# The Major Histocompatibility Complex of the Turkey

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Lee D. Chaves

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

Kent M. Reed, Ph.D. Advisor

May 2010

## Acknowledgements

I wish to thank all the faculty, students and staff members at the University of Minnesota and othe institutions for their knowledge and assistance throughout the years of my graduate education.

My Advisor: Dr. Kent Reed.
Thank you for the years of dedication and patience to have me as a student. As a young, stary eyed undergraduate Iquickly understood how much I could learn working for you, and by the time I recievied my Ph. D. I finally knew what that was.

My Committee.
I deeply appreciate all if the distinct insites you have provided me throughout the years in our correspondence and conversations.

My Fellow Students.
I have enjoyed all the activities and suggestions you have provided during my thesis work at the University.

My Friends and Family.
Thanks for the years of support and encouragement. You may not have always understood what I was doing, but you always knew why.

This research was supported by grants from the University of Minnesota Agriculture Experiment Station and the Cooperative State Research, Education and Extension Service, U.S. Department of Agriculture (2004-35205-14217).

## Abstract

The ability to identify between self and foreign pathogens is a key function of the vertebrate immune system. To achieve this, vertebrates have developed complex genetic and epigenetic mechanisms to provide sufficient variability to respond to continuously evolving pathogens or transformed cells they face. One of these systems is the large and highly polymorphic major histocompatibility complex (MHC), responsible for the presentation of endogenous and exogenous peptide antigens to T cells. The degree of polymorphism, number of gene copies, and the co-dominant expression of genes of the mammalian MHC allows for a great number possible antigens to be presented. Compared to mammals, the chicken MHC (MHC-B) is greatly condensed, containing the major histocompatibility antigens in a region of just 50 kb with a reduction in gene copy numbers and a lack of co-dominant expression. In addition to the MHC-B, the chicken has a second MHC-like region (MHC-Y) located on the same microchromosome as B yet it is genetically unlinked. The work presented in this dissertation physically and genetically maps the homologous MHC regions in the turkey and identifies genes within these regions. Further work surveyed the polymorphism content within commercial and wild turkeys. A final study assayed the genome-wide diversity of the individual from which the MHC was sequenced to determine the appropriateness of this source DNA for whole genome sequencing. These data provide helpful background information to advance turkey whole genome sequencing and develops genomic resources for the study of the effect of MHC alleles on the outcome of pathogenic infections in the turkey.

**Table of Contents**

Page
i Acknowledgements
ii Abstract
iii Table of Contents
v List of Tables
vii List of Figures

**1 Chapter 1: Introduction and Literature Review**
2 Discovery of the Major Histocompatibility Complex
4 The Major Histocompatibility Complex Locus
5 The Function of Histocompatibility Molecules
7 Major Histocompatibility Alleles Can Protect or Promote Disease
8 The Avian Major Histocompatibility Complex
11 The Turkey MHC
13 Why Turkeys?
14 Research Overview

16 **Chapter 2: Characterization of the Turkey MHC Chromosome through Genetic and Physical Mapping**
17 Title Page
19 Introduction
21 Materials and Methods
24 Results
29 Discussion

**40 Chapter 3: Defining the Turkey MHC: Sequence and Genes of the B-Locus**
41 Title Page
42 Introduction
44 Materials and Methods
49 Results
58 Discussion

**97 Chapter 4: Initial Sequence Analysis of the Turkey MHC-Y Locus Supports Ancient Origins of Immunological Gene Clusters**
98 Title Page
99 Introduction
101 Materials and Methods
102 Results
106 Discussion

**123 Chapter 5: Haplotype Variation, Recombination, and Gene Conversion within the Turkey MHC-B**
124 Title Page

125 Introduction
127 Materials and Methods
128 Results
132 Discussion

**160 Chapter 6: A Locus-Wide Approach to Assessing Variation in the Avian MHC: The B-Locus of the Wild Turkey**
161 Title Page
162 Introduction
165 Materials and Methods
166 Results
169 Discussion

**196 Chapter 7: Genome-Wide Genetic Diversity of "Nici", the DNA Source for the CHORI-260 Turkey BAC Library and Candidate for Whole Genome Sequencing**
197 Title Page
198 Summary
165 Microsatellite Variation
166 Sequence Variation
169 Discussion

**211 Chapter 8: Conclusions and Future Studies**

**218 Reference List**

**235 Appendix: Approval for inclusion of previously published materials in thesis by publishers**

## List of Tables


**Chapter 2:**
Page 32 Table 1: MHC Loci Amplified in the Turkey
Page 33 Table 2: Overlapping Oligonucleotides Used to Generate Overgo Probes Used in CHORI 260 Library Array Hybridization
Page 34 Table 3: Blast Analysis of Representative Subclone Sequences from MHC BAC Clones

**Chapter 3:**
Page 67 Table 1: PCR and internal sequencing primers and amplicon size of the 3` sequenced turkey MHC-B region.
Page 68 Table 2: Coding Sequence Comparisons of Homologous Turkey and Chicken MHC-B Loci
Page 69 Table 3: Selective Pressures Identified Through Pairwise dN/dS on Class IIβ Loci Varies Between Species and Loci
Page 70 Table 4: Turkey Class IIβ β1 Domain and Expression Variation

**Chapter 4:**
Page 112 Table 1: Loci Identified in the BAC Clone 159B18

**Chapter 5:**
Page 138 Table 1: Primers and Reaction Conditions for Loci Examined in Commercial Breeder Birds
Page 139 Table 2: Polymorphisms Identified in Commercial Turkeys
Page 152 Table 3: Single Nucleotide Polymorphism Characterizations
Page 153 Table 4: Haplotypes Derived from 52 Individuals and 11 Fully Defined SNPs.
Page 154 Table 5: FastPHASE Reconstructed Haplotypes from 49 SNP Loci

**Chapter 6:**
Page 174 Table 1: Subpopulations and Regional Representatives Examined in the Current Study
Page 175 Table 2: Primers and Reaction Conditions for Loci Examined in Wild Turkeys
Page 176 Table 3: Polymorphisms Identified in Wild Turkeys
Page 187 Table 4: Characterization of Single Nucleotide Polymorphisms Identified Within Wild Turkey Subpopulations
Page 188 Table 5: MHC haplotypes in Wild Turkey Inferred from 37 SNPs and DIPs
Page 190 Table 6: Local Comparison for Sample Locations with 4 or More Individuals

| | | |
|---|---|---|
| **Chapter 7:** | Page 205 | Table 1: Number of Alleles Identified at Microsatllite Loci in the Turkey |
| | Page 208 | Table 2: Summary of Heterozygosity Scores Determined by Genotyping 78 Microsatellite Markers in 43 Individuals |
| | Page 209 | Table 3: Summary of Individual Heterozygosities Determined by Genotyping 147 Microsatellite Markers Across the P Generation and F1 Males of the UMN/NTBF Population and Nici |

## List of Figures


| | | |
|---|---|---|
| **Chapter 1:** | Page 8 | Figure 1: The Chicken MHC Chromosome |
| **Chapter 2:** | Page 35 | Figure 1: Silver-Stained Turkey Metaphase Chromosomes |
| | Page 36 | Figure 2: Dual Color FISH of BAC Clones to Turkey Metaphase Chromosomes |
| | Page 37 | Figure 3: FISH Of Mgasat2 Repetitive DNA Telomere Repeat and rDNA to Turkey Metaphase Chromosomes |
| | Page 38 | Figure 4: Hybridization of BAC Clone 97E05 and Internal Telomere Repeat to the Same Arm of the Microchromosome |
| | Page 39 | Figure 5: Current Model of the Turkey MHC Chromosome Based on Genetic and Physical Mapping |
| **Chapter 3:** | Page 71 | Figure 1: Sequence Features of the Turkey MHC-B Region |
| | Page 73 | Figure 2: Comparative Idenitity Matrix Plot of the Turkey and Chicken MHC-B Loci |
| | Page 74 | Figure 3: Genbank Accession File of the Turkey MHC-B Region |
| | Page 89 | Figure 4: Phylogenetic Analysis of the Major Avian MHC Antigen Gene Coding Sequences |
| | Page 90 | Figure 5: Phylogenetic Analysis if Class IIβ β1 Sequence from Turkey, Chicken, and Quail |
| | Page 91 | Figure 6: Semi-Quantitative PCR Analysis of Turkey MHC Class I Loci Identified in the OCB02 Bird |
| | Page 92 | Figure 7: Class I Gene Promoters |
| | Page 93 | Figure 8: Southern Hybridization of BAC and Genomic DNA with a Class IIβ Probe |
| | Page 94 | Figure 9: Class IIβ Loci in Turkey were Generated from a Primordial Rearrangement of a Chicken-Like MHC-B Locus |
| | Page 95 | Figure 10: Class IIβ OCB02 Spleen RT-PCR and Genomic Sequencing Results |
| | Page 96 | Figure 11: Class IIβ Promoter Regions Share Significant Homology between Species and Loci |
| **Chapter 4:** | Page 113 | Figure 1: Contig Assemblies of BAC Clone 159B18 |
| | Page 114 | Figure 2: Genbank Accession of Clone 159B18 SC1 |
| | Page 117 | Figure 3: Dot Homology Plot of SC1 Against Itself Illustrates Local Areas of Repetitive Alignments |
| | Page 118 | Figure 4: Southern Hybridization Identifies a Single Repetitive Array of MgaSat2 in Clone 159B18 |

| | | |
|---|---|---|
| | Page 119 | Figure 5: Representative Example of GC Rich Islands within the 3` End of SC1 |
| | Page 120 | Figure 6: Comparison of 2 Turkey and 4 Chicken MHC-Y Class I Loci Indicate the Single Class I Locus Identified in 159B18 is a Pseudogene |
| | Page 122 | Figure 7: Model of MHC Chromosomal Divergence between Mammals and Avians from a Hypothetical Reptilian Progenitor |
| **Chapter 5:** | Page 155 | Figure 1: The Turkey MHC-B Locus and Amplicons Examined in this Study |
| | Page 156 | Figure 2: Minor Allele Frequencies of B-Locus SNPs Observed in Commercial Turkeys |
| | Page 157 | Figure 3: Proportion of Haplotypes Derived from Analysis of 11 “Minimal Shared” SNPs |
| | Page 158 | Figure 4: Phylogenetic Relationships Amoung the 45 Fastphase Assembled Turkey MHC-B Haplotypes |
| | Page 159 | Figure 5: Genetic Map of the Turkey MHC-B Locus |
| **Chapter 6:** | Page 191 | Figure 1: Historical Geographical Locations of Wild Turkey Subpopulations |
| | Page 192 | Figure 2: Minor Allele Frequencies of SNPs and DIPs Identified Wild Turkeys |
| | Page 193 | Figure 3: Allele Frequencies from 45 Loci Used in Haplotype Inference |
| | Page 194 | Figure 4: The Limited Extent of Linkage Disequilibrium across the MHC-B Region Identified from Wild Turkeys |
| | Page 195 | Figure 5: Phylogenetic Relationships of Inferred Haplotypes from Wild and Selected Commercial Turkeys |
| **Chapter 7:** | Page 210 | Figure 1: Heterozygosity Comparisons of the Turkey Macrochromosomes of P Generation, F1 Hybrids, and Nici |

# Chapter 1

# Introduction and Literature Review

"The discovery of the phenomenon of linkage, and the development of its possibilities now being realized, indicate a shorter and more subtle method of attack upon some of the genetic problems that are uppermost in the mind of the man who makes his living out of poultry... if we can, through linkage, relate the invisible with the visible, the task of analysis becomes immeasurably lightened, and the possibility of control is brought appreciably nearer" R.C. Punnett (1).

