T.; Funahashi, T. Role of aquaporin-7 and aquaporin-9 in glycerol metabolism; involvement in obesity. *Handb Exp Pharmacol* 2009, 10.1007/978-3-540-79885-9_12, 233-249, doi:10.1007/978-3-540-79885-9_12.
523. da Silva, I.V.; Rodrigues, J.S.; Rebelo, I.; Miranda, J.P.G.; Soveral, G. Revisiting the metabolic syndrome: the emerging role of aquaglyceroporins. *Cell Mol Life Sci* 2018, *75*, 1973-1988, doi:10.1007/s00018-018-2781-4.
524. Edwards, S.L.; Beesley, J.; French, J.D.; Dunning, A.M. Beyond GWASs: illuminating the dark road from association to function. *Am J Hum Genet* 2013, *93*, 779-797, doi:10.1016/j.ajhg.2013.10.012.
525. Giral, H.; Landmesser, U.; Kratzer, A. Into the Wild: GWAS Exploration of Non-coding RNAs. *Front Cardiovasc Med* 2018, *5*, 181, doi:10.3389/fcvm.2018.00181.
526. Zhang, F.; Lupski, J.R. Non-coding genetic variants in human disease. *Hum Mol Genet* 2015, *24*, R102-110, doi:10.1093/hmg/ddv259.
527. Zhang, X.; Bailey, S.D.; Lupien, M. Laying a solid foundation for Manhattan--'setting the functional basis for the post-GWAS era'. *Trends Genet* 2014, *30*, 140-149, doi:10.1016/j.tig.2014.02.006.
528. Visscher, P.M.; Wray, N.R.; Zhang, Q.; Sklar, P.; McCarthy, M.I.; Brown, M.A.; Yang, J. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* 2017, *101*, 5-22, doi:10.1016/j.ajhg.2017.06.005.
529. Gallagher, M.D.; Chen-Plotkin, A.S. The Post-GWAS Era: From Association to Function. *Am J Hum Genet* 2018, *102*, 717-730, doi:10.1016/j.ajhg.2018.04.002.
530. Huang, L.; Wang, C.; Rosenberg, N.A. The relationship between imputation error and statistical power in genetic association studies in diverse populations. *Am J Hum Genet* 2009, *85*, 692-698, doi:10.1016/j.ajhg.2009.09.017.
531. Jagannathan, V.; Gerber, V.; Rieder, S. Comprehensive characterization of horse genome variation by whole-genome sequencing of 88 horses. *Anim Genet* 2019, *50*, 74-77, doi:10.1111/age.12753.
532. Daetwyler, H.D.; Capitan, A.; Pausch, H.; Stothard, P.; van Binsbergen, R.; Brondum, R.F.; Liao, X.; Djari, A.; Rodriguez, S.C.; Grohs, C., et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nature Genetics* 2014, *46*, 858-865, doi:10.1038/ng.3034.
533. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A., et al. Finding

the missing heritability of complex diseases. *Nature* 2009, *461*, 747-753, doi:10.1038/nature08494.
534. Stranger, B.E.; Forrest, M.S.; Dunning, M.; Ingle, C.E.; Beazley, C.; Thorne, N.; Redon, R.; Bird, C.P.; de Grassi, A.; Lee, C., et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science (New York, N.Y.)* 2007, *315*, 848-853, doi:10.1126/science.1136678.
535. Clifton, N.E.; Pocklington, A.J.; Scholz, B.; Rees, E.; Walters, J.T.; Kirov, G.; O'Donovan, M.C.; Owen, M.J.; Wilkinson, L.S.; Thomas, K.L., et al. Schizophrenia copy number variants and associative learning. *Mol Psychiatry* 2017, *22*, 178-182, doi:10.1038/mp.2016.227.
536. Gagliano, S.A.; Sengupta, S.; Sidore, C.; Maschio, A.; Cucca, F.; Schlessinger, D.; Abecasis, G.R. Relative impact of indels versus SNPs on complex disease. *Genetic Epidemiology* 2019, *43*, 112-117, doi:10.1002/gepi.22175.
537. Bartonicek, N.; Clark, M.B.; Quek, X.C.; Torpy, J.R.; Pritchard, A.L.; Maag, J.L.V.; Gloss, B.S.; Crawford, J.; Taft, R.J.; Hayward, N.K., et al. Intergenic disease-associated regions are abundant in novel transcripts. *Genome Biology* 2017, *18*, 241, doi:10.1186/s13059-017-1363-3.
538. The Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012, *489*, 57-74, doi:10.1038/nature11247.
539. Giuffra, E.; Tuggle, C.K. Functional Annotation of Animal Genomes (FAANG): Current Achievements and Roadmap. *Annu Rev Anim Biosci* 2019, *7*, 65-88, doi:10.1146/annurev-animal-020518-114913.
540. Degner, J.F.; Pai, A.A.; Pique-Regi, R.; Veyrieras, J.B.; Gaffney, D.J.; Pickrell, J.K.; De Leon, S.; Michelini, K.; Lewellen, N.; Crawford, G.E., et al. DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature* 2012, *482*, 390-394, doi:10.1038/nature10808.
541. Trynka, G.; Sandor, C.; Han, B.; Xu, H.; Stranger, B.E.; Liu, X.S.; Raychaudhuri, S. Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nature Genetics* 2013, *45*, 124-130, doi:10.1038/ng.2504.
542. Vaz-Drago, R.; Custodio, N.; Carmo-Fonseca, M. Deep intronic mutations and human disease. *Human Genetics* 2017, *136*, 1093-1111, doi:10.1007/s00439-017-1809-4.
543. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermuller, J.; Hofacker, I.L., et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science (New York, N.Y.)* 2007, *316*, 1484-1488, doi:10.1126/science.1138341.
544. Rinn, J.L.; Chang, H.Y. Genome Regulation by Long Noncoding RNAs. *Annual Review of Biochemistry* 2012, *81*, 145-166, doi:10.1146/annurev-biochem-051410-092902.
545. Mercer, T.R.; Dinger, M.E.; Mattick, J.S. Long non-coding RNAs: insights into functions. *Nature Reviews Genetics* 2009, *10*, 155, doi:10.1038/nrg2521.
546. Sun, M.; Kraus, W.L. From discovery to function: the expanding roles of long noncoding RNAs in physiology and disease. *Endocrine Reviews* 2015, *36*, 25-64, doi:10.1210/er.2014-1034.

547. Macintyre, G.; Jimeno Yepes, A.; Ong, C.S.; Verspoor, K. Associating disease-related genetic variants in intergenic regions to the genes they impact. *PeerJ* 2014, *2*, e639, doi:10.7717/peerj.639.
548. Batty, G.D.; Barzi, F.; Woodward, M.; Jamrozik, K.; Woo, J.; Kim, H.C.; Ueshima, H.; Huxley, R.R.; Asia Pacific Cohort Studies, C. Adult height and cancer mortality in Asia: the Asia Pacific Cohort Studies Collaboration. *Ann Oncol* 2010, *21*, 646-654, doi:10.1093/annonc/mdp363.
549. La Vecchia, C.; Decarli, A.; Negri, E.; Ferraroni, M.; Pagano, R. Height and the prevalence of chronic disease. *Revue d'epidemiologie et de sante publique* 1992, *40*, 6-14.
550. Perelman, J. Are chronic diseases related to height? Results from the Portuguese National Health Interview Survey. *Econ Hum Biol* 2014, *15*, 56-66, doi:10.1016/j.ehb.2014.06.001.
551. Lawlor, D.A.; Ebrahim, S.; Davey Smith, G. The association between components of adult height and Type II diabetes and insulin resistance: British Women's Heart and Health Study. *Diabetologia* 2002, *45*, 1097-1106, doi:10.1007/s00125-002-0887-5.
552. Smith, G.D.; Greenwood, R.; Gunnell, D.; Sweetnam, P.; Yarnell, J.; Elwood, P. Leg length, insulin resistance, and coronary heart disease risk: the Caerphilly Study. *J Epidemiol Community Health* 2001, *55*, 867-872.
553. Guerrero-Igea, F.J.; Lepe-Jimenez, J.A.; Garrido-Serrano, A.; Palomo-Gil, S. Association among hyperinsulinemia, family history of diabetes, and diminutive stature in normoglycemic premenopausal women. *Diabetes care* 2001, *24*, 602-603.
554. Silva, E.C.; Martins, I.S.; de Araujo, E.A. [Metabolic syndrome and short stature in adults from the metropolitan area of Sao Paulo city (SP, Brazil)]. *Ciencia & saude coletiva* 2011, *16*, 663-668.
555. Bozorgmanesh, M.; Hadaegh, F.; Zabetian, A.; Azizi, F. Impact of hip circumference and height on incident diabetes: results from 6-year follow-up in the Tehran Lipid and Glucose Study. *Diabet Med* 2011, *28*, 1330-1336, doi:10.1111/j.1464-5491.2011.03343.x.
556. Janghorbani, M.; Amini, M. Associations of hip circumference and height with incidence of type 2 diabetes: the Isfahan diabetes prevention study. *Acta Diabetologica* 2012, *49 Suppl 1*, S107-114, doi:10.1007/s00592-011-0351-4.
557. Asao, K.; Kao, W.H.; Baptiste-Roberts, K.; Bandeen-Roche, K.; Erlinger, T.P.; Brancati, F.L. Short stature and the risk of adiposity, insulin resistance, and type 2 diabetes in middle age: the Third National Health and Nutrition Examination Survey (NHANES III), 1988-1994. *Diabetes Care* 2006, *29*, 1632-1637, doi:10.2337/dc05-1997.
558. Brown, D.C.; Byrne, C.D.; Clark, P.M.; Cox, B.D.; Day, N.E.; Hales, C.N.; Shackleton, J.R.; Wang, T.W.; Williams, D.R. Height and glucose tolerance in adult subjects. *Diabetologia* 1991, *34*, 531-533.
559. Nuesch, E.; Dale, C.; Palmer, T.M.; White, J.; Keating, B.J.; van Iperen, E.P.; Goel, A.; Padmanabhan, S.; Asselbergs, F.W.; Investigators, E.P.-N., et al. Adult height, coronary heart disease and stroke: a multi-locus Mendelian randomization meta-analysis. *Int J Epidemiol* 2016, *45*, 1927-1937, doi:10.1093/ije/dyv074.

560. Bosy-Westphal, A.; Plachta-Danielzik, S.; Dorhofer, R.P.; Muller, M.J. Short stature and obesity: positive association in adults but inverse association in children and adolescents. *Br J Nutr* 2009, *102*, 453-461, doi:10.1017/S0007114508190304.
561. La Batide-Alanore, A.; Tregouet, D.A.; Sass, C.; Siest, G.; Visvikis, S.; Tiret, L. Family study of the relationship between height and cardiovascular risk factors in the STANISLAS cohort. *Int J Epidemiol* 2003, *32*, 607-614.
562. Olatunbosun, S.T.; Bella, A.F. Relationship between height, glucose intolerance, and hypertension in an urban African black adult population: a case for the "thrifty phenotype" hypothesis? *J Natl Med Assoc* 2000, *92*, 265-268.
563. Nelson, C.P.; Hamby, S.E.; Saleheen, D.; Hopewell, J.C.; Zeng, L.; Assimes, T.L.; Kanoni, S.; Willenborg, C.; Burgess, S.; Amouyel, P., et al. Genetically determined height and coronary artery disease. *N Engl J Med* 2015, *372*, 1608-1618, doi:10.1056/NEJMoa1404881.
564. Barker, D.J. The intrauterine origins of cardiovascular and obstructive lung disease in adult life. The Marc Daniels Lecture 1990. *J R Coll Physicians Lond* 1991, *25*, 129-133.
565. Palmer, J.R.; Rosenberg, L.; Shapiro, S. Stature and the risk of myocardial infarction in women. *Am J Epidemiol* 1990, *132*, 27-32.
566. Liu, G.; Liu, J.; Li, N.; Tang, Z.; Lan, F.; Pan, L.; Yang, X.; Hu, G.; Yu, Z. Association between leg length-to-height ratio and metabolic syndrome in Chinese children aged 3 to 6 years. *Prev Med Rep* 2014, *1*, 62-67, doi:10.1016/j.pmedr.2014.11.002.
567. Meyre, D.; Boutin, P.; Tounian, A.; Deweirder, M.; Aout, M.; Jouret, B.; Heude, B.; Weill, J.; Tauber, M.; Tounian, P., et al. Is glutamate decarboxylase 2 (GAD2) a genetic link between low birth weight and subsequent development of obesity in children? *J Clin Endocrinol Metab*2005, *90*, 2384-2390, doi:10.1210/jc.2004-1468.
568. Wegner, L.; Andersen, G.; Sparso, T.; Grarup, N.; Glumer, C.; Borch-Johnsen, K.; Jorgensen, T.; Hansen, T.; Pedersen, O. Common variation in LMNA increases susceptibility to type 2 diabetes and associates with elevated fasting glycemia and estimates of body fat and height in the general population: studies of 7,495 Danish whites. *Diabetes* 2007, *56*, 694-698, doi:10.2337/db06-0927.
569. McCoy, A.M.; McCue, M.E. Validation of imputation between equine genotyping arrays. *Anim Genet* 2013, 10.1111/age.12093, doi:10.1111/age.12093.
570. Akey, J.M.; Ruhe, A.L.; Akey, D.T.; Wong, A.K.; Connelly, C.F.; Madeoy, J.; Nicholas, T.J.; Neff, M.W. Tracking footprints of artificial selection in the dog genome. *Proc Natl Acad Sci USA* 2010, *107*, 1160-1165, doi:10.1073/pnas.0909918107.
571. Gusev, A.; Lee, S.H.; Trynka, G.; Finucane, H.; Vilhjalmsson, B.J.; Xu, H.; Zang, C.; Ripke, S.; Bulik-Sullivan, B.; Stahl, E., et al. Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *Am J Hum Genet* 2014, *95*, 535-552, doi:10.1016/j.ajhg.2014.10.004.
572. Xu, H.; Guan, Y. Detecting local haplotype sharing and haplotype association. *Genetics* 2014, *197*, 823-838, doi:10.1534/genetics.114.164814.
573. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3--new capabilities and interfaces. *Nucleic Acids Res* 2012, *40*, e115, doi:10.1093/nar/gks596.