*Discovery of the Major Histocompatibility Complex*

The major histocompatibility complex (MHC) is the most polymorphic and gene dense region in the vertebrate genome (2). Many of the genes of the MHC have immune-related functions; in particular they function to provide the organism a way to distinguish between self, non-self, and intracellular infection. Immunologists first identified the region due to its dramatic effects on graft rejection and extensive work has characterized the region in humans and mice. Early on, tissue/tumor graft rejection was identified to be strongly regulated by a genetic histocompatibility (MHC Class I) locus in mice, H-2 (3).

The human leukocyte antigen (HLA) complex was more serendipitously identified by antibodies against alloantigens generated in blood transfusions. Simultaneously, MHC Class II genes were identified and mapped to the same region due to their regulatory role in the expansion and control of the immune defense. In 1980 George D. Snell, Jean Dausset, and Baruj Benacerraf were awarded the Nobel Prize in Physiology or Medicine "for their discoveries concerning genetically determined structures on the cell surface that regulate immunological reactions" (4-6).

The recognition of the importance of the MHC was tempered by the fact that the natural function of the MHC was still not fully understood. At the time, it was generally understood that the MHC must have a different purpose in nature as tissue transplantation is not a naturally occurring phenomenon. The role of MHC molecules to present antigens to T-cells was clearly articulated, however, the molecular interactions of antigen presentation were not well defined. The Nobel Prize in Physiology or Medicine for 1996 was awarded to Peter C. Doherty and Rolf M. Zinkernagel, for their discoveries concerning "the specificity of the cell mediated immune defense", identifying that T-cells require pathogen-specific antigens and self-specific MHC molecules for proper effector cell function (7,8). The genetic mechanism providing the ability of an organism to generate a diverse repertoire of T-cell receptors, able to recognize a nearly unlimited set of antigenic peptides bound to classical MHC molecules, was identified by Chien et al (9) in 1984. The mechanism of somatic recombination within the T-cell receptor loci was remarkably similar to that previously identified in the paralogous B-cell receptor loci. For this discovery in B-cells, Susumu Tonegawa was awarded the Nobel Prize in Physiology or Medicine for 1987 (10).

*The Major Histocompatibility Complex Locus*

In addition to humans and mice, the MHC region of several species of economic and agricultural importance including cattle, pigs, horses, and chickens have been studied (11,12). Mammalian MHCs vary in size but generally are considered to be between 3-5 Mb (11). However, based on conserved interspecies synteny and intraspecies linkage disequilibrium, an extended map of the human MHC (HLA) has been generated spanning a length of 7.6 Mb (13). Many of the genes within this region are associated with processing and presentation of antigens to T cells and provide a form of immune surveillance. These genes include the transporter associated with antigen processing (TAP and TAPBP) genes and low molecular weight proteosome genes (LMP), and classical and non-classical MHC Class Iα and Class IIα and β genes, respectively (13-15). Additional unique features of the extended human MHC are the large, super-clusters of genes encoding histones, tRNAs, butyrophilins, olfactory receptors, tripartite motif (TRIM) containing proteins, and zinc finger proteins (13).

The mammalian MHC appears to have several highly conserved features such as overall gene content and organization, demonstrating the evolutionary significance of gene retention in this region. Mammalian MHCs are generally subdivided into three regions. The class I region contains mostly genes encoding classical and non-classical MHC Class Iα molecules. The class II region contains many genes for classical and non-classical MHC Class IIα and β molecules, as well as many of the antigen processing genes present in the MHC (TAPs and LMPs). Separating these two regions is the class III region that contains genes generally not involved in antigen presentation such as

complement proteins, cytokines (TNFs and LTs), and highly conserved loci (HLA-B associated transcripts) not known to participate in immune processes (13).

*The Function of Histocompatibility Molecules*

Classical Class I MHC molecules, comprised of a Class Iα chain and a promiscuous β2 microglobin chain, are expressed on most nucleated cells to bind and present endogenously generated peptide antigens for $CD8^+$ T cell receptor recognition (16). Classical Class II MHC molecules, generally containing monogamous α and β chains, are typically only expressed on select antigen presenting cells (macrophages, dendritic cells, B cells) to bind and present exogenous peptide antigens for $CD4^+$ T cell recognition. Classical MHC molecules are well expressed and highly polymorphic, whereas non-classical MHC molecules (such as HLA-E,F,G and DM, DO) are generally monomorphic with alternative functions such as NK cell receptors and antigen processing (17,18).

The immune system must be able to recognize and adapt to a diverse and constantly evolving world of potential pathogens and determine what is “self” and what is “non-self”. B-cell receptors (and their secreted form, antibodies) recognized three dimensional features (epitopes) of antigens. T-cell receptors recognize linear peptides, derived from intracellular and extracellular degraded proteins, bound to MHC molecules. B and T-cell receptors are generated through somatic recombination or gene conversion mechanisms to rearrange variable gene segments and insert random nucleotides to create a unique and diverse repertoire of receptors (19). Whereas lymphocytes rely on somatic gene rearrangement to generate a diverse set of receptors, MHC molecules rely on

multiple loci, genetic polymorphisms and co-dominate expression to allow for a diverse set of peptides to be bound for T-cell presentation.

The mechanism of antigen processing and presentation has been recently reviewed by Vyas et al. (20). The three dimensional shape and amino acid residues within the peptide binding groove of MHC molecules determine the length and amino acid composition it may bind. Class I molecules are expressed by all nucleated cells and normally bind peptides of 8-10 amino acids endogenously generated through protein degradation by the proteosome. These peptides are transported, by TAP proteins, into the endoplasmic reticulum where they are loaded onto nascently synthesized Class I-$\beta_2$ microglobin complexes. Class I molecules present peptides to $CD8^+$ T-cells. In this manner, Class I molecules provide a form of self surveillance for intracellular pathogens or transformed cells. Class II molecules are normally only expressed on specific antigen presenting cells (dendritic cells, macrophages, and B-cells) and generally bind peptides from 13 to 20 amino acids in length. These peptides are generally derived from extracellular proteins through phagocytosis. The protein-antigen containing phagosome is fused with a lysosomal vesicle containing acid hydrolases and chaperoned Class II molecules. This phagolysosome degrades proteins into peptides suitable for Class II binding. The binding of peptides to MHC molecules stabilize the complex and allow for translocation to the cell membrane and subsequent presentation to T-cells. Class II molecules present antigens to $CD4^+$ T-cells which regulate and control the immune response.

*Major Histocompatibility Alleles Can Protect Against or Promote Disease*

In order to bind a diverse level of peptides, the peptide binding region of MHC molecules need to be highly variable. MHC genes are the most polymorphic loci in the vertebrate genome. The single human Class I locus (HLA-B) has over 800 reported alleles (21). The extreme polymorphism and multigenic nature of the MHC molecules is essential to the immune functions for both the individual and the species. The MHC must be able to present an extremely diverse set of both self and antigenic peptides to T-cells and MHC polymorphisms occur most often in the peptide binding region. Variations in the amino acid residues (polar, non-polar, charged, and size) in the peptide binding groove alter the types of peptides each allele can bind and present. The large number of pathogens encountered by the species results in a high level of positive selection at the MHC. This selection likely generates the extreme level of diversity found at the MHC, where diversity both in the species and population is literally a matter of life and death (2).

The ability of different Class I and Class II alleles to bind and present a diverse set of antigens can affect the way an animal responds to pathogens and progression of disease. In humans, alleles of the HLA-B locus can severely affect the progression, regression, and control of disease in patients infected with HIV (21). These effects are likely due to the ability of different Class I molecules (alleles) from infected cells to successfully present antigenic peptides to $CD8^+$ T cells, triggering a cytotoxic effect. Allelic variants of Class II genes have been identified as contributing to the susceptibility or resistance to several chronic inflammatory diseases including celiac disease, Type 1 diabetes, rheumatoid arthritis, and multiple sclerosis (22). The mechanism is unclear and perhaps resembles a molecular mimicry mechanism where certain alleles are capable of

presenting foreign peptides to CD4$^+$ T-cells, which in turn activate (help) B and/or T effector cells for specific self-molecules. Studies in dogs have associated MHC haplotypes to the incidence/severity of diabetes mellitus, primary immune-mediated hemolytic anemia, and hypothyroid disease (23-26). Equine recurrent uveitis has a strong association with a particular MHC haplotype (27).

*The Avian Major Histocompatibility Complex*

In contrast to mammals, avian MHC loci are considerably more condensed, containing fewer genes with shorter introns and untranslated regions (28). The chicken MHC locus (MHC-B) was the second vertebrate major histocompatibility complex characterized and contains many of the genes associated with mammalian MHC loci (29). A portion of the chicken B-F/B-L region spanning approximately 92 kb has been sequenced from the B12 haplotype. The locus can be further divided into two regions, the B-F/B-L region and the B-G region separated by about 150 kb (Figure 1). The B-F/B-L

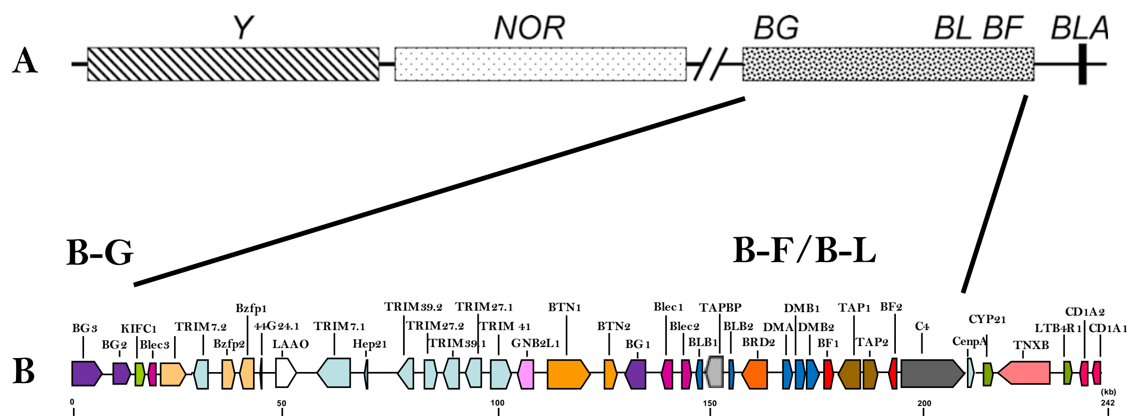



**Figure 1: The Chicken MHC Chromosome. A. Chromosomal content and hypothesized locus order of the MHC chromosome (adapted from Miller et al. (29)) B. Gene content and position of the MHC-B in chicken (adapted from Kaufman et al and Shiina et al. (AL023516, AB268588, (28,30)).**

region contains approximately 20 genes including Class Iα (B-F) and Class IIβ (B-L) genes. A single monomorphic Class IIα gene has been genetically mapped 5 cM from the B-F/B-L locus (31).

Additional genes present in the B-F/B-L region include two B30 (butyrophilin-like) genes, two lectin-like genes (one an NK cell receptor), a BG locus, TAPBP, BRD2, DMA, two DMB genes, TAP 1 and 2, complement protein C4, and a histone H3 gene. BG genes are highly polymorphic, polygenic, membrane-bound molecules with distinct immunoglobin-like regions whose functions are still undetermined (32). Further genomic analysis described additional loci upstream from the 92 kb sequence similar to TRIM genes in the B12 haplotype (33). CD1 loci which encode MHC-like molecules but are located on non-MHC chromosomes in mammals were found linked to the chicken MHC-B region (34,35). Recently published work has fully sequenced and assembled a 242 kb region of the Red Jungle Fowl (RJF, B21-like haplotype) encompassing the B-F/B-L region and extending significantly both upstream to the B-G region, downstream to the CD1 loci, and included zinc finger genes, TRIM loci, and a large array of tRNAs (30) . A further analysis resequenced and examined the level and etiology of diversity among 14 chicken haplotypes encompassing the B-F/B-L region (36).

In 2004 Shiina et al. (37) identified a portion of the quail (*Coturnix japonica*) MHC-B region and found considerable differences with the chicken. Although the two species belong to the same order (*Galliformes*) and family (*Phasianidae*), the chicken region (92 kb) is roughly half the size of the homologous region in the quail (~180 kb). While the overall gene families and gene orders are generally conserved, the total number of genes differs remarkably. All chicken haplotypes examined contain two Class Iα, two

Class IIβ genes, and a single BG gene. In contrast, the quail haplotype has seven Class Iα, ten Class IIβ, and eight BG genes. The number of NK receptor/Lectin-like genes also differs between the chicken and quail. Whether this dramatic variation is a distinction between the two species or a consequence of the particular haplotypes sequenced is yet to be determined. Further, it appears there may be some extensive flexibility with the quail species allowing for significant variation of not only the sequence of MHC alleles, but of the total number of loci as well (38). Sequence data from additional avian species are needed to explore such variations.

Prior to sequencing of the B-locus, Briles et al. (39) identified a second MHC-like region in the chicken. Originally denoted Restriction Fragment Pattern-Y (Rfp-Y, MHC-Y), this MHC locus is genetically independent of the MHC-B locus. Hybridization studies have mapped genes related to those found in the B-locus including Class Iα and Class IIβ-like loci to the Y locus (40,41). The Class I loci appear to be non-classical, polymorphic MHC Class I genes (42). A portion of the chicken Y locus has been sequenced and shown to contain at least one Class I-like gene (43). Several challenge studies have been performed to examine the effect of the chicken MHC-Y genotype and host response to pathogens. Wakenell et al. showed a significant, and independent, effect of both the MHC-B and Y genotypes to the response to Marek's disease virus (44). However, alternative studies using different virus and chicken strains showed limited effects of MHC-Y genotype on disease outcome (45,46).

Although unlinked genetically, the two loci (B and Y) have been placed on the same nucleolar organizer region (NOR)-containing microchromosome in the chicken (GGA16) through fluorescence in situ hybridization (FISH) and trisomic mapping

(47,48). In the chicken, the cytogenetic order of the MHC-B and MHC-Y loci was not precisely established. Bloom and Bacon (49) reported that the NOR occupies most of the distal end of a microchromosome. However, due to the tight linkage of the Y locus and the NOR, and presence of rDNA sequences within the Y locus cosmid cluster, it has also been inferred that the NOR may be located between the two MHC loci. It was hypothesized that the highly repetitive rDNA sequence block contributes to the enhanced recombination identified in this chromosome (47,48). However, recent efforts have definitively placed the two MHC loci adjacent and the NOR distal (M.E. Delany, unpublished results).

Similar to mammals, the two MHC Class Iα and two MHC Class IIβ genes in the MHC-B region are highly polymorphic in the chicken. Moreover, specific MHC-B haplotypes have been found to strongly affect the outcome of Marek's disease infection (e.g. tumor growth), the successful vaccination against Marek's, as well as survival studies against the minute Rous Sarcoma virus (50,51). The chicken B21 haplotype bestows a remarkable level of resistance to Marek's disease virus (52). Conversely, the B19 haplotype confers an equally noteworthy degree of susceptibility to the same pathogen (52). Other haplotypes show similar resistance/susceptibility effects for the Rous Sarcoma Virus (53). Whereas the B4 and B15 haplotypes confer a susceptibility to tumor progression, the B12 haplotype renders the bird resistant to tumor progression. In addition to these viruses, other studies have found an association of the MHC and response to bacterial colonization (54,55).

*The Turkey MHC*

Despite the marked progress in characterizing the chicken MHC, information regarding the turkey MHC has been limited. Early tissue grafting experiments identified immune reactions similar to those dependant on Class I and Class II molecules (56). However, a lack of serological reagents, genome sequence information, and highly inbred lines has limited turkey genetic researchers. Some of the first molecular experiments utilized RFLPs for MHC typing in turkeys (57,58). These studies took advantage of a chicken genomic DNA clone (Class IIβ gene) for use as a hybridization probe to genotype a random-bred subline segregating 4 separate and distinct MHC haplotypes. Southern hybridizations of *Pvu* II digested genomic DNA from nearly 400 individuals of four separate experimental lines identified at least seven shared haplotypes (59) and verified linkage disequilibrium of these genotypes relative to histocompatibility (58). Using similar techniques, an even larger study on 34-60 turkeys from each of eleven commercial breeder lines, as well as a sample of wild turkeys, identified the same haplotypes previously reported in the experimental lines (60). A single predominate haplotype (A) was present in the commercial lines and was also represented in the wild turkeys. The wild birds also had the highest proportion of unique haplotypes. Studies of an experimental Beltsville Small White line identified two additional haplotypes bringing the total number of commonly identified haplotypes to nine (61). Because RFLPs are dependent on relatively few nucleotide polymorphisms much of the genetic diversity in the MHC was likely missed.

Recent molecular biology studies of the turkey have used PCR to amplify and sequence portions of MHC genes. One study identified up to three MHC Class IIβ loci that shared strong homology to the chicken (62). Two additional studies identified

polymorphisms in the turkey TAP and TAPBP genes (63,64). Challenge studies in turkeys found subtle associations of RFLP (haplotypes) for four MHC haplotypes to *Pasteurella multocida* or Newcastle disease virus infection outcomes (65). However, again, a single RFLP may not be sufficient to define a single haplotype.

*Why Turkeys?*

Turkeys are the fourth largest agriculturally important livestock species, with annual production approaching eight billion pounds and valued at nearly $3.7 billion (USDA-NASS April 28, 2008) and in total, overall poultry production is valued near $30 billion (66). Although it is clear that genetic background is a major determinant in immune response to pathogens, this component has not been fully assessed in turkeys. For example, evidence exists that turkey lines selected for greater production qualities exhibit reduced immune response to pathogens as compared to random bred controls (67). Additionally, the emergence of the highly pathogenic avian influenza H5N1 virus and the systemic culling of regional flocks (140 million birds) have cost over $10 billion from 2003 to 2005 (World Bank, November 2005). In 1997, avian metapneumovirus (aMPV) was first detected in a flock of turkeys in Minnesota. Morbidity rates within infected flocks were between 50-100% with a high incidence of mortality (up to 30%; (68)). Studies primarily at the University of Minnesota attempted to understand the pathogenesis of MPV without considering the important role of the host genome in the response to infection (68-71). Because avian MHC haplotypes have been extensively associated with disease resistance/susceptibility, the ability to identify genetic

determinates resulting in disease resistance could reduce the costs associated with disease and prevent the spread of potentially zoonotic diseases.

Comparative studies between turkey and chicken have demonstrated high levels of sequence similarity, gene conservation, and extensive homology across the genomes (72-75) despite the estimated 20-60 million years separating their common ancestor (76,77). It is likely that they share significant similarity within the MHC: in structure, sequence, expression, and functional relationships to disease. However, since the MHC is also the most variable region of the vertebrate genome, an in-depth study of the turkey is required.