574. Frischknecht, M.; Jagannathan, V.; Plattet, P.; Neuditschko, M.; Signer-Hasler, H.; Bachmann, I.; Pacholewska, A.; Drogemuller, C.; Dietschi, E.; Flury, C., et al. A Non-Synonymous HMGA2 Variant Decreases Height in Shetland Ponies and Other Small Horses. *PloS One* 2015, *10*, e0140749, doi:10.1371/journal.pone.0140749.
575. Woodward, J. Bi-allelic SNP genotyping using the TaqMan(R) assay. *Methods Mol Biol* 2014, *1145*, 67-74, doi:10.1007/978-1-4939-0446-4_6.
576. Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D.; Team, R.C. *nlme: Linear and Nonlinear Mixed Effects Models*, R package version 3.1-131: 2017.
577. Lenth, R.V. Least-Squares Means: The R Package lsmeans. *J Stat Softw* 2016, *69*, 33, doi:10.18637/jss.v069.i01.
578. Gonzalez, J.R.; Armengol, L.; Sole, X.; Guino, E.; Mercader, J.M.; Estivill, X.; Moreno, V. SNPassoc: an R package to perform whole genome association studies. *Bioinformatics* 2007, *23*, 644-645, doi:10.1093/bioinformatics/btm025.
579. Norton, E.M.; Schultz, N.E.; Rendahl, A.K.; Geor, R.J.; Mickelson, J.R.; McCue, M.E. Heritability of metabolic traits associated with equine metabolic syndrome in Welsh ponies and Morgan horses. *Equine Vet J* 2019, 51, 475-480.
580. Feng, S.; Wang, S.; Chen, C.C.; Lan, L. GWAPower: a statistical power calculation software for genome-wide association studies with quantitative traits. *BMC Genetics* 2011, *12*, 12, doi:10.1186/1471-2156-12-12.
581. Cleynen, I.; Van de Ven, W.J. The HMGA proteins: a myriad of functions (Review). *Int J Oncol* 2008, *32*, 289-305.
582. Lettre, G.; Jackson, A.U.; Gieger, C.; Schumacher, F.R.; Berndt, S.I.; Sanna, S.; Eyheramendy, S.; Voight, B.F.; Butler, J.L.; Guiducci, C., et al. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nature Genetics* 2008, *40*, 584-591, doi:10.1038/ng.125.
583. Weedon, M.N.; Lango, H.; Lindgren, C.M.; Wallace, C.; Evans, D.M.; Mangino, M.; Freathy, R.M.; Perry, J.R.; Stevens, S.; Hall, A.S., et al. Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genetics* 2008, *40*, 575-583, doi:10.1038/ng.121.
584. Assessment of Insulin and Glucose Dynamics by Using an Oral Sugar Test in Horses Weedon, M.N.; Lettre, G.; Freathy, R.M.; Lindgren, C.M.; Voight, B.F.; Perry, J.R.B.; Elliott, K.S.; Hackett, R.; Guiducci, C.; Shields, B., et al. A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nature Genetics* 2007, *39*, 1245-1250, doi:10.1038/ng2121.
585. Yang, T.L.; Guo, Y.; Zhang, L.S.; Tian, Q.; Yan, H.; Guo, Y.F.; Deng, H.W. HMGA2 is confirmed to be associated with human adult height. *Ann Hum Genet* 2010, *74*, 11-16, doi:10.1111/j.1469-1809.2009.00555.x.
586. Gudbjartsson, D.F.; Walters, G.B.; Thorleifsson, G.; Stefansson, H.; Halldorsson, B.V.; Zusmanovich, P.; Sulem, P.; Thorlacius, S.; Gylfason, A.; Steinberg, S., et al. Many sequence variants affecting diversity of adult human height. *Nature Genetics* 2008, *40*, 609-615, doi:10.1038/ng.122.
587. Lango Allen, H.; Estrada, K.; Lettre, G.; Berndt, S.I.; Weedon, M.N.; Rivadeneira, F.; Willer, C.J.; Jackson, A.U.; Vedantam, S.; Raychaudhuri, S., et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010, *467*, 832-838, doi:10.1038/nature09410.

588. Lanktree, M.B.; Guo, Y.; Murtaza, M.; Glessner, J.T.; Bailey, S.D.; Onland-Moret, N.C.; Lettre, G.; Ongen, H.; Rajagopalan, R.; Johnson, T., et al. Meta-analysis of Dense Genecentric Association Studies Reveals Common and Uncommon Variants Associated with Height. *Am J Hum Genet* 2011, *88*, 6-18, doi:10.1016/j.ajhg.2010.11.007.
589. Rimbault, M.; Beale, H.C.; Schoenebeck, J.J.; Hoopes, B.C.; Allen, J.J.; Kilroy-Glynn, P.; Wayne, R.K.; Sutter, N.B.; Ostrander, E.A. Derived variants at six genes explain nearly half of size reduction in dog breeds. *Genome Research* 2013, *23*, 1985-1995, doi:10.1101/gr.157339.113.
590. Zhou, X.; Benson, K.F.; Ashar, H.R.; Chada, K. Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature* 1995, *376*, 771-774, doi:10.1038/376771a0.
591. Battista, S.; Fidanza, V.; Fedele, M.; Klein-Szanto, A.J.; Outwater, E.; Brunner, H.; Santoro, M.; Croce, C.M.; Fusco, A. The expression of a truncated HMGI-C gene induces gigantism associated with lipomatosis. *Cancer Research* 1999, *59*, 4793-4797.
592. Ligon, A.H.; Moore, S.D.; Parisi, M.A.; Mealiffe, M.E.; Harris, D.J.; Ferguson, H.L.; Quade, B.J.; Morton, C.C. Constitutional rearrangement of the architectural factor HMGA2: a novel human phenotype including overgrowth and lipomas. *Am J Hum Genet* 2005, *76*, 340-348, doi:10.1086/427565.
593. Voight, B.F.; Scott, L.J.; Steinthorsdottir, V.; Morris, A.P.; Dina, C.; Welch, R.P.; Zeggini, E.; Huth, C.; Aulchenko, Y.S.; Thorleifsson, G., et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nature Genetics* 2010, *42*, 579-589, doi:10.1038/ng.609.
594. Bai, H.; Liu, H.; Suyalatu, S.; Guo, X.; Chu, S.; Chen, Y.; Lan, T.; Borjigin, B.; Orlov, Y.L.; Posukh, O.L., et al. Association Analysis of Genetic Variants with Type 2 Diabetes in a Mongolian Population in China. *J Diabetes Res* 2015, *2015*, 613236, doi:10.1155/2015/613236.
595. Ng, M.C.; Shriner, D.; Chen, B.H.; Li, J.; Chen, W.M.; Guo, X.; Liu, J.; Bielinski, S.J.; Yanek, L.R.; Nalls, M.A., et al. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS Genetics* 2014, *10*, e1004517, doi:10.1371/journal.pgen.1004517.
596. Saxena, R.; Elbers, C.C.; Guo, Y.; Peter, I.; Gaunt, T.R.; Mega, J.L.; Lanktree, M.B.; Tare, A.; Castillo, B.A.; Li, Y.R., et al. Large-scale gene-centric meta-analysis across 39 studies identifies type 2 diabetes loci. *Am J Hum Genet* 2012, *90*, 410-425, doi:10.1016/j.ajhg.2011.12.022.
597. Hulsmans, M.; Geeraert, B.; De Keyzer, D.; Mertens, A.; Lannoo, M.; Vanaudenaerde, B.; Hoylaerts, M.; Benhabiles, N.; Tsatsanis, C.; Mathieu, C., et al. Interleukin-1 receptor-associated kinase-3 is a key inhibitor of inflammation in obesity and metabolic syndrome. *PloS One* 2012, *7*, e30414, doi:10.1371/journal.pone.0030414.
598. Li, H.; Cuartas, E.; Cui, W.; Choi, Y.; Crawford, T.D.; Ke, H.Z.; Kobayashi, K.S.; Flavell, R.A.; Vignery, A. IL-1 receptor-associated kinase M is a central regulator of osteoclast differentiation and activation. *J Exp Med* 2005, *201*, 1169-1177, doi:10.1084/jem.20041444.

599. Tan, Q.; Majewska-Szczepanik, M.; Zhang, X.; Szczepanik, M.; Zhou, Z.; Wong, F.S.; Wen, L. IRAK-M deficiency promotes the development of type 1 diabetes in NOD mice. *Diabetes* 2014, *63*, 2761-2775, doi:10.2337/db13-1504.
600. Kettunen, J.; Demirkan, A.; Würtz, P.; Draisma, H.H.M.; Haller, T.; Rawal, R.; Vaarhorst, A.; Kangas, A.J.; Lyytikäinen, L.-P.; Pirinen, M., et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nat Commun* 2016, *7*, 11122, doi:10.1038/ncomms11122
601. Adamski, J. Genome-wide association studies with metabolomics. *Genome Medicine* 2012, *4*, 34-34, doi:10.1186/gm333.
602. Crawford NF, Thomas MG, Holt TN, Speidel SE, Enns RM. Heritabilities and genetic correlations of pulmonary arterial pressure and performance traits in Angus cattle at high altitude. *J Anim Sci*. 2016;94(11):4483-90. Epub 2016/11/30. doi: 10.2527/jas.2016-0703.
603. Lupi TM, Leon JM, Nogales S, Barba C, Delgado JV. Genetic parameters of traits associated with the growth curve in Segurena sheep. *Animal* 2016;10(5):729-35.
604. Ahlgren J, Uimari P. Heritability of hypothyroidism in the Finnish Hovawart population. *Acta Veterinaria Scandinavica*. 2016;58(1):39. doi: 10.1186/s13028-016-0221-8.
605. Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, de Andrade M, Feenstra B, Feingold E, Hayes MG, Hill WG, Landi MT, Alonso A, Lettre G, Lin P, Ling H, Lowe W, Mathias RA, Melbye M, Pugh E, Cornelis MC, Weir BS, Goddard ME, Visscher PM. Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics*. 2011;43(6):519-25. doi: 10.1038/ng.823.

Appendix A: Chapter 2 Supplemental Materials

Supplemental Methods

Phenotype Data: Information collected on all individuals included morphometric, environmental and biochemical measurements. Morphometric measurements included body condition score, height, length, neck circumference, and girth circumference. Environmental management included diet (amount and type of hay, grain, pelleted feed, supplements, hours on pasture per day, and complete dietary analysis) and exercise regimen (daily hours in a stall, turnout and forced exercise). Biochemical measurements included insulin (INS) and glucose (GLU) after an 8 hour fast and 75 minutes post oral sugar challenge (OST) using 0.15ml/kg light Karo Syrup as previously described [406]. Additional samples taken at baseline included: triglycerides (TG), non-esterified fatty acids (NEFA), adiponectin, leptin and ACTH.

Biochemical Measurements	Assay
Insulin and INS-OST	Siemen's TKIN1 Insulin Coat-A-Count Kit
Glucose and GLU-OST	YSI 2300 STAT Plus glucose and lactate analyzer
Triglycerides	Millipore Sigma's TR0100 Serum Triglyceride Determination kit
NEFA	Wako Diagnostics' HR Series NEFA kit
Adiponectin	Millipore Sigma's EZHMWA-64K Human High Molecular Weight Adiponectin ELISA
Leptin	Millipore Sigma's XL-85K Multi-Species Leptin RIA
ACTH	Siemen's LKAC1 ACTH kits

Horses with a history of laminitis were defined as individuals that had previously been diagnosed with laminitis by a veterinarian, had radiographic changes consistent with laminitis, or had laminitic rings at the time of sampling. Horses that had signs of acute, active laminitis (i.e. pain, lameness) at the time of sampling were excluded from the study. Horses with a history of laminitis that could be contributed to an inciting factor other the EMS (such as horses with a history of endotoxemia, grain overload, trauma, support limb

laminitis, corticosteroid administration, or pars pituitary intermedia dysfunction) were excluded from the study. A total of 58 WP and 43 Morgan horses had a history of laminitis.

Horses with a previous diagnosis of EMS were defined as individuals that had previously been diagnosed with EMS by a veterinarian, although diagnostic criteria varied between veterinarians. For inclusion in the study, each farm had to have at least one control and one horse with clinical signs consistent with EMS under the same management. Clinical signs consistent with EMS included horses with regional adiposity, hyperinsulinemia or an exaggerated response to the OST, elevations in TG, and decreased levels of adiponectin at the time of sampling. Horses with a history or phenotypic appearance of pars pituitary intermedia dysfunction (hirsutism, polyuria/polydipsia, muscle wasting) or elevated ACTH values (based on seasonal reference ranges) were excluded from the study.

Heritability Estimates Using Different Methods: In the main text, we reported h^2_{SNP} estimates calculated in GCTA with the inclusion of the wGRM. A REML based approach was chosen due to its published use estimating h^2_{SNP} in hundreds of traits [413], versatility for use with domestic animal data [602-604], and evidence showing that it is more powerful when analyzing quantitative data as compared to a Haseman-Elston regression [405]. In addition to GCTA and the wGRM, heritability estimates were also calculated with GCTA and the standard GRM (computed in GCTA), as well as using LDAK's REML algorithm with the wGRM. For GCTA, the default settings were used for analysis. For LDAK, the addition of --decay YES and --half-life 100 was included in constructing the wGRM to account for the high structure and LD in horse genotype data. We also used an alpha value of -0.25 (--power -0.25) as recommended by Speed et al to allow the average h^2_{SNP} to vary with MAF [413].

Assessment of Bias Due to Population Substructure: To determine if population substructure was inflating the h^2_{SNP} estimates, we split the genome into two groups of approximately equivalent numbers of SNPs to see if individuals shared more of their genome than what would be expected by chance [605]. The first group contained SNPs from autosomes 1 through 12, with a total of 723,378 SNPs in the WP and 588,093 SNPs in the Morgan. The second group contained SNPs from autosomes 13 through 31, with a total of 704,959 SNPs in the WP and 570,738 SNPs in the Morgan. For each trait, separate

h^2_{SNP} estimates were calculated for each group using GCTA with a wGRM comprised only of the SNPs within their respective groups. The difference between the summed estimates of both groups and the total h^2_{SNP} estimate from inclusion of all SNPs was calculated to determine if population substructure was causing inflation of the h^2_{SNP} estimate. Based on the high LD in horses, we set a cutoff of a difference greater than 0.05 for indication of population substructure and inflation of the h^2_{SNP} estimates.

Results:

Heritability Estimates Using Different Methods: In the main text, we reported h^2_{SNP} estimates calculated in GCTA with the inclusion of the wGRM. However, we also compared three methods: GCTA with standard GRM, GCTA with the wGRM, and LDAK with the wGRM. The results of all three methods are shown in **Supplementary Table A3**. In general, the three methods produced very similar estimates of h^2_{SNP}. The exception was ACTH in the Morgans; however, this estimate was still with the range of the SE.

Assessment of Bias Due to Population Substructure: We tested for the presence of population substructure by comparing the difference between the summed h^2_{SNP} estimates obtained from splitting the genome into two sections and the original estimate from the full data set. Higher differences between the two calculations indicate that individuals within the population are more genetically similar across chromosomes than what is expected by chance. Across both breeds, the differences between the summed values for the split SNP dataset and the h^2_{SNP} estimates using the full SNP data set ranged from 0.001-0.095 (**Supplemental Table A5**). Based on our cutoff, both adiponectin (summed h^2_{SNP} of 0.818 with a difference of 0.095) and NEFA (summed h^2_{SNP} of 0.831 with a difference of 0.085) in the Morgans had a difference that suggested potential inflation of the h^2_{SNP} estimates. The remaining 16 models (of 18 models, i.e., 9 traits in each breed) had differences less than 0.05, indicating that inflation due to population substructure was unlikely.