*Research Overview*

The research described herein provides a genomic understanding of the turkey MHC and develops reagents to help standardize host genetics in pathogen challenge studies and/or commercial selection for specific MHC genotypes. The objective of this project was to characterize the turkey major histocompatibility complex through sequencing, genetic and physical mapping, and analysis of variation within commercial and wild turkey populations. Additionally, the functionality of specific loci has been assessed by expression analysis. Chapter 2 describes the identification of genomic clones containing the MHC loci and their physical and genetic map positions. In chapter 3, the core MHC-B locus of the turkey is presented, with comparative and functional analysis. Chapter 4 provides an initial survey of the turkey MHC-Y and provides some evolutionary evidence of the origin and conservation of MHC throughout the vertebrate lineage. A resequencing survey and haplotype estimation of commercial turkeys is

described in Chapter 5 along with a similar survey of wild turkeys in Chapter 6. Finally, due to the unique level of MHC homozygosity identified in NT WF06 (the DNA source for the CHORI-260 turkey BAC library), a whole genome analysis of this bird and assessment of its appropriateness for whole genome sequencing is provided in Chapter 7. Concluding remarks and future directions for studying the turkey MHC and, by proxy, other poultry species are described in Chapter 8.

# Chapter 2

## Characterization of the Turkey MHC Chromosome through Genetic and Physical Mapping

# Characterization of the Turkey MHC Chromosome through Genetic and Physical Mapping[1]

**Chaves L.D., S.B. Krueth, and K.M Reed**

Department of Veterinary and Biomedical Sciences,

College of Veterinary Medicine,

University of Minnesota, St Paul, MN 55108

Published as:

Chaves L.D., S.B. Krueth, and K.M Reed. Characterization of the Turkey MHC Chromosome through Genetic and Physical Mapping. *Cytogenet Genome Res.* 117, 213-220 (2007).

[1] Manuscript published in Cytogenet Genome Res 117:213–220 (2007)

Previous studies in the chicken have identified a single microchromosome (GGA16) containing the ribosomal DNA (rDNA) and two genetically unlinked MHC regions, MHC-B and MHC-Y. Chicken DNA sequence from these loci was used to develop PCR primers for amplification of homologous fragments from the turkey (*Meleagris gallopavo*). PCR products were sequenced and overgo probes were designed to screen the CHORI 260 turkey BAC library. BAC clones corresponding to the turkey rDNA, MHC-B and MHC-Y were identified. BAC end and subclone sequencing confirmed identity and homology of the turkey BAC clones to the respective chicken loci. Based on subclone sequences, single nucleotide polymorphisms (SNPs) segregating within the UMN/NTBF mapping population were identified and genotyped. Analysis of SNP genotypes found the B and Y to be genetically unlinked in the turkey. Silver staining of metaphase chromosomes identified a single pair of microchromosomes with nucleolar organizer regions (NORs). Physical locations of the rDNA and MHC loci were determined by fluorescence *in situ* hybridization (FISH) of the BAC clones to metaphase chromosomes. FISH clearly positioned the rDNA distal to the Y locus on the q-arm of the MHC chromosome and the MHC-B on the p-arm. An internal telomere array on the MHC chromosome separates the B and Y loci.

## INTRODUCTION

The major histocompatibility complex (MHC) is considered the most polymorphic and gene dense region in the vertebrate genome. Immunologists first identified the region due to its dramatic effects on graft rejection and extensive work has been done to characterize the region in mice and humans. Many of the genes within this region are associated with antigen presentation to T cells including classical and non-classical MHC class I and class II genes. Classical class I MHC molecules, comprised of a class I α chain and a promiscuous $\beta_2$ microglobin chain, are expressed on most nucleated cells to bind and present endogenous peptide antigens for CD8+ T cell receptor recognition (16). Classical class II MHC molecules, generally containing monogamous α and β chains, are usually only expressed on select antigen presenting cells (macrophages, dendritic cells, B cells) which bind and present exogenous antigens for CD4+ T cell recognition. Classical MHC molecules are highly expressed and polymorphic, whereas non-classical MHC molecules are generally monomorphic with alternative functions such as NK cell receptors and antigen processing (17). The ability of different class I and class II alleles to bind antigens can affect the way an animal responds to pathogens and disease.

In addition to humans and mice, the MHC region of several species of economic and agricultural importance including cattle, pigs, horses, and chickens have been studied (11,78). The mammalian MHC appears to have some highly conserved features including overall gene content and organization, demonstrating the evolutionary significance of retention in this region. Mammalian MHCs are generally divided into three regions. The class I region contains mostly genes encoding classical and non-classical MHC class Iα

molecules. The class II region contains many genes for classical and non-classical MHC class IIα and β molecules. Separating these two regions is the class III region that contains genes generally not involved in antigen presentation such as complement proteins and TNF genes.

In contrast to mammalian genomes, avian MHC loci are considerably more condensed, containing fewer genes with shorter introns and untranslated regions (28,37). The chicken MHC-B locus was the second vertebrate major histocompatibility complex characterized and contains many of the genes associated with mammalian MHC loci (29). The locus can be further divided into two regions, the B-F/B-L region and the B-G region located ~200 kb away. A portion of the chicken B-F/B-L region spanning ~92 kb has been sequenced. This region contains approximately 20 genes including class I and class IIβ genes (28). The monomorphic class IIα gene has been genetically mapped to ~5 cM from the B locus (29).

Additional work by Briles et al. (39) identified a second MHC-like region, originally denoted Restriction Fragment Pattern-Y (Rfp-Y, MHC-Y), genetically independent of the B locus. Genes related to but distinct from those found in the B-locus including class I and class IIβ-like genes have been mapped to the Y locus (Miller et al., 1994). Notably present are what appear to be polymorphic non-classical MHC class I genes (42). A portion of the chicken Y locus has been sequenced and shown to contain at least one class I-like gene (43).

Although unlinked genetically, the two loci (B and Y) have been placed on the same nucleolar organizer region (NOR)-containing microchromosome in the chicken (GGA16) through fluorescence in situ hybridization (FISH) experiments (47) and

trisomic mapping (48). In the chicken the cytogenetic order of the MHC-B and MHC-Y loci is not precisely established. Bloom and Bacon (49) reported that the NOR occupies most of the distal end of a microchromosome. However, due to the tight linkage of the Y locus and the NOR and the presence of rDNA sequences within the Y locus cosmid cluster, it has also been suggested that the NOR may be located between the two MHC loci and thus contributes to the enhanced recombination identified in this chromosome (47,48). This study was undertaken to identify homologous MHC-B and MHC-Y regions in the turkey genome, determine their genetic linkage relationships, and their chromosomal locations.

## MATERIALS AND METHODS

*Genomic Amplification and Sequencing*

Genbank sequences were used to design primers for polymerase chain reaction (PCR) to amplify turkey genomic DNA. Primers to amplify 18s rDNA were designed based on the published turkey sequence[2]. Primers for the B and Y loci were designed from chicken sequences (Table 1). Primers for the B locus were designed to anneal within exons and amplify intervening introns of two genes, BTN2 and C4[3]. Y locus primers (Rfp-Y F 5`TATGCTGAGAGATGGAAGCA3`, Rfp-Y R 5`ACTATTACCCCAAAGCAACG3`) were obtained from P. Mariani (unpublished) or based on the partial Y region sequenced in chicken[4] (43). PCR reactions were performed using NT WF06 2002 E0010 genomic DNA, the same individual used to generate the CHORI 260 turkey BAC library.

---

[2] Genbank AJ419877
[3] Genbank AL023516
[4] Genbank AJ277927

PCR reactions (50 μL total volume) included approximately 50 ng genomic DNA, 1.5 mM $MgCl_2$, 25 pmol each primer, 100 μM dNTPs, and 1 U Hotstar Taq polymerase (Qiagen Inc, Valencia CA.). Amplifications were performed with the following reaction conditions: 15 min at 94° C; 40 cycles of 30 s at 94° C, 30 s at annealing temperature, 1 min at 72° C; and a final extension of 5 min at 72° C. Products were electrophoresed through 1% agarose, purified using a Qiaquick gel extraction kit, and sequenced on an automated DNA sequencer (ABI). Sequences were aligned to the chicken sequence to confirm identity.

*BAC Library Screening and Subcloning*

Sequences obtained from turkey were used to develop overgo probes for the B and Y loci (Table 2). The CHORI 260 BAC library array was screened by overgo hybridized as described by Ross et al. (79) with pooling of additional overgo probes (unpublished results). Positive BAC clones were identified and grown overnight in LB media containing 25 μg/mL chloramphenicol. BAC DNA was prepared using QIAprep columns (Qiagen, Inc.) and presence of the original target sequence was confirmed by PCR using the 18s rDNA, BTN2, C4, and Rfp-Y locus primers (Table 1).

Confirmed clones were prepared using a Qiagen large construct kit (Qiagen, Inc.) and end sequenced with vector primers (T7- TAATACGACTCACTATAGGG, gSP6-GTTTTTTGCGATCTGCCGTTTC). To generate internal subclones from BACs containing the MHC loci, purified DNA from clones 97E5 and 159B18 was subjected to digestion with restriction enzymes (*Bam* HI, *Eco* RI, and *Hind* III) according to manufacturers recommendations (New England Biolabs, MA). The digested DNA was

electrophoresed through 1% agarose and 1-3kb fragments were gel extracted and purified with Qiaquick columns (Qiagen, Inc). The resulting fragments were ligated into like-digested pBluescript II sk+ (Stratagene, Inc) pretreated with calf intestinal phosphatase (New England Biolabs). Ligated DNA was transformed into electrocompetent XLI Blue MRF` cells (Stratagene, Inc) and plated on LB agar containing ampicillin, Xgal, and IPTG. Individual clones were picked and grown overnight in LB containing ampicillin (100 μg/mL). Plasmids were purified with a QIAprep 96 turbo kit (Qiagen, Inc), and sequenced on an ABI automated sequencer.

*Genetic Linkage Analysis*

In order to identify putative single nucleotide polymorphisms (SNPs) for linkage mapping, MHC fragments amplified by PCR (Table 1) and additional Y locus fragments (identified by subcloning) were amplified and sequenced in the UMN/NTBF resource population ((80), Table 1). SNPs for three loci were found to occur within the recognition site of restriction enzymes indicating they could be genotyped by PCR-RFLP. Genetically informative markers were genotyped in the mapping families by the direct addition of the appropriate restriction enzyme to the PCR products (Table 1). The resulting restriction fragments were electrophoresed on 1-3% agarose gels and alleles were manually scored. Linkage analysis was performed using Locusmap software (81).

*Cytogenetic Analysis*

Turkey embryonic fibroblast cell lines and metaphase chromosome preparations were generated as described in Reed et al. (82). To identify the NOR in turkeys, silver

staining was performed according to Howell and Black (83). For physical mapping, BAC clones and a subclone derived from BAC 159B18 containing a turkey-specific repetitive element (MgaSat2[5]) were labeled for FISH. Briefly, BACs and plasmids were subjected to nick-translation with aminoallyl dUTP and labeled with Alexa Fluor amine-reactive dyes (Invitrogen, Inc) according to the manufacturer's protocol. Metaphase spreads were denatured and dehydrated using standard protocols. For each hybridization, approximately 200 ng of each probe, 500 ng turkey competitor DNA and 600 ng sheared salmon sperm DNA was used in 12 μl of hybridization solution (50% formamide, 10% dextran, 1X SSC). Probes were denatured at 75° C for 5 minutes and allowed to anneal at 37° C for 3-5 minutes. The mixture was added to prewarmed slides and allowed to hybridize overnight in a moist chamber at 37° C. Slides were washed and stained with ProLong Gold antifade reagent with DAPI (Molecular Probes, P36931, now Invitrogen, Inc). To identify chromosome ends and interstitial telomeres, an Alexa Fluor 488 tagged oligonucleotide (CCCTAACCCTAACCCTAACCCTAAC) corresponding to the vertebrate telomere repeat was used. Hybridizations containing the telomere probe were performed without blocking DNA and post-hybridization stringency washes were done at room temperature. Photo microscopic images were captured with Metamorph V6.2.1 software using a SPOT Insight camera mounted on a Nikon E800 microscope.

## RESULTS

*Genomic Amplification and Sequencing*

In order to confirm the presence of homologous chicken MHC regions in the turkey genome and to generate turkey specific sequence, DNA from the MHC B and Y

[5] Genbank AJ238569

regions was amplified and products of the appropriate size were purified and sequenced. Sequences were compared to Genbank and Ensembl databases by Blastn to verify correct amplification. All sequenced products were highly homologous to the published chicken B[6] and Y[7] sequences with E values between 4.00E-41 and 0.0.

*BAC Library Screening and Subcloning*

The CHORI 260 turkey BAC library was screened to obtain clones containing B and Y regions for use in downstream applications, e.g. physical mapping and sequencing. Turkey-specific DNA sequence was used to develop a set of overgo probes to identify BAC clones (Table 2). Pooled overgo hybridizations yielded many positives that were then sorted by PCR and/or BAC end sequencing. PCR and BAC end sequencing assigned two BACs (01H21, 97E05) to the B locus and one BAC (159B18) to the Y locus.

The two B locus clones were approximately the same length by gel electrophoresis and terminated at the same *Eco* RI site in the 5`UTR of the TAP1 gene[8] . The other end sequences[9] appeared ~500 bp apart within a TRIM7 gene based on alignment with the chicken genome. The BAC clone 97E05 appeared to be the larger of the two and was consequently subjected to further analysis. Assuming a high degree of homology between turkey and chicken, this BAC clone appears to encompass ~75% of the described chicken B-locus, potentially only lacking the genes for TAP2, a class Iα gene, and the complement protein C4.

The clone 159B18 was confirmed to contain MHC-Y associated sequences by PCR amplification and sequencing. However, end sequencing of this clone yielded a

---

[6] Genbank AL023516
[7] Genbank AJ277927
[8] Genbank DX922435
[9] Genbank DX922431 and DX922434

short sequence on one end[10] that was not homologous to any chicken sequence (2004 assembly) while the other end[11] was highly similar (E value 0.0) to the MgaSat2 satellite repetitive element (Genbank AJ238569). Many BAC clones were positive for 18s rDNA, however only a single clone was selected for further use. End sequencing of the clone 170K02 in both directions and subsequent Blastn analysis identified it as containing *M. gallopavo* 18s rDNA (E=0.0).

Several subclones from 97E05 contained sequences with significant homology to genes located in the published chicken B-locus. Some subclones from 159B18 contained sequences similar to MHC class I molecules and the partial Y sequence from the chicken (Table 3). Subclone sequencing and subsequent Blast analysis revealed additional clones with homology to other genes, regions, and/or with no significant homology (unpublished results). These results indicate these two BAC clones from the turkey (97E05 and 159B18) represent homologous regions to the two chicken MHC loci. Further sequencing of these clones is currently underway. Of interest is the degree of synteny between the turkey and chicken within the MHC. The quail, with a similar evolutionary divergence with respect to chicken as the turkey, has a high level of variation in this genome region, with the quail MHC-B locus nearly twice the size of the chicken (37).

*Genetic Linkage Analysis*

Despite their location on the same microchromosome, the B and Y loci are genetically unlinked in the chicken (29). In order to ascertain if a similar genetic relationship exists in the turkey, DNA from the F1 individuals of the UMN/NTBF

[10] Genbank DX922430
[11] Genbank DX922429

resource population was amplified by PCR and sequenced with MHC-B and Y primer sets. Additional loci generated from the BAC corresponding to these regions were examined to increase statistical significance and ensure alleles were in linkage phase (Table 1). Informative SNPs were identified in the introns of the BTN2 and C4 genes (B locus) and in an anonymous region of the Y locus.

Segregation analysis of the MHC-B and Y SNPs established a linkage group containing the BTN2 and C4 markers (MHC-B) but excluded the MHC-Y marker (Rfp-Y BAM F3). The genetic distance between BTN2 and C4 was 0.6 cM (LOD = 36.32) with a single informative recombinant within the mapping population. As observed in the chicken, the turkey MHC-B and Y loci were genetically unlinked as marker alleles were in phase and contained over 100 informative individuals for each locus (Table 1).

*Cytogenetic Analysis*

In the chicken the NOR resides on the long arm of a microchromosome (GGA16). Silver staining of turkey fibroblasts identified two acrocentric microchromosomes with high levels of active transcription associated with the NOR (Fig 1). These chromosomes were highly stained on the majority of the q-arm, with a small unstained region near the centromere, localizing the NOR on the distal end of the q-arm.

BAC clones identified as homologous to chicken MHC-B, MHC-Y, and rDNA were labeled by nick-translation incorporating Alexa Fluor dyes (Invitrogen, Inc). Two color FISH placed both MHC BACs on the same acrocentric microchromosome (Fig 2A). The B locus BAC (97E05) hybridized to the p-arm, displaying unique spots representing each chromosome arm (Fig 2A). A single broad hybridization site was

consistently observed for the Y locus BAC (159B18) indicating a more centromeric position due to the constriction of the chromosome arms created by the centromere (Fig 2A). Interphase cells hybridized with B and Y BACs displayed a minimal distance between the MHC loci further suggesting the rDNA repetitive sequences (~5-7 Mb in chickens, (84)) are located telomeric of Y locus.

In order to confirm the co-localization of the MHC loci and rDNA, two color FISH was performed with the rDNA-containing BAC (170K02) and each MHC BAC. The rDNA BAC hybridized to most of the q-arm. Hybridization results confirmed that the rDNA and both MHC loci are indeed on the same microchromosome in turkeys as in chickens (Fig 2B and C). Simultaneous hybridization of the B, Y, and rDNA BACs display the linear order of all three loci on the same chromosome (Fig 2D). In the turkey the NOR (rDNA) is on the distal end of the q-arm, the Y region is centromeric, and the B region is on the distal end of the p-arm.

Both Miller et al. (48) and Fillon et al. (47) suggested the high rate of recombination observed in the chicken between the B and Y loci was caused by the intervening repetitive rDNA. Since in turkeys it is apparent the NOR is located distal to the B and Y loci it was of interest to identify other chromosomal elements that may contribute to the apparent high rate of recombination. Delany et al. (85) identified an ultra-long telomere array on chromosome 16 in the chicken and suggested that a combination of repetitive elements contribute to the high rate of recombination found in certain chicken chromosomes. End sequencing of the MHC-Y BAC revealed the presence of a repetitive turkey satellite sequence (MgaSat2).

A BAC subclone containing the MgaSat2 repeat and a telomere-specific

oligonucleotide were hybridized to turkey chromosomes. The MgaSat2 repeat showed a strong hybridization signal to the MHC chromosome (Fig 3A-B) indicating a high copy number of this particular repeat. Lighter signals were detected near the centromeres of several other chromosomes, however it cannot be determined if this is a centromere-specific repeat. The telomere probe revealed many sites throughout the turkey genome (Fig 3A and C) with strong signals indicative of ultra-long telomere arrays. Dual color FISH found the co-localization of the MgaSat2 repetitive element (Fig 3) with a large telomere array on the MHC chromosome. Two color FISH with the telomere probe and the rDNA BAC suggests that a large telomere occurs on the short arm of the MHC chromosome (Fig 3D). Two color FISH using the telomere and MHC-B probes clearly place the large telomere array internal (centromeric) to the MHC-B locus (Fig 4).

**DISCUSSION**

Although separated by approximately 40 million years of (76), numerous studies have described the highly homologous and syntenic relationship between the turkey and chicken genomes (72,80,86,87). Chicken DNA sequences were successfully utilized in the design of primers for amplification of turkey MHC regions indicating further homology between the species, even in the most polymorphic and positively selected region of the genome.

Screening the CHORI 260 BAC library resulted in three MHC containing clones, far less than expected in a library of 11X coverage. One reason may be due to the screening process. Pooling of overgo probes is an efficient method for large scale screening (76), however, the necessity to verify clones by PCR can add additional

limitations. The highly polymorphic nature of the MHC can limit the number of positives detected by hybridization through null hybridization-site producing substitutions. The overgo probes and PCR primers for the turkey MHC regions were designed from monomorphic PCR sequence, potentially limiting the hybridizations to one haplotype.

There may be other reasons why this region of the genome is not amenable to large insert cloning. For instance, these loci may encode gene products toxic to the host bacteria. Relative to other BAC clones, clones 97E05 and 01H21 indeed appeared to grow at ~50% the rate of other clones. Modern high-throughput clone-picking machines used to harvest BAC libraries may be less likely to pick slow growing (small) clones resulting in lower representation in the resulting libraries.

Early mapping of the chicken MHC was based on trisomic mapping. Bloom and Bacon (49) identified through serology and histology the co-inheritence of three MHC haplotypes and three NOR per cell. Likewise, through similar trisomic mapping and Southern blot procedures, Miller et al. (48) noted that the NOR (and MHC-B locus, by proxy) and the MHC-Y locus were also located on the same chicken microchromosome with a high level of recombination between the loci. Additionally, Fillon et al. (47) physically mapped cosmids representing the chicken B, Y and rDNA through FISH to the same microchromosome. In the latter two studies, the authors suggest that the NOR was likely located between the two MHC loci and that this large repetitive sequence block may contribute to the higher level of recombination found in this region.