	Glucose	GLU-OST	Insulin	INS-OST	NEFA	TG	Leptin	ADIPON	ACTH
Welsh ponies									
No Covariates	2022.83	2174.85	194.06	301.86	-191.55	326.78	675.08	2567.95	-20.13
Age	2024.69	2176.48	188.45	290.59	-193.72	327.44	677.06	2565.70	-51.97
Sex	2024.59	2176.34	187.21	296.74	-189.99	313.62	669.05	2554.63	-18.82
Season	**1976.98**	**2161.71**	196.85	287.41	-237.78	287.45	674.54	2561.84	-51.36
Age and Sex	2026.36	2177.76	**184.11**	288.38	-191.80	311.60	670.88	2554.74	-50.04
Age and Season	1978.95	2163.51	191.10	276.82	**-240.37**	287.68	676.54	2559.60	**-94.16**
Sex and Season	1978.36	2162.90	190.38	282.51	-236.28	276.59	**667.51**	**2550.46**	-49.41
Age, Sex, Season	1980.36	2164.50	187.15	**274.73**	-238.45	**274.06**	669.17	2550.48	-93.46
Morgan horses									
No Covariates	2078.92	2284.92	235.73	326.86	**-145.08**	-24.92	681.06	585.13	-65.13
Age	2078.37	2286.83	230.37	323.71	-143.15	-23.57	682.66	584.61	-100.03
Sex	2078.49	2282.70	237.65	328.58	**-145.77**	-27.39	673.77	586.62	-64.50
Season	2075.88	2260.96	201.09	296.53	-140.99	-31.93	649.68	582.34	-128.28
Age and Sex	2078.17	2284.65	232.19	325.25	-143.81	-26.22	675.54	586.21	-100.54
Age and Season	2076.33	2262.96	**196.17**	**292.75**	-139.10	-31.33	651.27	**581.71**	**-172.37**
Sex and Season	**2075.29**	**2260.69**	202.31	298.24	-141.37	**-38.29**	**634.74**	582.75	-126.29
Age, Sex, Season	2075.95	2262.67	197.62	294.61	-139.43	-38.18	636.60	582.33	-170.68

Supplemental Table A1: Table of Akaike information criterion (AIC) values obtained after fitting the covariates to a linear regression model. AIC values were obtained for each of the nine biochemical traits and a combination of four potential confounding covariates in each breed. Values in red were the minimal AIC values for the respective column. Covariates form the models with the minimum AIC values were used in subsequent SNP-based heritability estimates. Abbreviations: GLU-OST: glucose post oral sugar test, INS-OST: insulin post oral sugar test, NEFA: non-esterified fatty acids, TG: triglycerides, ADIPON: adiponectin.

	Welsh Ponies		Morgan Horses	
Rank	Unadjusted P-Values	Adjusted P-Values	Unadjusted P-values	Adjusted P-Values
1	1.36e-13	**<0.001**	1.66e-07	**<0.001**
2	3.14e-07	**<0.001**	1.94e-07	**<0.001**
3	1.75e-05	**<0.001**	9.50e-05	**<0.001**
4	1.60e-04	**<0.001**	1.16e-04	**<0.001**
5	3.95e-04	**<0.001**	7.77e-04	**<0.001**
6	4.63e-04	**<0.001**	1.16e-02	**0.046**
7	7.75e-03	**0.02**	1.78e-02	0.053
8	2.70e-02	0.054	2.61e-02	0.053
9	3.45e-02	0.054	5.68e-02	0.057

Supplemental Table A2: Unadjusted and Holm-Bonferroni adjusted p-values for heritability (h^2_{SNP}) estimates in Welsh ponies and Morgan horses. The unadjusted p-values from the h^2_{SNP} estimates in the Welsh ponies and Morgan horses were ranked from lowest to highest. Holm-Bonferroni adjusted p-values were calculated as follows: P-value*(9-rank+1). Adjusted p-values that were significant (those in bold) were less than the 0.05. Adjusted p-values which are shown at two significant figures are those where the threshold for <0.05 cutoff required clarification from rounding and were represented in the manuscript as **<0.05**.

	Glucose	GLU-OST	Insulin	INS-OST	NEFA	TG	Leptin	ADIPON	ACTH
Welsh ponies									
GCTA GRM									
h^2_{SNP} estimate	0.402	0.225	0.778	0.447	0.434	0.303	0.554	0.465	0.300
SE	0.129	0.136	0.106	0.137	0.136	0.139	0.132	0.138	0.149
P-Value	**<0.001**	**<0.05**	**<0.001**	**<0.001**	**<0.001**	**0.02**	**<0.001**	**<0.001**	**<0.05**
GCTA wGRM									
h^2_{SNP} estimate	0.408	0.226	0.808	0.440	0.434	0.313	0.554	0.488	0.305
SE	0.135	0.142	0.108	0.148	0.136	0.146	0.132	0.143	0.154
P-Value	**<0.001**	0.05	**<0.001**	**<0.001**	**<0.001**	**0.02**	**<0.001**	**<0.001**	0.05
LDAK wGRM									
h^2_{SNP} estimate	0.408	0.226	0.808	0.428	0.467	0.316	0.573	0.488	0.303
SE	0.135	0.142	0.108	0.147	0.141	0.145	0.138	0.143	0.154
P-Value	**<0.001**	0.05	**<0.001**	**<0.001**	**<0.001**	**0.02**	**<0.001**	**<0.001**	0.05
Morgan horses									
GCTA GRM									
h^2_{SNP} estimate	0.174	0.510	0.547	0.331	0.684	0.242	0.442	0.841	0.316
SE	0.153	0.155	0.170	0.164	0.164	0.157	0.155	0.156	0.193
P-Value	0.2	**<0.001**	**<0.001**	**0.04**	**<0.001**	0.08	**<0.001**	**<0.001**	0.2
GCTA wGRM									
h^2_{SNP} estimate	0.208	0.565	0.592	0.359	0.746	0.273	0.486	0.913	0.408
SE	0.172	0.175	0.195	0.185	0.188	0.176	0.177	0.181	0.215
P-Value	0.05	**<0.001**	**<0.001**	**<0.05**	**<0.001**	0.05	**<0.001**	**<0.001**	0.06
LDAK wGRM									
h^2_{SNP} estimate	0.208	0.565	0.592	0.359	0.746	0.273	0.486	0.913	0.408
SE	0.172	0.175	0.195	0.185	0.188	0.176	0.177	0.181	0.215
P-Value	0.05	**<0.001**	**<0.001**	**<0.05**	**<0.001**	0.05	**<0.001**	**<0.001**	0.06

Supplemental Table A3: Summary table of heritability (h^2_{SNP}) estimates using three methods. Table presents h^2_{SNP}, standard error (SE) and p-values for nine biochemical traits in both Morgan horses and Welsh ponies using GCTA with the standard GRM, GCTA with the linkage disequilibrium corrected GRM (wGRM) and LDAK with the wGRM. Covariates included in the model were based on AIC values. P-values are adjusted by a Holm-Bonferroni correction, those in bold were <0.05, and those in red had an unadjusted p-value of >0.05. Abbreviations: GLU-OST: glucose post oral sugar test, INS-OST: insulin post oral sugar test, NEFA: non-esterified fatty acids, TG: triglycerides, ADIPON: adiponectin

	Glucose	GLU-OST	Insulin	INS-OST	NEFA	TG	Leptin	ADIPON	ACTH
Welsh ponies									
Age and Sex									
h^2_{SNP} estimate	0.537	0.300	0.808	0.463	0.560	0.445	0.477	0.454	0.510
SE	0.127	0.137	0.108	0.143	0.116	0.124	0.139	0.135	0.129
P-Value	**<0.001**	**0.004**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**
Age, Sex, Farm									
h^2_{SNP} estimate	0.229	0.00	0.367	0.125	0.160	0.667	0.363	0.099	0.274
SE	0.161	0.157	0.175	0.180	0.162	0.154	0.162	0.178	0.178
P-Value	0.4	>0.9	0.1	>0.9	0.8	>0.9	0.05	>0.9	0.4
Age, Sex, Season									
h^2_{SNP} estimate	0.427	0.241	0.827	0.440	0.466	0.313	0.612	0.488	0.298
SE	0.136	0.145	0.106	0.148	0.142	0.146	0.136	0.143	0.154
P-Value	**0.001**	0.06	**<0.001**	**0.002**	**<0.001**	**0.02**	**<0.001**	**<0.001**	0.06
Morgan horses									
Age and Sex									
h^2_{SNP} estimate	0.285	0.634	0.776	0.463	0.743	0.391	0.671	0.920	0.575
SE	0.163	0.179	0.180	0.186	0.189	0.180	0.173	0.181	0.213
P-Value	**0.01**	**<0.001**	**<0.001**	**0.006**	**<0.001**	**0.009**	**<0.001**	**<0.001**	**0.009**
Age, Sex, Farm									
h^2_{SNP} estimate	0.00	0.198	0.574	0.432	0.00	0.212	0.199	0.992	0.666
SE	0.253	0.227	0.225	0.212	0.248	0.242	0.228	0.198	0.225
P-Value	>0.9	0.9	**0.03**	0.05	>0.9	0.9	0.9	**0.001**	**0.03**
Age, Sex, Season									
h^2_{SNP} estimate	0.219	0.568	0.589	0.459	0.750	0.293	0.485	0.916	0.413
SE	0.173	0.175	0.195	0.185	0.188	0.177	0.173	0.182	0.216
P-Value	0.1	**<0.001**	**0.004**	**<0.05**	**<0.001**	0.05	**<0.001**	**<0.001**	0.1

Supplemental Table A4: Comparison of heritability (h^2_{SNP}) estimates of nine biochemical traits in the Welsh ponies and Morgans with the inclusion of different covariates. Heritability estimates were performed in GCTA with the linkage disequilibrium corrected GRM (wGRM). P-values are adjusted by a Holm-Bonferroni correction bolded values were <0.05 after correction. P-values in red are those which the unadjusted p-value was >0.05. Abbreviations: GLU-OST: glucose post oral sugar test, INS-OST: insulin post oral sugar test, NEFA: non-esterified fatty acids, TG: triglycerides, ADIPON: adiponectin, SE: standard error.

	GLU	GLU-OST	INS	INS-OST	NEFA	TG	Leptin	ADIPON	ACTH
Welsh ponies									
Sum h^2_{SNP} estimates	0.417	0.229	0.817	0.456	0.478	0.323	0.578	0.510	0.298
Section1 h^2_{SNP}	0.264	0.165	0.282	0.00	0.108	0.106	0.177	0.00	0.278
Section 2 h^2_{SNP}	0.153	0.064	0.536	0.456	0.370	0.217	0.402	0.510	0.020
Diff h^2_{SNP} estimates	0.009	0.003	0.009	0.016	0.044	0.010	0.024	0.022	0.007
Morgan horses									
Sum h^2_{SNP} estimates	0.215	0.611	0.603	0.373	0.831	0.274	0.502	0.818	0.400
Section1 h^2_{SNP}	0.00	0.00	0.292	0.110	0.032	0.165	0.229	0.00	0.138
Section 2 h^2_{SNP}	0.215	0.611	0.313	0.263	0.799	0.109	0.273	0.818	0.262
Diff h^2_{SNP} estimates	0.007	0.046	0.011	0.014	0.085	0.001	0.016	0.095	0.008

Supplemental Table A5: **Assessment of bias due to population substructure**. Heritability (h^2_{SNP}) estimates when splitting the genome into two sections, with section 1 including chromosomes 1-11 and section 2 including chromosomes 13-31. Covariates included in the analysis were based on the model analyses of Table 1 and Supplementary Table 1. The Table provides the individual section h^2_{SNP} estimates, the sum of h^2_{SNP} estimates for both sections, and the difference between the summed value and that original h^2_{SNP} estimate that used the full data set. Heritability was calculated using GCTA and the linkage disequilibrium corrected GRM (wGRM). . Abbreviations: GLU: glucose, GLU-OST: glucose post oral sugar test, INS: insulin, INS-OST: insulin post oral sugar test, NEFA: non-esterified fatty acids, TG: triglycerides, ADIPON: adiponectin, SE: standard error

Response Variable	Human h2 Estimates	Human h^2_{SNP} Estimates	Horse h^2_{SNP} Estimates
Glucose	0.14-0.81	0.33	0.21-0.41
Insulin	0.43-0.51	0.23	0.59-0.81
Insulin Sensitivity			
Glucose-OST	0.17-0.24	-	0.23-0.57
Insulin-OST	-	-	0.36-0.44
HOMA-IR	0.38-0.48	-	-
Dyslipidemia			
NEFA	-	-	0.43-0.75
Triglycerides	0.17-0.60	0.47	0.27-31
Total cholesterol	0.42-0.62	-	-
HDL	0.42-0.63	0.48	-
LDL	0.58	-	-
Measures Obesity			
BMI	0.39-0.68	0.34	-
WC	0.27-0.46	-	-
Adipokines			
Leptin	0.28-0.55	-	0.49-0.55
Adiponectin	0.51	-	0.49-0.91
Other			
MetS	0.11-0.38	-	-
MetS Score	0.15-0.34	0.24-0.25	-
Systolic BP	0.16-0.28	0.30	-
Diastolic BP	0.12-0.38	-	-

Supplemental Table A6: Comparison of narrow-sense heritability estimates for metabolic syndrome in humans and horses using a pedigree based (h2) or SNP based (h^2_{SNP}) analysis. Abbreviations: OST (oral sugar test), HOMA-IR (homeostatic models assessment of insulin resistance), NEFA (non-esterified fatty acids), HDL (high density lipoproteins), LDL (low-density lipoproteins), BMI (body mass index), WC (waste circumference), MetS (metabolic syndrome as a binary trait and typically defined as the presence of three or more components of MetS), BP (blood pressure).

Appendix B: Chapter 3 Supplemental Materials

Welsh Pony	Section	Sire	Dam
6229	B	B	B
6230	B	B	B
6231	B	B	B
6234	B	B	B
6235	B	B	B
6236	B	B	B
6280	B	B	B
6281	B	B	B
6333	B	B	A
6334	H	A	H (0.5B_0.5Arab)
6362	A	A	A
6427	B	B	Inactive
6509	NA	-	-
6510	NA	-	-
6548	B	B	A
6549	B	B	A
6564	A	A	A
6566	A	A	A
6567	NA	-	-
6568	A	A	A
6569	A	A	A
6572	A	A	A
6573	NA	-	-
6574	NA	-	-
6575	A	A	A
6575	A	A	A
6576	A	A	A
6578	A	A	A
6579	B	B	B
6580	B	B	B
6581	B	B	B
6582	B	B	B
6583	A	A	A
6584	P	Arab	H (0.5B_0.25A_0.25TB)
6585	NA	-	-
6586	P	Arab	H (0.5B_0.25A_0.25TB)
6587	B	B	B
6603	NA	-	-
6604	A	A	A

Welsh Pony	Section	Sire	Dam
6605	A	A	A
6606	D	D	D
6607	A	A	A
6608	A	A	A
6611	A	A	A
6612	B	B	B
6635	B	B	B
6636	B	B	B
6637	D	D	D
6638	B	B	B
6639	B	B	B
6640	B	B	B
6641	B	B	B
6642	B	B	B
6643	B	B	B
6644	B	B	B
6660	B	Inactive	Inactive
6661	B	B	B
6678	NA	-	-
6690	B	A	B
6691	B	B	B
6692	B	B	B
6693	B	B	B
6694	B	B	B
6695	B	B	B
6696	B	B	B
6697	B	B	B
6698	B	B	B
6699	NA	-	-
6700	B	B	B
6701	B	B	B
6702	B	B	B
6703	B	B	B
6704	B	B	B
6705	B	B	B
6706	B	B	B
6707	B	B	B
6708	B	B	B
6709	B	B	B
6710	B	B	B

Welsh Pony	Section	Sire	Dam
6711	B	B	B
6712	B	B	B
6713	B	B	B
6714	B	B	B
6715	A	A	A
6716	B	B	B
6718	A	A	A
6719	B	B	B
6720	B	B	B
6741	H	A	Unknown
6742	H	A	Unknown
7661	B	B	B
7662	B	B	B
7663	A	A	A
7664	C	C	Inactive
7665	B	B	B
7666	D	D	D
7667	D	D	D
7668	B	B	B
7669	B	B	B
7670	B	B	B
7671	B	B	B
7672	A	A	A
7673	A	A	A
7674	A	A	A
7675	A	A	A
7676	D	D	D
7677	B	B	B
7678	B	B	B
7679	B	B	B
7680	B	B	B
7681	B	Inactive	Inactive
7682	B	B	B
7683	B	Inactive	Inactive
7684	B	B	B
7685	B	B	B
7686	H	B	H (0.5A_0.5WB)
7688	B	B	B
7689	A	A	A
7690	A	A	A