Linkage analysis found the two MHC loci in turkeys to be genetically unlinked as they are in the chicken. The high rate of recombination between the MHC-B and Y loci is counterintuitive given their proximity to a centromere, which are classically regions of

reduced recombination. Delany et al. (85) suggested that repetitive elements such as the ultra-long telomere arrays of chromosomes 9, 16, 28 and W, contribute to the high rate of recombination found in certain chicken chromosomes similar to that identified in hamsters by Ashley and Ward (88). Perhaps the MgaSat2 sequence, the interstitial ultra-long telomere array, or both may be contributing to the recombinational hotspot found on the turkey MHC chromosome. Interestingly, the MgaSat2 sequence is not found in the chicken genome. Possibly an alternative repeat may be present near the chicken Rfp-Y.

Chromosomal hybridizations confirmed the location of the two homologous turkey MHC loci on the same microchromosome. Hybridizations clearly position the rDNA distal on the q-arm of the MHC chromosome and the MHC-B on the p-arm (Fig 2B). The MHC-Y locus can be tentatively placed on the centromeric end of the q-arm based on hybridization results, the close linkage of the chicken MHC-Y locus and the NOR, as well as the presence of rDNA genes within the same cosmid cluster as Y genes reported in Miller et al. (48). The minimal distance observed between the MHC loci at both metaphase and interphase further indicate that the rDNA repetitive sequences are located telomeric of Y locus and not between the two MHC loci. A difference in chromosome structure may exist between the two bird species due to a chromosomal rearrangement between. However further analysis of the MHC chromosome (GGA16) in the chicken would be of interest to determine exact physical order of the MHC loci.

**Table 1**: MHC loci amplified in the turkey. For each locus, the PCR primers, amplicon size, annealing temperature (TM) and Blast results are given. Restriction enzyme used for PCR-RFLP genotyping and number of informative meioses are given for polymorphic loci mapped within the UMN/NTBF population.

| Locus | Primer Sequence | Size (bp) | Tm (C) | Genbank Accession | Blast Result Accession | E-value | SNP / Enzyme | Informative Meioses |
|---|---|---|---|---|---|---|---|---|
| BTN-F | GGCTGAGAGGCAGAAAATTG | 1594 | 58 | DQ836346 | AL023516 | 0.0 | *Tsp* 509I | 172 |
| BTN-R | ACCTCCACATCCCAAAAGTG | | | | | | | |
| C4-F | TGAGATAGGGGCTTTTGTGG | 417 | 58 | DQ836347 | AL023516 | 2.00E-83 | *Mbo* I | 44 |
| C4-R | CAGCTCTTCCGCACCAGCAG | | | | | | | |
| C4b F | GCTGTGGGTCACATGCAG | 977 | 58 | DQ836348 | AL023516 | 6.00E-120 | *Mbo* I | 252 |
| C4b R | TCTCACCTTCCAGTCCTTCC | | | | | | | |
| Rfp-Y BAM F3 F | GGTGGGAGCTGAGGAGTACC | 653 | 58 | DQ836351 | BX929356.1 | 4.00E-41 | *Bsr* DI | 120 |
| Rfp-Y BAM F3 R | GTGACATCGCAGTCCTTGAG | | | | | | | |
| YFV3 F | CACTCCCTGCGCTACTTC | 266 | 58 | DQ836353 | AJ277927 | 1.00E-53 | N/A | N/A |
| YFV3 R | CCACGCTCACCTTTACTTTT | | | | | | | |
| MgRfp-Y F | CGTGCTGAAAAGGAGAGGTG | 682 | 58 | DQ836349 | AJ277927 | 1.00E-111 | N/A | N/A |
| MgRfp-Y R | CCTTCCCTGCAGACAGAAAG | | | | | | | |
| Rfp-Y HindIII 2 F | GGTAACATCCCATTGGCTTC | 409 | 58 | DQ836352 | AC175394 | 6.00E-117 | N/A | N/A |
| Rfp-Y HindIII 2 R | CTGGTAGACGGAGGAAAAGC | | | | | | | |
| Rfp-Y BAM A1 F | AGGATGGAGAACACCATTGC | 933 | 58 | DQ836350 | XM_415344 | 2.00E-70 | N/A | N/A |
| Rfp-Y BAM A1 R | GTTCACGATTGCAACAATGC | | | | | | | |
| Mg18s F | TCCCGTTCCTTGGATAACTG | 530 | 58 | MGA419877 | NA | NA | N/A | N/A |
| Mg18s R | CAGCTCGATCCCAAGATCC | | | | | | | |

**Table 2:** Overlapping oligonucleotides used to generate overgo probes used in CHORI 260 library array hybridization.

| Overgo | Sequence | MHC Locus | BACs Identified |
|---|---|---|---|
| BTN-2O1 | GCTCCGCCATTTTCTGAAGGAGA | MHC-B | 96E05, 01H21 |
| BTN-2O2 | GAAGCATCACCTCCTTCTCCTTC | | |
| C4O1 | TTGAGACGGAAATCAAAGAGGTG | MHC-B | NONE |
| C4O2 | TCACCTCCAAGCAGCACCTCTTTG | | |
| MHC-YO1F | GACAGTAAGAGCTGGAGGGCGCAG | MHC-Y | 159B18 |
| MHC-YO1R | CAGTATCTCCACAATGGGCTGCGCC | | |
| MHC-YO2F | CAGAGGGCCCAGGATGGTGAGGTGG | MHC-Y | 159B18 |
| MHC-YO2R | CCCAAGAACCAGTAGAGATCCACC | | |
| Mg 18s OF | TCCCGTTCCTTGGATAACTG | rDNA | 170K02 |
| Mg 18s OR | ATTAGCTCTAGAATTACCACAGTT | | |

**Table 3:** Blast analysis of representative subclone sequences from MHC BAC clones.

| Subclone | Genbank Accession | Blast Hit | Accession | E-value |
|---|---|---|---|---|
| 97E05 Bam 6 | DX922436 | *Gallus gallus* B Locus | AL023516 | 0.0 |
| 97E05 Bam 7 | DX922437 | *Gallus gallus* B Locus | AL023516 | 0.0 |
| 159B18 Bam 8 | DX922432 | *Gallus gallus* partial Rfp-Y region | AJ277927 | 1.00E-57 |
| 159B18 Hind 10 | DX922433 | *Gallus gallus* similar to MHC Rfp-Y class Iα chain | XM_415344 | 4.00E-70 |

**Figure 1:** Silver-stained turkey metaphase chromosomes. Arrows denote p-arms of an acrocentric microchomosomal pair with distinct NORs. Inset depicts the overlapping location of the two chromosomes.

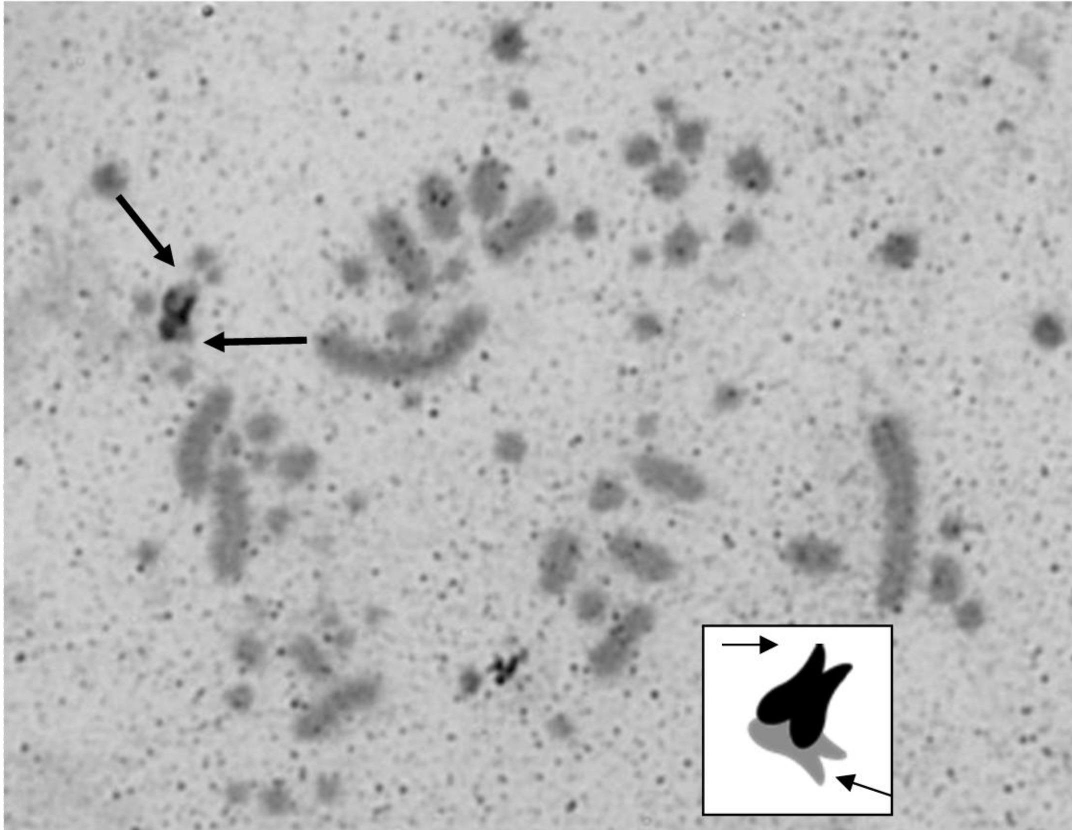


**Figure 2:** Dual color FISH of BAC clones to turkey metaphase chromosomes. Arrows indicate selected chromosomal p arms. **(A)** Hybridization of BAC clones 97E05 (MHC-B, green) and 159B18 (MHC-Y, red) to the same microchromosome. Insets indicate relative chromosomal distance between the two loci at metaphase **(a)** and in interphase **(b)**. **(B)** Hybridization of BAC clones 97E05 (MHC-B, green) and 170K02 (rDNA, red) to the same microchromosome. Inset shows enlarged view of the MHC chromosomal pair from a second individual. **(C)** Hybridization of BAC clones 159B18 (MHC-Y, green) and 170K02 (rDNA, red) to the same microchromosome. Inset shows enlarged view of the MHC chromosomal pair from a second individual. **(D)** Hybridization of BAC clones 97E05 (MHC-B, red), 159B18 (MHC-Y, green) and 170K02 (rDNA, red) to the same microchromosome. Inset shows enlarged view of the MHC chromosomes.

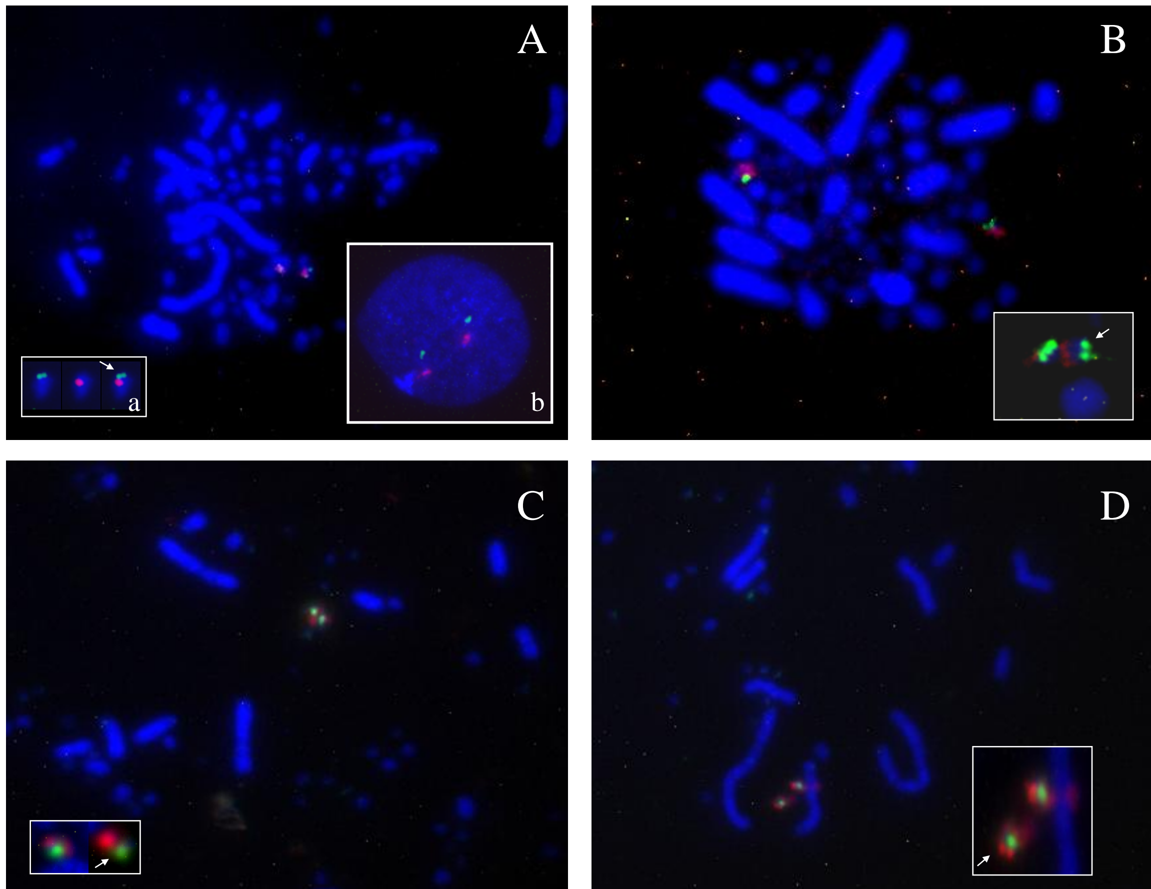


**Figure 3:** **(A-C)** FISH of MgaSat2 repetitive DNA (MHC-Y subclone) and telomere repeat to turkey metaphase chromosomes. **(A)** Dual color FISH of polyploid metaphase displaying hybridization of MgaSat2 repeat (red) and telomere repeat (green). **(B)** Same metaphase as in A with red signal enhanced. Arrows denote position of minor MgaSat2 sites on several microchromosomes. **(C)** Same metaphase in A with green signal enhanced. Arrows denote internal telomeres associated with the MHC chromosomes. **(D)** Dual color hybridization of BAC clone 170K02 (rDNA, red) and telomere repeat (green). Inset shows relative position of the internal telomere relative to the rDNA on the MHC chromosome.

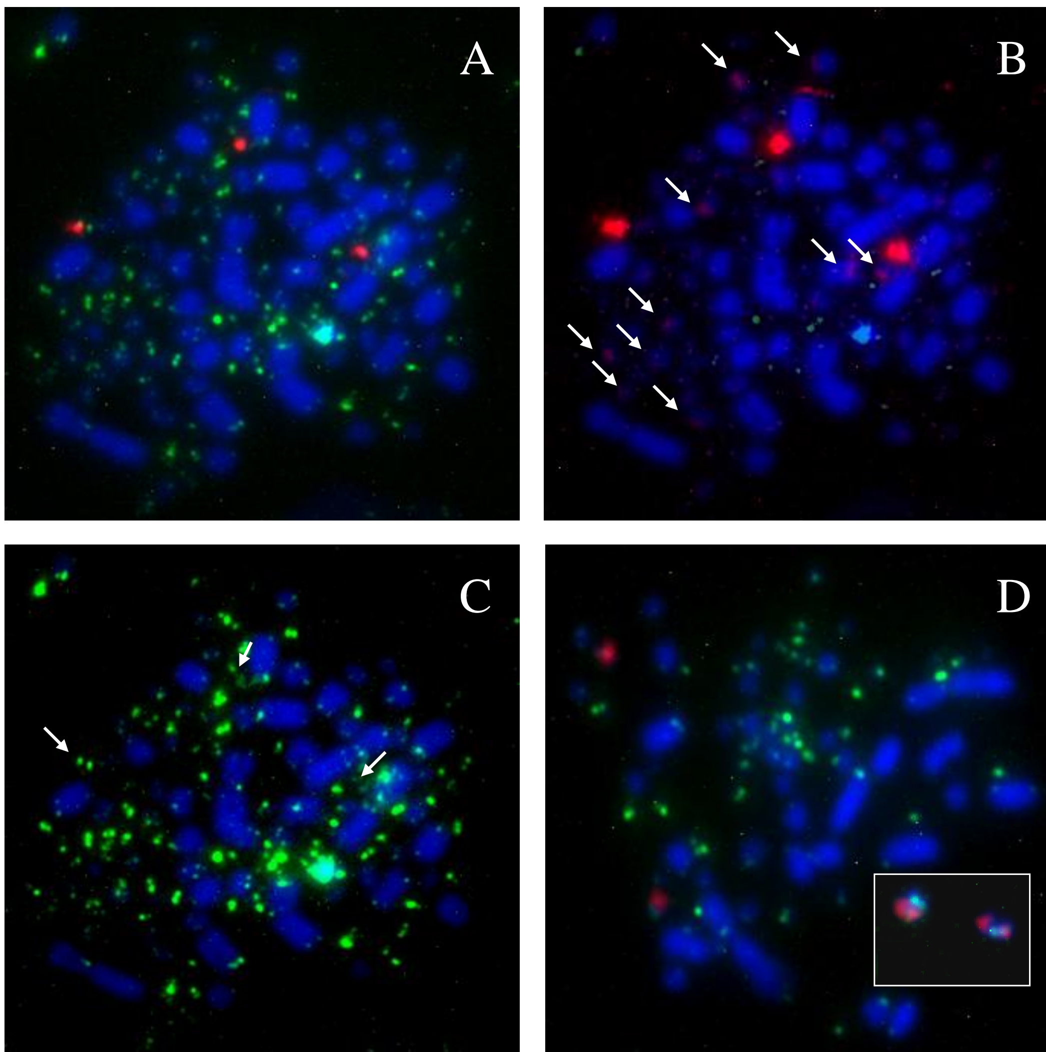


**Figure 4:** Hybridization of BAC clone 97E05 (MHC-B, red) and internal telomere repeat (green) to the same arm of the microchromosome. Inset shows enlarged view of the MHC chromosomal pair from a second individual.