Welsh Pony	Section	Sire	Dam
7691	A	A	A
7692	A	A	A
7693	A	A	A
7694	A	A	A
7695	A	A	A
7696	A	A	A
7697	A	A	A
7698	A	A	A
7699	A	A	A
7700	A	A	A
7701	D	D	D
7702	A	A	A
7703	C	A	C
7704	A	A	A
7705	A	A	A
7706	A	A	A
7707	A	A	A
7708	A	A	A
7709	A	A	A
7710	B	B	B
7711	A	A	A
7712	A	A	A
7713	A	A	A
7714	B	B	B
7716	B	B	B
7719	B	B	Inactive
7720	B	B	B
7721	B	B	B
7722	A	A	A
7723	H	H (0.5B_0.5TB)	B
7724	B	B	B
7725	H	B	Unknown
7726	P	H**	H**
7729	H	B	H (0.5A_0.5TB)
7730	P	H*	H**
7735	P	H*	H (0.5B_0.5Unknown)
7737	P	H*	H (0.5B_0.25A_0.25TB)
7739	P	H*	H (0.5B_0.5Unknown)
7740	H	H*	B
7741	A	A	A

Welsh Pony	Section	Sire	Dam
7743	C	A	D
7744	H	B	H (0.5A_0.5QH)
7745	D	D	D
7746	D	D	D
7747	D	D	D
7748	B	B	B
7749	D	D	D
7750	B	B	B
7751	B	B	B
7752	B	B	B
7753	D	D	C
7754	D	D	D
7755	D	D	D
7756	D	Inactive	Inactive
7757	B	B	B
7758	B	B	B
7759	B	B	B
7760	B	B	B
7761	B	B	B
7762	B	B	B
7763	A	A	A
7765	A	A	A
7766	A	A	A
7767	B	B	B
7768	B	B	B
7769	B	B	B
7770	A	A	A
7771	A	A	A
7772	B	B	B
7773	B	B	B
7774	D	D	D
7775	B	B	B
7776	B	B	A
7777	B	B	B
7778	B	B	B
7779	A	A	A
7780	B	B	B
7782	B	A	B
7783	B	A	B
7784	H	A	H (0.5B_0.5BRP)

Welsh Pony	Section	Sire	Dam
7785	A	A	A
7786	A	A	A
7787	A	A	A
7788	A	A	A
7789	A	A	A
7790	A	A	A
7791	A	A	A
7792	A	A	A
7793	A	A	A
7794	A	A	A
7795	B	B	B
7796	A	A	A
7797	A	A	A
7798	A	A	A
7799	A	A	A
7801	H	A	Inactive
7802	A	A	A
7803	A	A	A
7804	A	A	A
7805	A	A	A
7806	A	A	A
7807	B	B	B
7808	B	B	B
7809	B	B	B
7810	B	B	B
7811	B	B	B
7812	B	B	B
7813	B	B	B
7814	B	Inactive	Inactive
7815	B	B	B
7816	B	B	B
7817	B	B	B
7818	B	B	B
7819	B	B	B
7820	B	B	B
7821	B	B	B
7822	A	A	A
7823	B	B	B
7824	B	B	B
7829	B	B	B

Welsh Pony	Section	Sire	Dam
7831	B	B	B
7832	B	B	B
7833	B	B	B
7834	B	B	B
7835	B	B	B
7836	B	B	B
7837	B	B	B
7838	B	B	B
7839	B	B	B
7840	B	B	B
7841	B	B	B
7842	B	B	B
7843	B	B	B
7844	B	B	B
7845	B	B	B
7846	B	B	B
7847	B	B	B
7848	B	B	B
7849	B	B	B
7850	B	B	B
7851	B	B	B
7852	B	B	B
7853	B	B	B
7854	B	B	B
7855	B	B	B

Supplemental Table B1: Pedigree information for the 264 Welsh ponies sequenced on one of three SNP chip arrays. The column labeled at "Section" represents the section for the Welsh ponies in this study, followed by the Sire's section and the Dam's section. Inactive indicates a sire or dam which were no longer active in the database and registered section was not available. NA represents a study Welsh pony which was unregistered or in which the pedigree information was unavailable. Unknown indicates a pedigree for a sire or dam which could no longer be traced. For the dam or sire listed as a section H, additional breed information was provided if available.

Abbreviations: QH: Quarter horse, BRP: British Riding Pony, Arab: Arabian, WB: Warmblood, TB: Thoroughbred

* 0.6875B_0.125TB_0.125Arab_0.0625Unknown

**0.5B_0.25A_0.25Unknown

	10M Iterations Seeds 1-10		10M Iterations Seeds 11-20		20M Iterations Seeds 1-10		30M Iterations Seeds 1-10	
CHR	**SNPs_Sugg**	**SNPs_Sign**	**SNPs_Sugg**	**SNPs_Sign**	**SNPs_Sugg**	**SNPs_Sign**	**SNPs_Sugg**	**SNPs_Sign**
1	2	0	NA	NA	NA	NA	NA	NA
1	1	0	NA	NA	1	0	4	0
2	38	27	5	1	32	2	5	1
3	NA	NA	NA	NA	1	NA		
4	54	4	25	2	3	0	5	2
6	68	4	8	1	11	2	11	2
7	14	0	10	0	5	0	8	1
8	6	0	NA	NA	NA	NA	NA	NA
9	NA	NA	2	0	NA	NA	NA	NA
14	NA	NA	NA	NA	1	0	NA	NA
15	6	1	5	1	5	2	12	1
16	NA	NA	NA	NA	2	0	NA	NA
18	NA	NA	1	0	NA	NA	NA	NA
19	NA	NA	NA	NA	NA	NA	3	0
20	NA	NA	NA	NA	NA	NA	1	0
22	NA	NA	NA	NA	4	0	NA	NA
23	NA	NA	11	0	NA	NA	NA	NA
24	NA	NA	4	1	16	0	NA	NA

Supplemental Table B2: Repeatability across results for the Bayesian sparse linear mixed model (BSLMM) using the max gamma values from 10 million (M) iterations with seeds 1-10, 10M iterations with seeds 11-20, 20M iterations with seeds 1-10, and 30M iterations with seeds 1-10 for adiponectin concentrations in the Morgan horses. Regions which are highlighted in yellow indicate those which would have been identified as a region of interest (contained a minimum of five SNPs exceeding the suggestive threshold, with at least one SNP exceeding the threshold for genome wide significance). Abbreviations: SNPs_Sugg (total number of SNPs which exceeded the suggested threshold for genome-wide significance), SNPs_Sign (total number of SNPs which exceed the threshold for genome-wide significance)

Appendix C: Chapter 4 Supplemental Materials

Supplemental Table C1: Whole Genome Sequencing for Welsh ponies				
EMS_ID	**Total Reads**	**Average Read Length**	**Average Sequencing Depth (nuclear)**	**Hits Unique Reads (nuclear)**
EMS_28	94749945	128.9	156238380	7.9
EMS_363	79493183	129.2	125604533	6.7
EMS_369	67904285	102.2	109771004	4.5
EMS_373	66358100	102.3	92834439	3.7
EMS_376	131590489	104.6	203593797	8.5
EMS_457	76025659	129.6	114078968	5.8
EMS_657	91515978	129.1	150994191	7.7
EMS_676	80877817	129.1	131351193	6.7
EMS_697	86085735	129.0	142069679	7.2
EMS_699	86881548	129.0	143257878	7.3
EMS_737	84172015	129.3	134460983	6.8
EMS_738	88063255	129.1	144871377	7.4
EMS_739	86795272	129.0	144293851	7.3
EMS_765	85901668	129.0	141878173	7.2
EMS_790	77984605	129.2	126556755	6.4
EMS_794	82227090	129.3	131739979	6.7
EMS_812	79792809	129.4	125624597	6.4
EMS_820	84065313	129.1	138711729	7.0

Supplemental Table C1: Whole genome sequencing summary data for Welsh ponies.

Supplemental Table C2: Whole Genome Sequencing for Morgan horses				
EMS_ID	**Total Reads**	**Average Read Length**	**Average Sequencing Depth (nuclear)**	**Hits Unique Reads (nuclear)**
EMS_9	81065821	97.5	117315572	4.6
EMS_49	148929760	100.3	236828180	9.4
EMS_50	87056999	97.9	127080299	5.0
EMS_91	81214822	99.6	112563159	4.5
EMS_93	104266887	103.5	143919132	6.0
EMS_134	97948336	100.1	139507318	5.6
EMS_246	106345544	101.0	142200305	5.8
EMS_259	104530622	100.2	153194262	6.2
EMS_265	104915827	100.6	149865399	6.0
EMS_279	195023299	105.7	274671254	11.7
EMS_333	191762166	105.0	292455094	12.2
EMS_336	208434711	99.7	302521708	12.1
EMS_355	169999656	101.0	269295263	10.8
EMS_395	85216006	96.6	127846307	4.9
EMS_479	84377490	97.3	120207175	4.7
EMS_595	213209544	104.0	330204194	13.7
EMS_605	86507364	97.9	128299775	5.0
EMS_611	105409660	101.3	154462677	6.3

Supplemental Table C2: Whole genome sequencing summary data for Morgan horses.

Appendix D: Chapter 5 Supplemental Methods:

Description for GWAS Custom Code: A Bayesian Sparse Linear Mixed Model (BSLMM) [430], available in the software program Genome-wide Efficient Mixed Model Association (GEMMA) [379], was used to rank SNPs based on the number of times in 10 million iterations that a SNP was estimated to have a large effect. This step was repeated 10 times and the maximum beta-value for each SNP was used for final ranking. Chromosomes were then divided into 500KB segments, and the top and two adjacent SNPs were kept within each segment. These SNPs were then used to build the select SNP GRM based on a stepwise feature selection, where each SNP was kept only if it is determined that it significantly improves the null model, which included both random and fixed effects. If a SNP was selected, inclusion of that SNP becomes the new null model for testing of the next SNP. Once the select SNP GRM was built, a linear mixed model, using the software program FaST-LMM [431], was performed with the select SNP GRM in place of the full GRM. FaST-LMM's algorithm tests each SNP individually for an effect on the phenotype using a maximum likelihood estimation [431]. SNPs within 1MB of the tested SNP were excluded from the select GRM to avoid double fitting of the model.

Description of LDAK software analyses: A full description of LDAK and the analyses available is available in Speed, *et. al.* [411,413]. Briefly, LDAK's algorithm uses restricted maximum likelihood to estimate the variance explained by all SNPs for a given phenotype. Unlike other mixed linear models available to estimate h2, LDAK incorporates an LD weighted genetic relationship matrix (LD-GRM) and a scaling factor to account for the effect of minor allele frequency on h2. The main output file includes the h2 estimate, standard deviation (SD), log likelihood of the estimate, and the p-value for the log likelihood. For genomic partitioning, the LD-GRM is comprised only of SNPs from the specified region; h2 is then estimated from this subset of SNPs. For our analysis, we utilized our ROI identified on GWAS and *di* statistic, and included age, sex and section as covariates. The top predictors approach fit the top SNPs from the GWAS as covariates in the analysis using LDAK's --top-preds function. The output from this analysis includes the genetic variance explained by the top predictors, the genetic variance explained by the

remainder of the SNPs, and then the sum of these values as the overall estimate of heritability. A SD is not provided for the top predictors.

Description of Random Sub-setting of Data: We also performed random sub-setting of the data by removing 10% of the population using the software package R's random number generator without replacement.[412] LD-GRMs were constructed from the reduced cohort and heritability estimates were calculated for total heritability, genomic partitioning of our ROI, and with the top ECA6 SNPs from the GWAS as covariates. This process was repeated 100 times and the average of all heritability estimates and SD were calculated and compared with the original estimates.

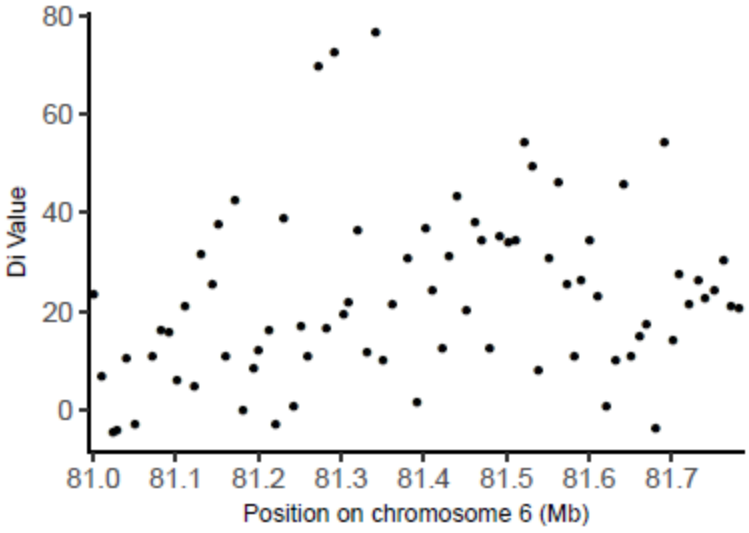
*Description of TaqMan SNP Genotyping assay***:** A TaqMan SNP genotyping assay was utilized to efficiently genotype individuals for the HMGA2 c.83G>A variant. Forward (CTTCAGCCCAGGGACAAC) and reverse (AAGCAGCAGCAAGTCAGT) PCR primers were designed to produce an 80 base pair amplicon that included the HMGA c.83G>A variant. Locked nucleic acid (LNA) probes, with a 5' fluorescent reporter dye and 3' quencher, were designed for allelic discrimination between the G allele (5HEX/AG+A+GA+G+G+ACG/3IABkFQ) and the A allele (56-FAM/AG+A+GA+G+A+A+CGC/3IABkFQ) as shown below.

GGCGGACTCGGGGCGGCTGAGGCCAGCGGCTGCAGCGGCGGTAGCGGCGGC
GGCGGCGGGAGGCAGGATGAGCGCACGCGGTGAGGGCGCCGGGCAGCCGTC
CA**CTTCAGCCCAGGGACAAC**CTGCCGCCCCGGCGCCTCAGA**AGAGAGG/AA**
CGCGGCCGACCCAGG**AAGCAGCAGCAAGTCAGT**ACGCGGGCGGGGTGGGG
GCACCAGCCCGCCTCCGCGCCCTCCGCGAAGGCCCGGCCACGCGCGGCCCCG
AGCGCGGGAGCCGGGTCGCCGCGCGCCGCCCGCCGGCCGGAGGCGGG

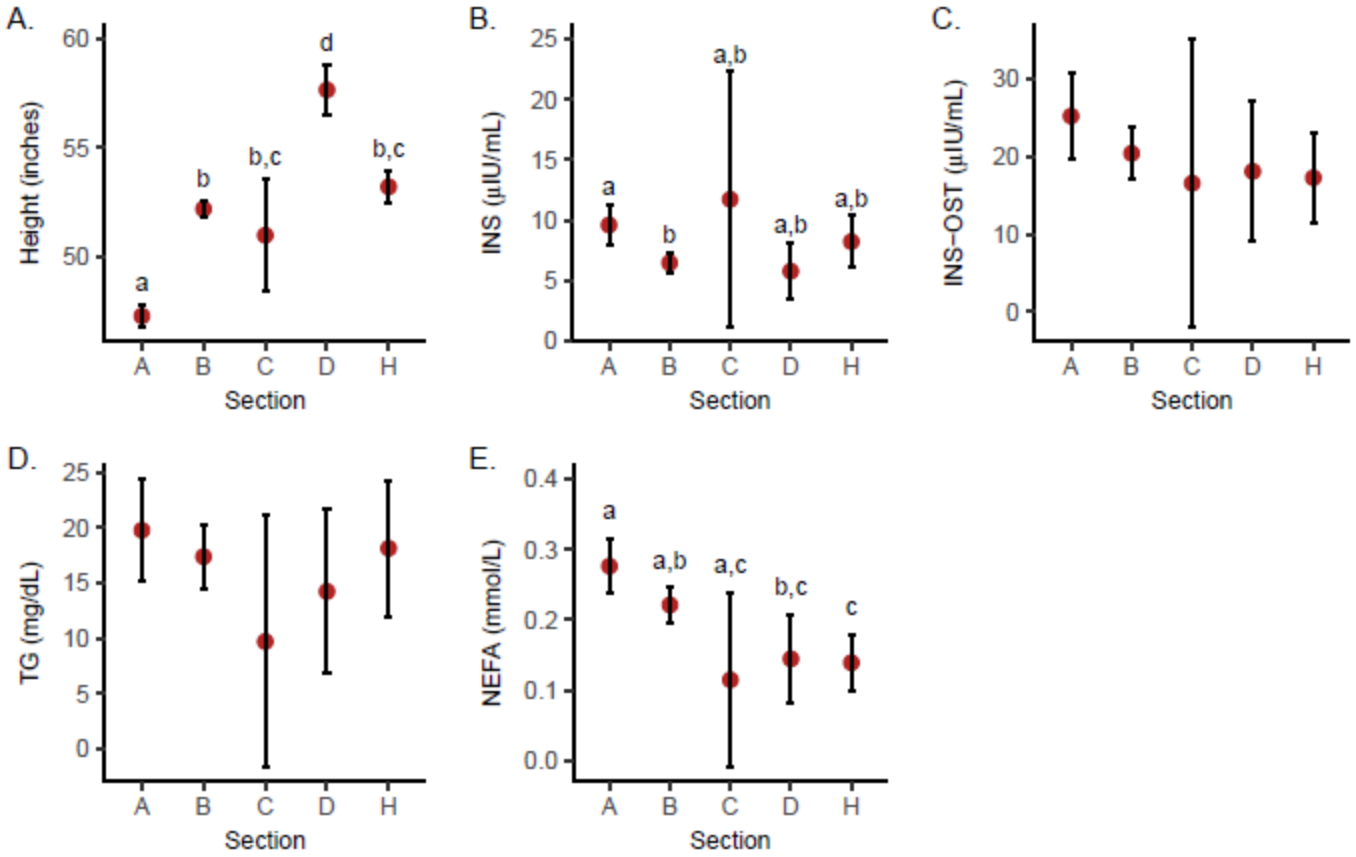
Supplemental Methods Figure 1: HMGA2 sequence approximately 150 bp 5' and 3' of the c.83G>A variant. The forward primer is indicated by green text and the reverse primer is indicated in red text. Orange text indicates sequence targeted by the florescent-tagged LNA probes.

Reaction components and volumes for each reaction were as follows: 5µL of 5ng/uL DNA template, 1µL forward primer, 1 µL reverse primer, 0.1 µL HEX probe, 0.1µL FAM probe,

5µL PrimeTime® Gene Expression Master Mix, 4µL betaine and 3µL molecular biology grade water for a final volume of 19.2µL. Final cycling protocol was as follows: Cycle 1 (1 repeat): 95^0C for 3 minutes. Cycle 2 (40 repeats): step 1: 95^0C for 5 seconds and step 2: 61^0C for 30 seconds.



Supplemental Figure D1. Local *di* values for the equine chromosome 6 (ECA6) segment analyzed further in this study. Each d_i value is plotted on the y axis and the ECA6 position in bp is shown on the x axis.



Supplemental Figure D2: Least-square mean estimates and 95% confidence intervals for height or EMS phenotypes and section in a population of 283 registered Welsh ponies adjusting for both age and sex. Height (A), insulin (B), INS-OST (C), triglycerides (D), and NEFA (E). Abbreviations: INS-OST = insulin after an oral sugar test, NEFA = non-esterified fatty acids.

Section	Pedigree	Height Requirements
A	A X A	Up to 12.2 hands (50 in)
B	AXB or BXB	Up to 14.2 hands (58 in)
C	At least one C or D parent	Up to 13.2 hands (54 in)
D	At least one C or D parent	Over 13.2 hands (54 in)
H	At least one registered purebred Welsh parent	No height limit
P	At least one registered half Welsh parent	No height limit

Supplemental Table D1: **Breed requirements for Welsh pony sections based on pedigree and height requirements.**

Breed	Number of Individuals
Yakutian	9
Welsh Pony	44
Warmblood	18
Thoroughbred	25
Standardbred	40
Przewalski	13
Morgan	61
Maremmano	22
Lusitano	21
Icelandic	18
Hanoverian	8
French Trotter	21
Franchese Montagne	30
Belgian	22
Arabian	36
Quarter Horse	75
Total	**463**

Supplemental Table D2: Reference population used for the calculation of *di*.

Exon	EquCab2	EquCab3	HMGA2 Horse Exon Sequence	Forward Primer	Reverse Primer	BP
1	N/A	80374152 - 80374262	ATGAGCGCACGCGGTGAGG GCGCCGGGCAGCCGTCCACT TCAGCCCAGGGACAACCTGC CGCCCCGGCGCCTCAGAAGA GAGGACGCGGCCGACCCAG GAAGCAGCAGCAA	CTCGTCCTCCAGCCCTATC	CGTGCACAATAGCGAAAGT C	491
2	81392745 - 81392831	80376870 - 80376956	GAGCCAACCGGTGAGCCCTC TCCTAAGAGACCCAGGGGAA GACCCAAAGGCAGCAAAAA CAAGAGTCCCTCCAAAGCAG CTCAAAAG	GTTCCAACCCTTCTGTCGA G	ACTGGGTTTTGCAGTAGTCA	623
3	81402791 - 81402841	80386916 - 80386966	AAAGCAGAAGCCACTGGAG AAAAACGGCCAAGAGGCAG ACCTAGGAAATGG	AAACGGGGCAGAGGAAT CTA	GAGCGTCTCCTGGAAAGAA C	458
4	81503615 - 81503647	80487792 - 80487824	CCACAACAAGTCGTTCAGAA GAAGCCTGCTCAG	GACCATGTATAAACACCC TTTAACC	GGTTTTTAATCACACAACCA CAG	383
5	81515174 - 81515221	80499351 - 80499398	GACAATGTTGCCTTGCCTGG GAAAGACCATCTAGGCAATC TTATGTGTCTACTACTCTTTA TAAATGCTGCTTGA	GCAGAACCTGCTGGAGTC AC	TGTGGGCAAGTGAATAATT G	398

Supplemental Table D3: PCR primers for Sanger sequencing and annotation for *HMGA2* exon sequencing. Based on poor annotation of the *HMGA2* gene in EquCab2 reference genome, we did a full reconstruction of the gene. Notably, Ensembl has this gene positioned for the horse at equine chromosome 6 (ECA6): 81,197,462-81,402,841 in contrast to NCBI position at ECA6: 81,389,151-81,518,054. Neither assembly included the ~1.4 kilobases annotated by Frishchknecht et al, including exon 1 and the 5' UTR (GenBank: LN8490000.1). Based on our annotation of exons 2-5, the NCBI position appears more accurate and corresponds with the most predominant peak identified in the haplotype analysis for baseline insulin (ECA6: 81,381,221-81,583,507). Base pair locations for EquCab2 and EquCab3 are also provided.

Supplemental Table D4: PCR primers and annotation for *IRAK3* exon sequencing.

Exon	EquCab2	EquCab3	IRAK3 Horse Exon Sequence	Forward Primer	Reverse Primer	BP
1	N/A	80681675 - 80681889	AATTTCCGCGGTTGTGTAACTGCCCCCCG CGGGCGCGCAGCGGCCTGGCCTCGCCTCC CTCCTCCACCGGGCCCTGCTCTCCGGCGG CAGAGCTATGGCCTGGGCGGCAGCGGCG GGCAGCGGCGGGGCCCGCGGCCAGCTCT CGGCGCACACGCTCCTCTTCGACCTGCCT CCCGTGCTGCTGGGCGAGCTCTGCGCCGT CCTGGACAGCTGCGACGGCGCGCTCGGCT GGCGCCGCCTCGGTGAGTGCGCCCGGGCG GGCGGGGGCTCCGCGCACGCTCCGTGCCC CCCGCGCCGGGGGTCGCTCCACCGGCCCC AGGACTGGGGCCTCCGCGCACGCTCCGTG CGCCCCGCGCCGTGCTGGGGCCGCTCCGC CCGTTCCTGGACCGAGTCTCCGTG	AATTTCCGCGGTTGTG TAAC	CACGGAGACTCGGTCC AG	399
2	81711764 - 81711955	80696010 - 80696192	GCGGAGCGACTTTCAAGCAGCTGGCTGGA TGTTCGTCACATTGAAAAGTATGTAGACC AAGGGAAAAGTGGAACGAGAGAATTGCT TTGGTCCTGGGCACAGAAAAACAAGACC ATCGGTGACCTTTTACAGATCCTCCAGGA GATGGGGCATCATCGAGCTATCCATTTAA TTACAAACCATGGTAAACAC	GTTGTCACTGCCTCCG ATC	AAGTTTGGCAAGAAG GAAGGA	580
3	81720696 - 81720760	80704941 - 80705005	GAGCAGCCTTGAATCCTTCAGAGCAGAGT CACCTGGGAGATGGATTTCCAAGCATGTT ACCCAAG	TTCAATGGAAATGACA CTGAGC	TCCTGAATCCCCAACT AAACA	372
4	81721655 - 81721709	80705900 - 80705954	GAAACAACCAATGTCACAGTGGATAATGT TCTTATTCCTAAACATAATGAAAAAG	AGAGTTGGTATGGAA GCCTT	CACCCATCAGAACCAT GTGT	437
5	81722977 - 81723128	80707222 - 80707373	GAATATTGTTTAAACCTTCTATCAGCTTTC AAAACATCACAGAAGGAACCAAAAATTT CCACAAAGACTTCCTAATTGGAGAAGGG GAGATTTTTGAGGTGTACAGAGTGGAGAT CCAAAACCGAACGTATGCCGTTAAATTAT TTAAACAG	GCTCTCCTGACTTTCC ACTG	TCAGTCATTTCTCCAG TCACC	451
6	81726877 - 81726941	80711122 - 80711186	GAGAAAAAAATGCAATGTAAGCAACAAT GGAAGAGCTTTTTATCTGAGCTTGAAGTT TTACTACT	AGCGGTGGTTCTGATT GTTT	AGGAATATACCAAGG CAGATGT	418

Supplemental Table 4 (cont): PCR primers and annotation for *IRAK3* exon sequencing						
Exon	**EquCab2**	**EquCab3**	**IRAK3 Horse Exon Sequence**	**Forward Primer**	**Reverse Primer**	**BP**
7	81731400 - 81731514	80715645 - 80715759	GTTTCATCATCCAAACATTCTGGAGTTGGCTGCA TATTTTACAGAGAGTGACAAGTTCTGCCTGGTTT ATCCGTATATGAGAAATGGGTCCCTTTTTGACAG ACTGCAGTGTGTA	AATCTGTGTCACG TGTCTGG	GTCCTCTGCTTTCTTG GGAA	418
8	81732314 - 81732432	80716559 - 80716677	GGTAACACAGCCCCGCTCTCTTGGCACATTCGAA TCAGTATCTTAATAGGAGCGTGCAAGGCCATCC AGTATTTGCACAACATCGAGCCGTGCTCAGTTGT CTGTGGCAGCATCTCCAG	AGGTCGTCAGTA GTAGAGGA	AGAGTTCTTCACACGA GCAC	406
9	81745387 - 81745585	80729635 - 80729833	TGCAAACATACTTTTGGATGATCAGTTTCAACCC AAACTAACTGATTTTGCCGTGGCGCACTTCCGAC CCCACCTTGAACACCAGCACTGCACCATCAGCGT GACCGGCTGCAACAGGAAACACCTGTGGTACAT GCCCGAGGAGTACGTCAGGCAGGGCAGACTCAC CGTCAAAACCGACGTCTACAGCTTTGGGATT	GCTCTAGTTCGTG GAAAATTGC	TTTATATTTTATTGCTT GACTGACTGC	384
10	81745841 - 81745903	80730089 - 80730151	GTAATCATGGAAGTTCTGACAGGTTGTAAAGTG GTGTTGGATGAGCCAAAGCACATCCAGCTG	TGCAGTCAGTCA AGCAATAAA	ATTTTCTGTGGTGCCT GGTT	661
11	81745992 - 81746156	80730240 - 80730404	AGGGATCTTCTTATGGAATTGATGGAAAAGAGA GGCCTTGATTCATGTCTCTCATATCTAGATAAGA AAGTGTGTCCCTGTCCTCGGAATTTCTCTGCCAA GCTGTTCTCTTTGGCGGGCCAGTGTGCTGCAACA CGGGCCAAGTTGAGACCATCGATGGATGAA	TGCAGTCAGTCA AGCAATAAA	ATTTTCTGTGGTGCCT GGTT	661

Exon	EquCab2	EquCab3	IRAK3 Horse Exon Sequence	Forward Primer	Reverse Primer	BP
12	81748556 - 81749029	80732804 - 80733277	GTCCTGACCGTTCTTGAGAGCACTCCGGCCAGCT TGTATTTTGCTGAAGACCCTCCCGCCTCACTGAA GTCCTTCAGGTGTCCTTCTCCTCTGTTCTTGGACA ACGTACCAAGTATTCCAGTGGAAAATGATGAAA ACCAGAATAACTCTTCCCTGCCTCCTGATAAAGC TTGGAGAAAAGAGAGAATGACTCAGAAAATTCC CTTTGAATGTAGCCAGTCTGAGGTGACGTTTCTG GGCTTTGAGAGAAAGACAGGGAGTCAGAGAAAT GAGGATGCTTGCAACATACCCAGTTCTTCTTGTG AAAAGAGTTGGTCTCCAAAGGATGCAGCTCCAT CCCAGGACTCCAGCACCTGTGGTGTGACTATGG ACCCTTCTGCAGAAGCTCTGGGCCAGTCTTACAG GAGCAGGCCAATGGAGATTAGCTGGTCTTCTGA ATTTTCCTGGAATGAATGTGAAGAGTACAAAAA GGAG	GCTCAGGGACCA TGTTTCTC	ATTTCTAAGCCACCCC GTTT	770

Supplemental Table 4 (cont): PCR primers and annotation for *IRAK3* exon sequencing.

Test	Height	Insulin	INS-OST	NEFA	Triglycerides
ANOVA F-statistic Additive	-3.78 SE=0.21 p=<2.2e-16 F=<2.2e-16	0.18 SE=0.031 p=7.8e-09 F=1.3e-10	0.18 SE=0.038 p=1.9e-06 F=1.2e-09	0.043 SE=0.014 p=2.1e-03 F=7.1e-04	0.14 SE=0.038 p=2.0e-04 F=9.2e-08
ANOVA F-statistic Recessive	-4.89 SE=0.317 p=<2.2e-16 F=<2.2e-16	0.28 SE=0.043 p=4.9e-10 **F=8.8e-12**	0.28 SE=0.052 p=1.3e-07 **F=9.3e-11**	0.074 SE=0.020 p=1.7e-04 **F=8.7e-05**	0.20 SE=0.053 p=2.0e-04 F=9.3e-08
ANOVA F-statistic Dominant	-6.21 SE=0.56 p=<2.2e-16 F=<2.2e-16	0.213 SE=0.071 p=2.8e-03 F=1.9e-05	0.19 SE=0.085 p=0.024 F=6.3e-05	0.029 SE=0.031 p=0.034 F=0.042	0.20 SE=0.085 p=0.018 F=5.1e-05
AIC Additive	**1403.8**	**251.2**	**343.4**	**-229**	**373**
AIC Recessive	1443	**249.2**	**341.1**	**-230.3**	**371.7**
AIC Dominant	1521.3	279.5	364.5	-216.9	390
AIC Co-Dominant	1584.9	267.2	352.3	-220.7	389.9

Supplemental Table D5: ANOVA results and Akaike information criterion (AIC) values for models of inheritance between the *HMGA2* c.83G>A variant and height and the four EMS traits significantly correlated with genotype. ANOVA results and AIC values for models of inheritance between the *HMGA2* c.83G>A variant and height and the four EMS traits significantly correlated with genotype. Deciding values are highlighted in red. For height, an additive model was the best fit model (lowest AIC). For the EMS traits, p-value for the f-statistic slightly favored the recessive model but the AIC could not differentiate between a recessive and additive model. For example, the AIC for the recessive insulin model was 249.2 and 251.2 for the additive model, which can be interpreted as the additive model being 0.36 [exp^((249.2-251.2)/2)] times as likely as the recessive model, concluding that there is insufficient information to support picking either model. Abbreviations: INS-OST = insulin post oral sugar test, NEFA = non-esterified fatty acids.

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.80499826.PC	80499826	81667046			
MNEc.2.6.80501042.PC	80501042	81668273			
MNEc.2.6.80501273.PC	80501273	81668520		X	
MNEc.2.6.80502406.PC	80502406	81669653			
MNEc.2.6.80503522.PC	80503522	81670789			
MNEc.2.6.80503614.PC	80503614	81670881			
MNEc.2.6.80503671.PC	80503671	81670938			
MNEc.2.6.80504332.PC	80504332	81671599			
MNEc.2.6.80504637.PC	80504637	81671904			
MNEc.2.6.80504799.PC	80504799	81672066			
MNEc.2.6.80505411.PC	80505411	81672678			
MNEc.2.6.80509032.PC	80509032	81676297			
MNEc.2.6.80512076.PC	80512076	81679341			
MNEc.2.6.80512513.PC	80512513	81679778			
MNEc.2.6.80513598.PC	80513598	81680863			
MNEc.2.6.80515954.PC	80515954	81683219			
MNEc.2.6.80516105.PC	80516105	81683370			
MNEc.2.6.80516221.PC	80516221	81683486			
MNEc.2.6.80518479.PC	80518479	81685747			
MNEc.2.6.80518512.PC	80518512	81685780			
MNEc.2.6.80518693.PC	80518693	81685961			
MNEc.2.6.80520101.PC	80520101	81687369		X	
MNEc.2.6.80521407.PC	80521407	81688626			
MNEc.2.6.80522351.PC	80522351	81689570		X	
MNEc.2.6.80523773.PC	80523773	81690992			
MNEc.2.6.80527094.PC	80527094	81694313		X	
MNEc.2.6.80527223.PC	80527223	81694442			
MNEc.2.6.80527481.PC	80527481	81694700			
MNEc.2.6.80533180.PC	80533180	81700398			
MNEc.2.6.80533647.PC	80533647	81700865			
MNEc.2.6.80544097.PC	80544097	81711319			
MNEc.2.6.80545253.PC	80545253	81712475			
MNEc.2.6.80545309.PC	80545309	81712531			
MNEc.2.6.80548131.PC	80548131	81715353			
MNEc.2.6.80554792.PC	80554792	81722014			
MNEc.2.6.80555907.PC	80555907	81723129			
MNEc.2.6.80557468.PC	80557468	81724690			
MNEc.2.6.80564229.PC	80564229	81731412			
MNEc.2.6.80567345.PC	80567345	81734532			
MNEc.2.6.80567618.PC	80567618	81734805			
MNEc.2.6.80568218.PC	80568218	81735405			
MNEc.2.6.80569235.PC	80569235	81736422			
MNEc.2.6.80569293.PC	80569293	81736480			
MNEc.2.6.80569852.PC	80569852	81737039			
MNEc.2.6.80570298.PC	80570298	81737485		X	
MNEc.2.6.80572371.PC	80572371	81739554			
MNEc.2.6.80572584.PC	80572584	81739767			
MNEc.2.6.80574750.PC	80574750	81741935			
MNEc.2.6.80574793.PC	80574793	81741978			
MNEc.2.6.80575672.PC	80575672	81742857			
MNEc.2.6.80576028.PC	80576028	81743213			

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.80579951	80579951	81747136			
MNEc.2.6.80580074.PC	80580074	81747259			
MNEc.2.6.80580136.PC	80580136	81747321			
MNEc.2.6.80580526.PC	80580526	81747711			
MNEc.2.6.80580756.PC	80580756	81747941			
MNEc.2.6.80582308.PC	80582308	81749493			
MNEc.2.6.80582595.PC	80582595	81749780			
MNEc.2.6.80585781.PC	80585781	81752969			
MNEc.2.6.80586564.PC	80586564	81753752			
MNEc.2.6.80587550.PC	80587550	81754738			
MNEc.2.6.80592143.PC	80592143	81759332			
MNEc.2.6.80592392.PC	80592392	81759581			
MNEc.2.6.80597639.PC	80597639	81764830			
MNEc.2.6.80597710.PC	80597710	81764901			
MNEc.2.6.80598790.PC	80598790	81765981			
MNEc.2.6.80599081.PC	80599081	81766272			
MNEc.2.6.80599420.PC	80599420	81766612			
MNEc.2.6.80602873.PC	80602873	81770065			
MNEc.2.6.80605857.PC	80605857	81773049		X	
MNEc.2.6.80608356.PC	80608356	81775548			
MNEc.2.6.80609214.PC	80609214	81776406			
MNEc.2.6.80613296.PC	80613296	81780489			
MNEc.2.6.80613431.PC	80613431	81780624			
MNEc.2.6.80617709.PC	80617709	81784902			
MNEc.2.6.80619822.PC	80619822	81787015			
MNEc.2.6.80620304.PC	80620304	81787497			
MNEc.2.6.80620478.PC	80620478	81787671			
MNEc.2.6.80620792	80620792	81787985			
MNEc.2.6.80621253.PC	80621253	81788446			
MNEc.2.6.80621281.PC	80621281	81788474			
MNEc.2.6.80622121.PC	80622121	81789314			
MNEc.2.6.80622788.PC	80622788	81789981			
MNEc.2.6.80623531.PC	80623531	81790724			
MNEc.2.6.80625415.PC	80625415	81792608			
MNEc.2.6.80625991.PC	80625991	81793184			
MNEc.2.6.80627668.PC	80627668	81794861			
MNEc.2.6.80628078.PC	80628078	81795271			
MNEc.2.6.80634102.PC	80634102	81801300		X	
MNEc.2.6.80635142.PC	80635142	81802340			
MNEc.2.6.80636041.PC	80636041	81803239		X	
MNEc.2.6.80636779.PC	80636779	81803977			
MNEc.2.6.80639056.PC	80639056	81806255			
MNEc.2.6.80639161.PC	80639161	81806360			
MNEc.2.6.80639787.	80639787	81806986		X	X
MNEc.2.6.80639984.PC	80639984	81807183			
MNEc.2.6.80640275.PC	80640275	81807474			
MNEc.2.6.80642478.PC	80642478	81809677			
MNEc.2.6.80648966.PC	80648966	81816165			
MNEc.2.6.80649778.PC	80649778	81816951			
MNEc.2.6.80651081.PC	80651081	81818254			
MNEc.2.6.80651346.PC	80651346	81818519			

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.80654729.PC	80654729	81821902			
MNEc.2.6.80655300.PC	80655300	81822473			
MNEc.2.6.80656692.PC	80656692	81823865		X	
MNEc.2.6.80657480.PC	80657480	81824653			
MNEc.2.6.80658576.PC	80658576	81825749		X	
MNEc.2.6.80659413.PC	80659413	81826586			
MNEc.2.6.80659646.PC	80659646	81826819			
MNEc.2.6.80659981.PC	80659981	81827154			
MNEc.2.6.80660228.PC	80660228	81827401			
MNEc.2.6.80663069.PC	80663069	81830242			
MNEc.2.6.80666273.PC	80666273	81833447			
MNEc.2.6.80666491.PC	80666491	81833665			
MNEc.2.6.80670901.PC	80670901	81838075			
MNEc.2.6.80672874.PC	80672874	81840048			
MNEc.2.6.80674503.PC	80674503	81841677			
MNEc.2.6.80678320.PC	80678320	81845492			
MNEc.2.6.80681512.PC	80681512	81848686			
MNEc.2.6.80684153.PC	80684153	81851327			
MNEc.2.6.80686536.PC	80686536	81853710			
MNEc.2.6.80686702.PC	80686702	81853876			
MNEc.2.6.80688324.PC	80688324	81855498		X	
MNEc.2.6.80692551.PC	80692551	81859725			
MNEc.2.6.80694638.PC	80694638	81861812			
MNEc.2.6.80694729.PC	80694729	81861903			
MNEc.2.6.80697067.PC	80697067	81864242			
MNEc.2.6.80697327.PC	80697327	81864502			
MNEc.2.6.80700969.PC	80700969	81868144			
MNEc.2.6.80701317.PC	80701317	81868492			
MNEc.2.6.80701518.PC	80701518	81868693		X	
MNEc.2.6.80702649.PC	80702649	81869824			
MNEc.2.6.80703890.PC	80703890	81871065			
MNEc.2.6.80708442.PC	80708442	81875617			
MNEc.2.6.80710843.PC	80710843	81878018			
MNEc.2.6.80714065.PC	80714065	81881240			
MNEc.2.6.80715143.PC	80715143	81882318			
MNEc.2.6.80719193.PC	80719193	81886368			
MNEc.2.6.80722266.PC	80722266	81889441			
MNEc.2.6.80722564.PC	80722564	81889739			
MNEc.2.6.80722978.PC	80722978	81890153			
MNEc.2.6.80724746.PC	80724746	81891921			
MNEc.2.6.80728189.PC	80728189	81895365			
MNEc.2.6.80728297.PC	80728297	81895473			
MNEc.2.6.80729934.PC	80729934	81897110			
MNEc.2.6.80732384.PC	80732384	81899560			
MNEc.2.6.80733830.PC	80733830	81901006			
MNEc.2.6.80736496.PC	80736496	81903675			
MNEc.2.6.80740274.PC	80740274	81907453			
MNEc.2.6.80740392.PC	80740392	81907571			
MNEc.2.6.80742798.PC	80742798	81909977			
MNEc.2.6.80742855.PC	80742855	81910034			
MNEc.2.6.80744646.PC	80744646	81911825			

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.80745841	80745841	81913021			
MNEc.2.6.80746678.PC	80746678	81913858			
MNEc.2.6.80747017.PC	80747017	81914197			
MNEc.2.6.80753124.PC	80753124	81920304			
MNEc.2.6.80753306.	80753306	81920486			
MNEc.2.6.80756605.PC	80756605	81923785			
MNEc.2.6.80756667.PC	80756667	81923847			
MNEc.2.6.80758159.PC	80758159	81925341			
MNEc.2.6.80758193.PC	80758193	81925375			
MNEc.2.6.80759692.PC	80759692	81926874			
MNEc.2.6.80763018.PC	80763018	81930200			
MNEc.2.6.80763384.PC	80763384	81930566			
MNEc.2.6.80763601.PC	80763601	81930783			
MNEc.2.6.80763751.PC	80763751	81930933			
MNEc.2.6.80766863.PC	80766863	81934045			
MNEc.2.6.80767981.PC	80767981	81935160			
MNEc.2.6.80768143.PC	80768143	81935322			
MNEc.2.6.80770476.PC	80770476	81937655			
MNEc.2.6.80772574.PC	80772574	81939753			
MNEc.2.6.80773747.PC	80773747	81940926			
MNEc.2.6.80783645.PC	80783645	81950822			
MNEc.2.6.80784128.PC	80784128	81951305		X	
MNEc.2.6.80785871.PC	80785871	81953011			
MNEc.2.6.80786333.PC	80786333	81953473			
MNEc.2.6.80787590.PC	80787590	81954730			
MNEc.2.6.80787822.PC	80787822	81954962			
MNEc.2.6.80790795.PC	80790795	81957936			
MNEc.2.6.80792111.	80792111	81959252			
MNEc.2.6.80792181.	80792181	81959322			
MNEc.2.6.80793324.PC	80793324	81960465			
MNEc.2.6.80794944.PC	80794944	81962085			
MNEc.2.6.80795347.PC	80795347	81962488			X
MNEc.2.6.80795503.PC	80795503	81962644			
MNEc.2.6.80796963.PC	80796963	81964104			
MNEc.2.6.80797343.PC	80797343	81964484			
MNEc.2.6.80801661.PC	80801661	81968699			
MNEc.2.6.80802867.PC	80802867	81969905			
MNEc.2.6.80806293.PC	80806293	81973331			
MNEc.2.6.80806580.PC	80806580	81973618		X	
MNEc.2.6.80807204.PC	80807204	81974242			
MNEc.2.6.80815186.PC	80815186	81982237			
MNEc.2.6.80815571.PC	80815571	81982623			
MNEc.2.6.80817011.PC	80817011	81984046			
MNEc.2.6.80818111.PC	80818111	81985146			
MNEc.2.6.80819060.PC	80819060	81986095			
MNEc.2.6.80820491.PC	80820491	81987526			
MNEc.2.6.80821993.PC	80821993	81989028			
MNEc.2.6.80824601.PC	80824601	81991636		X	
MNEc.2.6.80830385.PC	80830385	81997419			
MNEc.2.6.80830446.PC	80830446	81997480			
MNEc.2.6.80831282.PC	80831282	81998316			

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.80832084.PC	80832084	81999118			X
MNEc.2.6.80832207.PC	80832207	81999241			
MNEc.2.6.80833819.PC	80833819	82000854			
MNEc.2.6.80834462.PC	80834462	82001497			
MNEc.2.6.80834740.PC	80834740	82001775		X	
MNEc.2.6.80837012.PC	80837012	82004048			
MNEc.2.6.80837085.PC	80837085	82004121			
MNEc.2.6.80842128.PC	80842128	82009162			
MNEc.2.6.80842603.PC	80842603	82009637			
MNEc.2.6.80844424.PC	80844424	82011458		X	
MNEc.2.6.80844664.PC	80844664	82011698		X	
MNEc.2.6.80846341.PC	80846341	82013375			X
MNEc.2.6.80846756.PC	80846756	82013790			
MNEc.2.6.80848712.PC	80848712	82015746		X	
MNEc.2.6.80849310.PC	80849310	82016344			
MNEc.2.6.80849467.PC	80849467	82016501			
MNEc.2.6.80850749.PC	80850749	82017783			
MNEc.2.6.80851196	80851196	82018230			
MNEc.2.6.80855422.PC	80855422	82022458			
MNEc.2.6.80859332.PC	80859332	82026311			
MNEc.2.6.80859678.PC	80859678	82026657			
MNEc.2.6.80861680.PC	80861680	82028659			
MNEc.2.6.80862656.PC	80862656	82029635			
MNEc.2.6.80863356.PC	80863356	82030335			X
MNEc.2.6.80864497.PC	80864497	82031476			
MNEc.2.6.80865169.PC	80865169	82032148			
MNEc.2.6.80865774.PC	80865774	82032753		X	
MNEc.2.6.80865916.PC	80865916	82032895			
MNEc.2.6.80867332.PC	80867332	82034311			
MNEc.2.6.80867552.PC	80867552	82034531			
MNEc.2.6.80868281.PC	80868281	82035260			
MNEc.2.6.80868393.PC	80868393	82035372			
MNEc.2.6.80870148.PC	80870148	NA			
MNEc.2.6.80871158.PC	80871158	NA			
MNEc.2.6.80872239.PC	80872239	NA			
MNEc.2.6.80873525.PC	80873525	82037961			
MNEc.2.6.80877589.PC	80877589	82042029			
MNEc.2.6.80879383.PC	80879383	82043822			
MNEc.2.6.80879864.PC	80879864	82044303			
MNEc.2.6.80883323.PC	80883323	82047762			
MNEc.2.6.80885397.PC	80885397	82049836			X
MNEc.2.6.80887049.PC	80887049	82051488			
MNEc.2.6.80889551.PC	80889551	82053990			
MNEc.2.6.80890927.PC	80890927	82055366			
MNEc.2.6.80891737.PC	80891737	82056176			
MNEc.2.6.80893186.PC	80893186	82057625			
MNEc.2.6.80893468.PC	80893468	82057907			
MNEc.2.6.80900288.PC	80900288	82064728			
MNEc.2.6.80902137.PC	80902137	82066602			
MNEc.2.6.80902997.PC	80902997	82067462			
MNEc.2.6.80903581.PC	80903581	82068046			

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.80907156.PC	80907156	82071589			
MNEc.2.6.80911139.PC	80911139	82075571			
MNEc.2.6.80915201.PC	80915201	82079633			
MNEc.2.6.80915493.PC	80915493	82079925			
MNEc.2.6.80915581.PC	80915581	82080013		X	
MNEc.2.6.80917882.PC	80917882	82082314			
MNEc.2.6.80919604.PC	80919604	82084036			X
MNEc.2.6.80919807.PC	80919807	82084239			
MNEc.2.6.80920347.	80920347	82084779			
MNEc.2.6.80926065.PC	80926065	82090497			
MNEc.2.6.80928463.PC	80928463	82092895			
MNEc.2.6.80929172.PC	80929172	82093604			
MNEc.2.6.80931312.PC	80931312	82095763			
MNEc.2.6.80933091.PC	80933091	82097541			
MNEc.2.6.80933683.PC	80933683	82098133			
MNEc.2.6.80933998.PC	80933998	82098448			
MNEc.2.6.80935302.PC	80935302	82099752			
MNEc.2.6.80935813.PC	80935813	82100263		X	
MNEc.2.6.80937306.PC	80937306	82101756			
MNEc.2.6.80939857.PC	80939857	82104307			
MNEc.2.6.80940129.PC	80940129	82104579			
MNEc.2.6.80942635.PC	80942635	82107085			
MNEc.2.6.80944105.PC	80944105	82108555			
MNEc.2.6.80945451.PC	80945451	82109901			
MNEc.2.6.80948795.PC	80948795	82113239			
MNEc.2.6.80948817.PC	80948817	82113261			
MNEc.2.6.80950340.PC	80950340	82114784			X
MNEc.2.6.80958104.PC	80958104	82122537			
MNEc.2.6.80958180.PC	80958180	82122613			
MNEc.2.6.80959050.PC	80959050	82123482			
MNEc.2.6.80960411.PC	80960411	82124573			
MNEc.2.6.80960850.PC	80960850	82125012			
MNEc.2.6.80961489.PC	80961489	82125651		X	
MNEc.2.6.80964432.PC	80964432	82128595		X	
MNEc.2.6.80968716.PC	80968716	82132880			
MNEc.2.6.80969726.PC	80969726	82133890			
MNEc.2.6.80971088.PC	80971088	82135252		X	
MNEc.2.6.80971929.PC	80971929	82136093			
MNEc.2.6.80976137.PC	80976137	82140301		X	
MNEc.2.6.80976600.PC	80976600	82140764			
MNEc.2.6.80976721.PC	80976721	82140885			
MNEc.2.6.80977285.PC	80977285	82141449			
MNEc.2.6.80977751.PC	80977751	82141915			
MNEc.2.6.80979433.PC	80979433	82143597			
MNEc.2.6.80980590.PC	80980590	82144754			
MNEc.2.6.80981134.PC	80981134	82145298			
MNEc.2.6.80981470.PC	80981470	82145634			
MNEc.2.6.80981691.PC	80981691	82145855			
MNEc.2.6.80984945.PC	80984945	82149109			
MNEc.2.6.80985350.PC	80985350	82149514			
MNEc.2.6.80987560.PC	80987560	82151724			

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.80988998.PC	80988998	82153162			
MNEc.2.6.80989356.PC	80989356	82153520	X		
MNEc.2.6.80994247.PC	80994247	82158411	X		
MNEc.2.6.81000288.PC	81000288	82164452	X	X	
MNEc.2.6.81001292.	81001292	82165456	X		
MNEc.2.6.81001824.PC	81001824	82165988	X	X	
MNEc.2.6.81004640.PC	81004640	82168804	X	X	
MNEc.2.6.81004787.	81004787	82168951	X		
MNEc.2.6.81004845.PC	81004845	82169009	X		
MNEc.2.6.81005329.PC	81005329	82169493	X		
MNEc.2.6.81011612.PC	81011612	82175776		X	
MNEc.2.6.81011857.PC	81011857	82176021			
MNEc.2.6.81012766.PC	81012766	82176930			
MNEc.2.6.81014209.PC	81014209	82178374			
MNEc.2.6.81018173.PC	81018173	82182338			
MNEc.2.6.81024712.PC	81024712	82188877			
MNEc.2.6.81028211.PC	81028211	82192376			
MNEc.2.6.81028417.PC	81028417	82192582			
MNEc.2.6.81029280.PC	81029280	82193445			
MNEc.2.6.81030458.PC	81030458	82194623			
MNEc.2.6.81040675.PC	81040675	82204840			
MNEc.2.6.81040813.PC	81040813	82204978			
MNEc.2.6.81040860.PC	81040860	82205025		X	
MNEc.2.6.81041470.PC	81041470	82205635			
MNEc.2.6.81041826.PC	81041826	82205991			
MNEc.2.6.81044828.PC	81044828	82208993			
MNEc.2.6.81046539.PC	81046539	82210704			
MNEc.2.6.81047385.PC	81047385	82211550			
MNEc.2.6.81048067.PC	81048067	82212232			
MNEc.2.6.81050533.PC	81050533	82214698			
MNEc.2.6.81053351.PC	81053351	82217515			
MNEc.2.6.81056567.PC	81056567	82220731			
MNEc.2.6.81057502.PC	81057502	82221666			
MNEc.2.6.81060020.PC	81060020	82224184			
MNEc.2.6.81066572.PC	81066572	82230736		X	X
MNEc.2.6.81072276.PC	81072276	82236440			
MNEc.2.6.81074062.PC	81074062	82238227			
MNEc.2.6.81074150.PC	81074150	82238315			
MNEc.2.6.81074374.PC	81074374	82238539		X	
MNEc.2.6.81074650.PC	81074650	82238815		X	
MNEc.2.6.81077322.PC	81077322	82241487			
MNEc.2.6.81082719.PC	81082719	82246884		X	
MNEc.2.6.81084493.PC	81084493	82248658			
MNEc.2.6.81084746.PC	81084746	82248911			
MNEc.2.6.81085399.PC	81085399	82249564		X	X
MNEc.2.6.81086049.PC	81086049	82250226			
MNEc.2.6.81088906.PC	81088906	82253083			
MNEc.2.6.81089958.PC	81089958	82254135		X	X
MNEc.2.6.81091853.PC	81091853	82256030			
MNEc.2.6.81092504.PC	81092504	82256681			
MNEc.2.6.81095746.PC	81095746	82269987			X

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.81092730.PC	81092730	82256907			
MNEc.2.6.81096691.PC	81096691	82260932			
MNEc.2.6.81097039.PC	81097039	82261280		X	
MNEc.2.6.81099063.	81099063	82263304			X
MNEc.2.6.81100715.PC	81100715	82264956			
MNEc.2.6.81101410.PC	81101410	82265651		X	X
MNEc.2.6.81105042.PC	81105042	82269283			
MNEc.2.6.81107367.PC	81107367	82271611			
MNEc.2.6.81107765.PC	81107765	82272009			
MNEc.2.6.81108269.PC	81108269	82272513	X		
MNEc.2.6.81110483.PC	81110483	82274730	X		
MNEc.2.6.81111923.PC	81111923	82276170	X		
MNEc.2.6.81113370.PC	81113370	82277617	X		
MNEc.2.6.81114838.PC	81114838	82279085	X		
MNEc.2.6.81116592.PC	81116592	82280839	X		
MNEc.2.6.81116663.PC	81116663	82280910	X	X	
MNEc.2.6.81117653.PC	81117653	82281900	X		
MNEc.2.6.81127246.PC	81127246	82291777	X		
MNEc.2.6.81131003.PC	81131003	82295534	X	X	
MNEc.2.6.81132667.PC	81132667	82297198	X		
MNEc.2.6.81138611.PC	81138611	82303142	X		
MNEc.2.6.81145454.PC	81145454	82309986	X		
MNEc.2.6.81146607.PC	81146607	82311139	X		
MNEc.2.6.81147917.PC	81147917	82312449	X	X	
MNEc.2.6.81148466.PC	81148466	82312998	X		
MNEc.2.6.81148841.PC	81148841	82313373	X		
MNEc.2.6.81149038.PC	81149038	82313570	X	X	
MNEc.2.6.81150141.PC	81150141	82314673	X		
MNEc.2.6.81150674.PC	81150674	82315206	X		
MNEc.2.6.81150862.PC	81150862	82315394	X		
MNEc.2.6.81151323.PC	81151323	82315855	X	X	
MNEc.2.6.81152002.PC	81152002	82316534	X	X	
MNEc.2.6.81152832.PC	81152832	82317364	X		
MNEc.2.6.81154362.PC	81154362	82318894	X		
MNEc.2.6.81155025.PC	81155025	82319557	X	X	
MNEc.2.6.81155688.PC	81155688	82320220	X		
MNEc.2.6.81155869.PC	81155869	82320401	X	X	
MNEc.2.6.81156390.PC	81156390	82320922	X		
MNEc.2.6.81156975.PC	81156975	82321507	X		
MNEc.2.6.81157724.PC	81157724	82322256	X		
MNEc.2.6.81159639.PC	81159639	82324171			
MNEc.2.6.81160030.PC	81160030	82324562			
MNEc.2.6.81161980.PC	81161980	82326512			
MNEc.2.6.81167162.PC	81167162	82331694		X	
MNEc.2.6.81168395.PC	81168395	82332927			
MNEc.2.6.81168833.PC	81168833	82333365	X	X	
MNEc.2.6.81171121.PC	81171121	82335653	X	X	
MNEc.2.6.81175201.PC	81175201	82339734	X	X	
MNEc.2.6.81176493.	81176493	82341026	X		
MNEc.2.6.81176653.	81176653	82341186	X		
MNEc.2.6.81176905.PC	81176905	82341438	X		

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.81181575.PC	81181575	82346108	X		
MNEc.2.6.81183704.PC	81183704	82348237		X	
MNEc.2.6.81185362.PC	81185362	82349895		X	
MNEc.2.6.81193866.PC	81193866	82358399			
MNEc.2.6.81196676.PC	81196676	82361209		X	
MNEc.2.6.81197672.PC	81197672	82362205			
MNEc.2.6.81198310.PC	81198310	82362843		X	
MNEc.2.6.81198765.PC	81198765	82363298			
MNEc.2.6.81199128.PC	81199128	82363661			
MNEc.2.6.81199762.PC	81199762	82364295			
MNEc.2.6.81200192.PC	81200192	82364725			
MNEc.2.6.81200578.PC	81200578	82365111		X	
MNEc.2.6.81200658.PC	81200658	82365191			
MNEc.2.6.81200913.PC	81200913	82365446			
MNEc.2.6.81201314.PC	81201314	82365847			
MNEc.2.6.81201348.PC	81201348	82365881		X	
MNEc.2.6.81201373.PC	81201373	82365906			
MNEc.2.6.81203959.PC	81203959	82368492		X	
MNEc.2.6.81205065.PC	81205065	82369598			
MNEc.2.6.81205686.PC	81205686	82370219			
MNEc.2.6.81210715.PC	81210715	82375248			
MNEc.2.6.81213956.PC	81213956	82378489			
MNEc.2.6.81215555.PC	81215555	82380088		X	
MNEc.2.6.81216957.	81216957	82381490		X	
MNEc.2.6.81218021.PC	81218021	82382554			
MNEc.2.6.81218092.PC	81218092	82382625			
MNEc.2.6.81218897.PC	81218897	82383430			
MNEc.2.6.81224318.PC	81224318	82388851			
MNEc.2.6.81227730.PC	81227730	82392263	X		
MNEc.2.6.81231288.PC	81231288	82395821	X		
MNEc.2.6.81231316.PC	81231316	82395849	X		
MNEc.2.6.81235378.PC	81235378	82399911	X	X	
MNEc.2.6.81236569.PC	81236569	82401102	X	X	
MNEc.2.6.81237287.PC	81237287	82401820	X	X	
MNEc.2.6.81242767.PC	81242767	82407300			
MNEc.2.6.81242979.PC	81242979	82407512			
MNEc.2.6.81245868.PC	81245868	82410401			
MNEc.2.6.81246188.PC	81246188	82410721			
MNEc.2.6.81250359.PC	81250359	82414892		X	
MNEc.2.6.81250799.PC	81250799	82415332		X	
MNEc.2.6.81252426.PC	81252426	82416959			
MNEc.2.6.81254844.PC	81254844	82419377			
MNEc.2.6.81257395.PC	81257395	82421926			
MNEc.2.6.81260218.PC	81260218	82424749			
MNEc.2.6.81260989.PC	81260989	82425520		X	
MNEc.2.6.81261983.PC	81261983	82426514			
MNEc.2.6.81263291.PC	81263291	82427823			
MNEc.2.6.81264617.PC	81264617	82429149			
MNEc.2.6.81265503.PC	81265503	82430033			
MNEc.2.6.81265743.PC	81265743	82430273			
MNEc.2.6.81265835.PC	81265835	82430365		X	

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.81274342.PC	81274342	82438510	X	X	
MNEc.2.6.81275539.PC	81275539	82439707	X		X
MNEc.2.6.81275716.PC	81275716	82439884	X		
MNEc.2.6.81280716.PC	81280716	82444884	X		
MNEc.2.6.81286888.PC	81286888	82451056	X	X	
MNEc.2.6.81288528.PC	81288528	82452696	X	X	
MNEc.2.6.81291848.PC	81291848	82456016	X		X
MNEc.2.6.81293583.PC	81293583	82457751	X	X	
MNEc.2.6.81300883.PC	81300883	82465051	X		
MNEc.2.6.81305201.PC	81305201	82469369			
MNEc.2.6.81307622.PC	81307622	82471790	X		
MNEc.2.6.81308993.PC	81308993	82473161	X		
MNEc.2.6.81309029.PC	81309029	82473197	X	X	
MNEc.2.6.81309194.PC	81309194	82473362	X	X	
MNEc.2.6.81309349.PC	81309349	82473517	X	X	X
MNEc.2.6.81310074.PC	81310074	82474242	X		
MNEc.2.6.81312462.PC	81312462	82476630	X		
MNEc.2.6.81312691.PC	81312691	82476859	X		
MNEc.2.6.81312805.PC	81312805	82476973	X		
MNEc.2.6.81313479.PC	81313479	82477647	X	X	
MNEc.2.6.81320405.PC	81320405	82484592	X	X	
MNEc.2.6.81322229.PC	81322229	82486416	X	X	X
MNEc.2.6.81329646.PC	81329646	82493449	X		
MNEc.2.6.81333137.PC	81333137	82496940		X	
MNEc.2.6.81333372.PC	81333372	82497175			X
MNEc.2.6.81340647.PC	81340647	82504450	X	X	
MNEc.2.6.81340998.PC	81340998	80204801	X		X
MNEc.2.6.81342894.PC	81342894	82506697	X		
MNEc.2.6.81344532.PC	81344532	82508335	X		X
MNEc.2.6.81348453.PC	81348453	82512256	X		X
MNEc.2.6.81347830.PC	81347830	82511633	X		
MNEc.2.6.81347974.PC	81347974	82511777	X		
MNEc.2.6.81352276.PC	81352276	82516073	X		
MNEc.2.6.81354701.PC	81354701	82518498			
MNEc.2.6.81356993.PC	81356993	82520790		X	
MNEc.2.6.81358887.PC	81358887	82522684			
MNEc.2.6.81361483.PC	81361483	82525280	X	X	
MNEc.2.6.81361520.PC	81361520	82525317	X		
MNEc.2.6.81365395.PC	81365395	82529192	X	X	
MNEc.2.6.81367697.PC	81367697	82531494	X		
MNEc.2.6.81368345.PC	81368345	82532142	X	X	
MNEc.2.6.81381221.PC	81381221	82545021	X	X	
MNEc.2.6.81382533.PC	81382533	82546333	X		X
MNEc.2.6.81392217.PC	81392217	82555953	X		
MNEc.2.6.81392910.PC	81392910	82556646		X	
MNEc.2.6.81396654.PC	81396654	82560390			
MNEc.2.6.81398600.PC	81398600	82562336			
MNEc.2.6.81400279.PC	81400279	82564015		X	X
MNEc.2.6.81405725.PC	81405725	82569461	X	X	X
MNEc.2.6.81407183.PC	81407183	82570919	X	X	
MNEc.2.6.81408708.PC	81408708	82572444	X		

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.81409678.PC	81409678	82573414	X		
MNEc.2.6.81411449.PC	81411449	82575185	X	X	
MNEc.2.6.81413254.PC	81413254	82576990	X	X	X
MNEc.2.6.81414435.PC	81414435	82578171	X		
MNEc.2.6.81415276.PC	81415276	82579013	X		
MNEc.2.6.81417632.	81417632	82581369	X		
MNEc.2.6.81421330.PC	81421330	82585066	X	X	X
MNEc.2.6.81431568.PC	81431568	82595304	X		X
MNEc.2.6.81432131.PC	81432131	82595867	X		X
MNEc.2.6.81434058.PC	81434058	82597794	X	X	
MNEc.2.6.81440665.PC	81440665	82604401	X	X	
MNEc.2.6.81441526.PC	81441526	82605262	X		
MNEc.2.6.81443295.PC	81443295	82607030	X		X
MNEc.2.6.81446353.PC	81446353	82610088	X		X
MNEc.2.6.81451782.	81451782	82615517	X	X	X
MNEc.2.6.81458759.PC	81458759	82622494	X	X	X
MNEc.2.6.81463114.PC	81463114	82626849	X	X	X
MNEc.2.6.81468176.PC	81468176	82631911	X		
MNEc.2.6.81469661.PC	81469661	82633396	X		
MNEc.2.6.81468256.PC	81468256	82631991	X		X
MNEc.2.6.81471494.PC	81471494	82635229	X		
MNEc.2.6.81473575.PC	81473575	82637310	X		X
MNEc.2.6.81474930.PC	81474930	82638665	X	X	
MNEc.2.6.81475049.PC	81475049	82638784	X	X	
MNEc.2.6.81476437.PC	81476437	82640172	X	X	X
MNEc.2.6.81481065.PC	81481065	82644800			X
MNEc.2.6.81482862.PC	81482862	82646597			
MNEc.2.6.81488131.PC	81488131	82651920		X	
MNEc.2.6.81490123.PC	81490123	82653915	X	X	X
MNEc.2.6.81492423.PC	81492423	82656215	X	X	X
MNEc.2.6.81494335.PC	81494335	82658127	X	X	
MNEc.2.6.81497380.PC	81497380	82661169	X	X	X
MNEc.2.6.81498052.PC	81498052	82661841	X	X	
MNEc.2.6.81503348.PC	81503348	82667137	X		
MNEc.2.6.81503730.PC	81503730	82667518	X		X
MNEc.2.6.81505709.PC	81505709	82669497	X	X	X
MNEc.2.6.81506349.PC	81506349	82670137	X		
MNEc.2.6.81507212.PC	81507212	82671000	X	X	X
MNEc.2.6.81507310.PC	81507310	82671098	X	X	X
MNEc.2.6.81508624.PC	81508624	82672412	X		
MNEc.2.6.81509429.	81509429	82673217	X		
MNEc.2.6.81510788.PC	81510788	82674576	X		
MNEc.2.6.81514907.PC	81514907	82678695	X	X	X
MNEc.2.6.81516706.	81516706	82680494	X	X	
MNEc.2.6.81521876.PC	81521876	82685665	X	X	
MNEc.2.6.81523773.PC	81523773	82687562	X	X	X
MNEc.2.6.81523837.PC	81523837	82687626	X		X
MNEc.2.6.81526828.PC	81526828	82690617	X	X	X
MNEc.2.6.81527285.PC	81527285	82691074	X		
MNEc.2.6.81528014.PC	81528014	82691803	X	X	
MNEc.2.6.81532654.PC	81532654	82696455	X		X

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.81533130.PC	81533130	82696931	X	X	
MNEc.2.6.81533415.PC	81533415	82697216	X	X	X
MNEc.2.6.81533975.PC	81533975	82697776	X	X	X
MNEc.2.6.81535308.PC	81535308	82699109	X		
MNEc.2.6.81536480.PC	81536480	82700281	X	X	
MNEc.2.6.81536796.PC	81536796	82700597	X		
MNEc.2.6.81541746.PC	81541746	82705547		X	
MNEc.2.6.81551984.PC	81551984	82715786	X	X	
MNEc.2.6.81553061.PC	81553061	82716863	X		X
MNEc.2.6.81554945.PC	81554945	82718747	X		
MNEc.2.6.81557294.	81557294	82721096	X	X	X
MNEc.2.6.81558781.PC	81558781	82722583	X		
MNEc.2.6.81560279.PC	81560279	82724081	X		
MNEc.2.6.81566120.PC	81566120	82729921	X	X	
MNEc.2.6.81568749.PC	81568749	82732548	X	X	X
MNEc.2.6.81575176.PC	81575176	82738976	X		
MNEc.2.6.81575713.PC	81575713	82739513	X	X	
MNEc.2.6.81576419.PC	81576419	82740219	X		
MNEc.2.6.81576767.PC	81576767	82740567	X		
MNEc.2.6.81577868.PC	81577868	82741668	X		
MNEc.2.6.81583349.PC	81583349	82747149			
MNEc.2.6.81583507.PC	81583507	82747307			
MNEc.2.6.81585047.PC	81585047	82748847			
MNEc.2.6.81589592.PC	81589592	82753392	X		
MNEc.2.6.81590012.PC	81590012	82753812	X		
MNEc.2.6.81591558.PC	81591558	82755358	X	X	
MNEc.2.6.81591919.PC	81591919	82755719	X		
MNEc.2.6.81600981.PC	81600981	82764783	X		
MNEc.2.6.81602184.PC	81602184	82765986	X		
MNEc.2.6.81602630.	81602630	82766432	X		
MNEc.2.6.81602938.PC	81602938	82766740	X	X	
MNEc.2.6.81603378.PC	81603378	82767180	X		
MNEc.2.6.81605181.PC	81605181	82768983	X	X	
MNEc.2.6.81605475.PC	81605475	82769277	X		
MNEc.2.6.81612750.PC	81612750	82776553	X		
MNEc.2.6.81614184.PC	81614184	82777987	X		
MNEc.2.6.81614934.PC	81614934	82778737	X		
MNEc.2.6.81615849.PC	81615849	82779652	X	X	
MNEc.2.6.81624548.PC	81624548	82788363	X		
MNEc.2.6.81625635.PC	81625635	82789450	X		
MNEc.2.6.81626239.PC	81626239	82790054	X	X	
MNEc.2.6.81634717.PC	81634717	82798532	X	X	
MNEc.2.6.81635994.PC	81635994	82799809	X		
MNEc.2.6.81637171.PC	81637171	82800986	X		
MNEc.2.6.81643314.	81643314	82807129	X		X
MNEc.2.6.81647854.PC	81647854	82811669	X		
MNEc.2.6.81649941.PC	81649941	82813756	X		
MNEc.2.6.81653476.PC	81653476	82817291			X
MNEc.2.6.81663656.PC	81663656	82827472			
MNEc.2.6.81664538.PC	81664538	82828354			
MNEc.2.6.81666343.PC	81666343	82830159		X	

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.81666520.PC	81666520	82830336			
MNEc.2.6.81669670.PC	81669670	82833486		X	
MNEc.2.6.81671214.PC	81671214	82835030	X		
MNEc.2.6.81694523.PC	81694523	82858459	X		
MNEc.2.6.81695495.PC	81695495	82859431	X	X	
MNEc.2.6.81695968.PC	81695968	82859904	X		
MNEc.2.6.81696642.PC	81696642	82860578	X		
MNEc.2.6.81700931.PC	81700931	82864787			
MNEc.2.6.81703726.PC	81703726	82867582			
MNEc.2.6.81705123.PC	81705123	82868979			
MNEc.2.6.81707385.PC	81707385	82871241	X	X	
MNEc.2.6.81710589.PC	81710589	82874445	X		
MNEc.2.6.81711770.PC	81711770	82875626	X		
MNEc.2.6.81715109.PC	81715109	82878965	X		
MNEc.2.6.81720090.PC	81720090	82883946	X		
MNEc.2.6.81722649.PC	81722649	82886505	X	X	
MNEc.2.6.81722944.PC	81722944	82886800	X		
MNEc.2.6.81725228.PC	81725228	82889084	X		
MNEc.2.6.81727663.PC	81727663	82891519	X		
MNEc.2.6.81728172.PC	81728172	82892028	X		
MNEc.2.6.81735717.PC	81735717	82899573	X		
MNEc.2.6.81740361.PC	81740361	82904217	X		
MNEc.2.6.81742978.PC	81742978	82906837	X		
MNEc.2.6.81749291.PC	81749291	82913150	X		
MNEc.2.6.81752110.PC	81752110	82915969	X	X	
MNEc.2.6.81753106.PC	81753106	82916923	X		
MNEc.2.6.81753656.PC	81753656	82917473	X		
MNEc.2.6.81755922.PC	81755922	82919739	X	X	
MNEc.2.6.81759471.PC	81759471	82923288	X		
MNEc.2.6.81764808.PC	81764808	82928625	X		
MNEc.2.6.81764849.PC	81764849	82928666	X	X	
MNEc.2.6.81766028.PC	81766028	82929796	X		
MNEc.2.6.81774220.PC	81774220	82937922	X		
MNEc.2.6.81774286.PC	81774286	82937988	X		
MNEc.2.6.81777995.PC	81777995	82941697	X		
MNEc.2.6.81780992.PC	81780992	82944695	X	X	
MNEc.2.6.81782298.PC	81782298	82946001	X		
MNEc.2.6.81784570.PC	81784570	82948273	X		
MNEc.2.6.81785714.PC	81785714	82949417	X		
MNEc.2.6.81788701.PC	81788701	82952403	X		
MNEc.2.6.81789212.PC	81789212	82952914	X		
MNEc.2.6.81791707.PC	81791707	82955409			
MNEc.2.6.81793853.PC	81793853	82957368			
MNEc.2.6.81795218.PC	81795218	82958733			
MNEc.2.6.81795964.PC	81795964	82959479			
MNEc.2.6.81796099.PC	81796099	82959614			
MNEc.2.6.81799715.PC	81799715	82963230			
MNEc.2.6.81801026.PC	81801026	82964541			
MNEc.2.6.81802027.PC	81802027	82965542			
MNEc.2.6.81802264.PC	81802264	82965779			
MNEc.2.6.81804903.PC	81804903	82968419			

Axiom MCEc2M SNP ID	EquCab2 BP Location	EquCab3 BP Location	Di	Height Assoc.	Insulin Assoc.
MNEc.2.6.81806445.PC	81806445	82969961			
MNEc.2.6.81808008.PC	81808008	82971523		X	
MNEc.2.6.81809066.PC	81809066	82972582			

Supplemental Table D6: EquCab2 and EquCab3 base pair (bp) position for SNPs on the Axiom MCEc2M within the region of interest on equine chromosome 6 (ECA6) bp positions 80,499,826-81,809,066. SNPs (presented by their Axiom MCEc2M SNP ID) within the entire region of interest were remapped to EquCab3.[427] EquCab3 coordinates were not provided for three SNPs as they did not have probes that mapped uniquely to EquCab3. SNPs which exceeded the threshold for genome wide significance on association analysis (Assoc) for height and baseline insulin are indicated by an X. Significant di windows are based on the average base pair position within a 10Kb window of SNPs. SNPs marked with an X represent 5Kb upstream and 5Kb downstream of the base pair location.

Breed	INS	INS-OST	GLU	GLU-OST	NEFA	TG	LEPTIN	ADIPON	ACTH
All Horses n=830	**-0.15 (-0.21,-0.08) p=<0.001**	-0.05 (-0.13,0.02) p=0.15	**-0.14 (-0.21,-0.08) p=<0.001**	**-0.10 (-0.18,-0.03) p=0.01**	-0.06 (-0.13,0.01) p=0.07	**0.12 (0.06,0.19) p=<0.001**	**0.12 (0.05,0.18) p=<0.001**	**-0.22 (-0.21,0.15) p=<0.001**	**-0.13 (-0.19,-0.06) p=<0.001**
Ponies n=301	**-0.33 (-0.42,-0.29) p=<0.001**	-0.15 (-0.26,-0.04) p=0.01	-0.14 (-0.25,-0.03) p=0.015	-0.02 (-0.13,0.1) p=0.75	-0.12 (-0.23,0.0) p=0.04	-0.14 (-0.09,-0.03) p=0.013	-0.09 (-0.20,-0.02) p=0.12	-0.07 (-0.05,-0.18) p=0.24	-.013 (-0.24,-0.02) p=0.02

Supplemental Table D7: Correlations between height and biochemical traits with the addition of seven ponies. Pearson's correlation coefficients were repeated with the inclusion of seven ponies representing three Shetland ponies, two Hackney ponies, and three British Riding ponies. Presented in the table are: Pearson's correlation coefficients, 95% confidence intervals and p-values for height, eight EMS biochemical traits, and ACTH for the entire cohort as well as just the ponies. All traits were corrected for age and sex prior to analysis. Significant p-values (<.0056) are in bolded text. Abbreviations: INS = insulin, INS-OST = insulin post oral sugar test, GLU = glucose, GLU-OST = glucose post oral sugar test, NEFA = non-esterified fatty acids, TG = triglycerides, ADIPON = adiponectin.