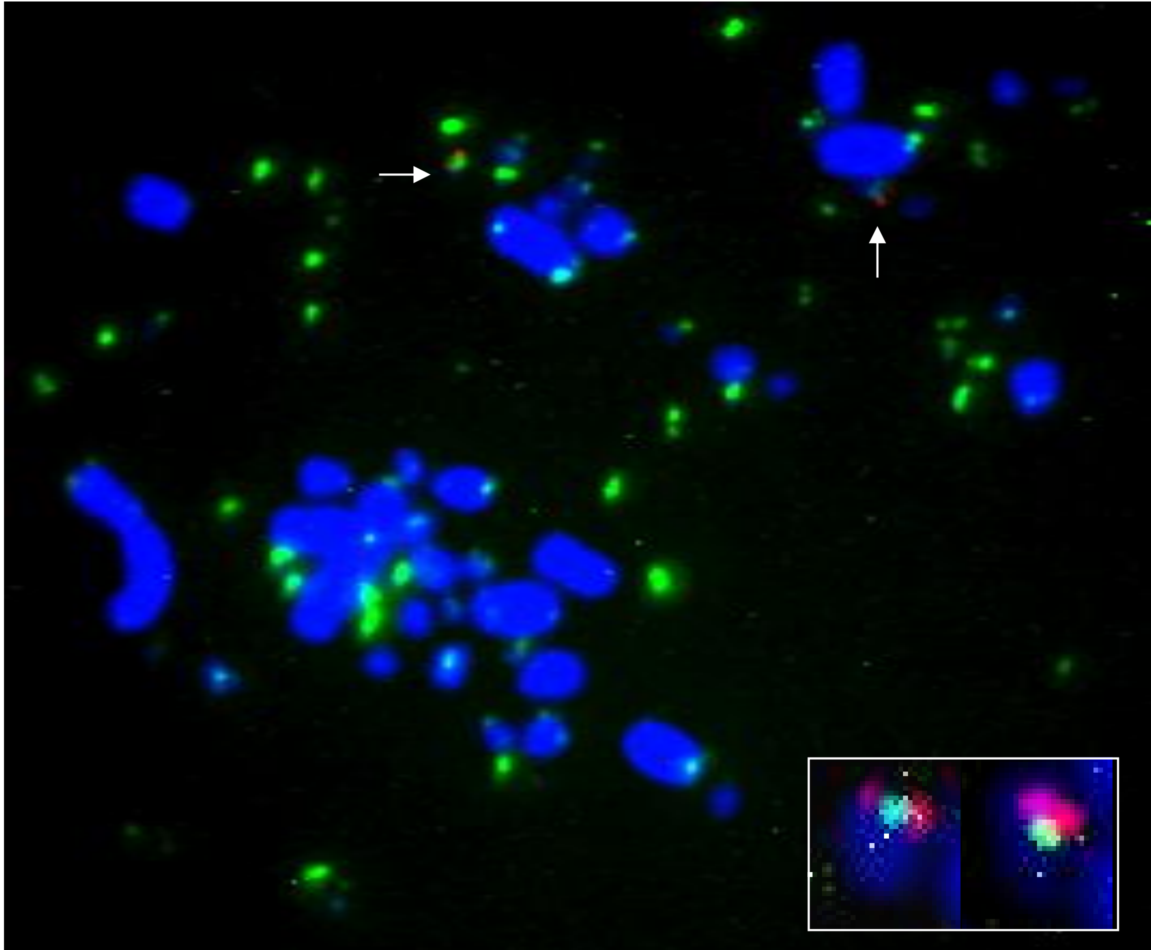
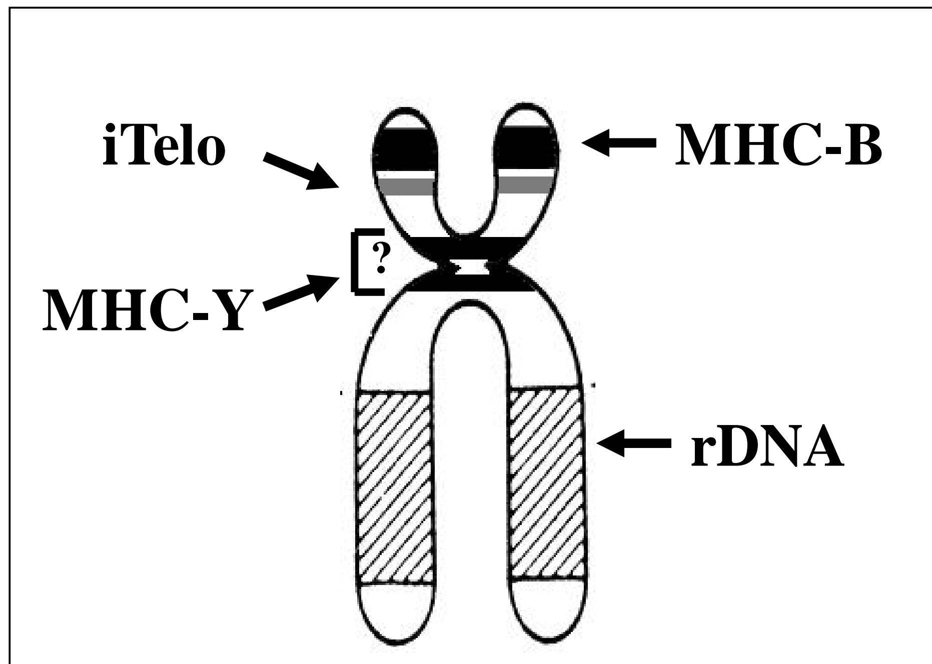


**Figure 5:** Current model of the turkey MHC chromosome based on genetic and physical mapping. MHC-Y is positioned on both sides of the centromere to denote the ambiguity of its precise location. (Adapted from Bacon and Bloom, 1985)



# Chapter 3

# Defining the Turkey MHC: Sequence and Genes of the B-Locus

# Defining the Turkey MHC: Sequence and Genes of the B-Locus

**Chaves L.D., S.B. Krueth, and K.M Reed**

Department of Veterinary and Biomedical Sciences,

College of Veterinary Medicine,

University of Minnesota, St Paul, MN 55108

## INTRODUCTION

The major histocompatibility complex is a genomic locus found in all jawed vertebrates and a key component in immune responses. Originally identified through tissue graft rejection experiments, the MHC locus has subsequently been found to contain several classes of genes responsible for antigen presentation to the host immune system. Specifically, classical MHC molecules encoded within the MHC possess a highly polymorphic peptide binding groove to bind peptide antigens through hydrophobic and/or hydrogen bonding and present them to T-cell receptors. The MHC Class Iα genes are expressed in all nucleated cells, interacting with β2-microglobin to present mostly endogenously generated peptides of 8 to 12 amino acids. MHC Class II molecules are heterodimers (α and β genes) primarily expressed on antigen presenting cells (dendritic cells, B-cells, and macrophages) and present mostly exogenously derived peptides of approximately 10-34 amino acids.

The chicken MHC has been defined as two genetically unlinked clusters, the MHC-B and -Y loci, located with the nucleolar organizer region (NOR) on the same microchromosome (GGA16) (39,47-49). The Y-locus contains lectin-like and non-classical MHC genes with varied effects on disease susceptibility (40,44,89,90). At least one Class I-like locus is polymorphic and transcribed (42). The chicken MHC-B is subdivided into two regions, the BG and the BL-BF. The BL-BF region contains the classical Class I and Class IIβ genes. In contrast to mammalian genomes which contain on average 6 paralogous genes for Class Iα, Class IIα, and Class IIβ genes (11), the chicken BL-BF locus has been described as a “minimal essential MHC” containing only two MHC Class Iα and two Class IIβ genes (28,29). A single monomorphic Class IIα gene located 5 centimorgans from the BF-BL region encodes a protein that is shared

between the Class IIβ to form the mature antigen (31). Furthermore, some chicken MHC haplotypes appear to predominantly express a single Class Iα and Class IIβ transcript thereby reducing the diversity of antigens presented (91,92). Interestingly, unlike mammals, the chicken has two C-type lectin-like genes, one of which is quite similar to NK complex loci (93). This minimal essential MHC is thus far exclusive to the chicken, as reports in other avian species have shown higher numbers of classical MHC loci. The limited repertoire of MHC molecules in the chicken—and the antigens they are able to present—has a remarkable effect on the species' ability to resist/resolve infectious disease including bacteria, viruses, and parasites (51,94,95).

Recent work in the turkey has identified two MHC regions homologous to the chicken B and Y loci (96). BAC clones containing portions of these regions were physically mapped to turkey metaphase chromosomes through fluorescent *in situ* hybridization and genetically mapped by segregation analysis using a resource population. Like the chicken, these two regions were genetically unlinked, located on the same NOR-containing microchromosome (96).

Separated by an estimated 50 million years (97), the genomes of the turkey and chicken have been shown to be highly homologous; chromosomal markers are present in both species in syntenic order (75). Gene sequence studies have found most coding and predicted amino acid sequences to be more than 90% identical (73,98). However, little is known of the similarities between the two species at the most variable region of the vertebrate genome, the MHC.  This work was undertaken to describe the core MHC sequence of the turkey and to compare this most variable genome region to homologous sequences derived from other avian species. The resources available for the chicken

(whole genome sequence, multiple MHC sequences, and close phylogenetic relationship to turkeys) provide excellent tools for comparative analysis. Results of this study provide genomic resources for the study of the effect of the turkey MHC in disease susceptibility and resistance and present insights into the evolutionary origins of the unique structure of the avian MHC.

## MATERIALS AND METHODS

### *Sequencing Strategy*

The CHORI-260 BAC library was generated with DNA from a female (NT WF06) of a partially inbred Nicholas commercial sub-line (Chaves et al. in press). A clone from the library (97E05) containing a portion of the MHC-B region was previously isolated (96). Screening for additional BAC clones was performed as previously described using overgo probes based on the end sequence of clone 97E05, as well as a PCR product corresponding to the CD1.1 gene[12] of the turkey (75). Additional BAC clones were identified. End sequencing of these clones[13] anchors them within the 97E05 clone with additional sequence extending further into the 5` BG region. No clones corresponding to the 3` MHC-B region were identified.

The BAC clone 97E05 was purified using the Qiagen Large Construct kit and shotgun subcloned. Plasmid DNA was isolated and sequenced by automated Sanger sequencing at the University of Washington High Throughput Genomics Unit (Seattle). Sequences were manually edited, aligned, and assembled using Sequencher software (Gene Codes, Corp.). Primer walking on selected subclones was performed to fill gaps.

[12] Genbank EU522671
[13] Genbank ET222701-4

Chicken genome sequence[14] from the MHC-B region (30) was used to develop primers for the amplification and sequencing of the 3' portion of the B-locus (sequence not included in the turkey BAC clone insert) TAP1-CenpA. The nascent turkey sequences were aligned with the chicken sequence for further primer design and sequencing (Table 1). PCR reactions were performed with genomic DNA (NT WF06) template using Taq Mastermix (Promega) supplemented with 1X Q solution (Qiagen Inc). Amplifications were performed for 30 cycles with 58° anneal temperature and 1 min/kb extension times. Heterogeneous PCR products (e.g. MHC Class Iα) were cloned using the QIAGEN PCR Cloning Kit (Qiagen Inc), transformed into DH5α cells (Invitrogen), and purified plasmid clones were sequenced using vector-specific primers.

*Gene Identification and Annotation*

Sequences were analyzed with the basic local alignment search tool (BLAST) and Softberry FGENESH (http://linux1.softberry.com/all.htm) to identify putative transcripts and homologies to known genes. Signal elements were identified with SignalP 3.0 (99). Comparisons between predicted gene sequences, available expressed sequence tags (ESTs) from poultry species, and the published chicken MHC sequence (28,30) were performed using Sequencher software (Gene Codes, Corp.). Repetitive elements were identified using REPEATMASKER and Tandem Repeats Finder (100) (http://tandem.bu.edu/trf/trf.basic.submit.html) and tRNAs elements were identified using tRNAScan (101). CpG islands were elicited with Softberry CpGfinder (http://linux1.softberry.com/all.htm) and GC content analysis was performed with 100 bp windows using Isochore (http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html).

[14] Genbank AB268588

*Interspecies comparisons*

Identity dot matrix plots and phylogenetic trees were drawn using PipMaker (http://bio.cse.psu.edu/pipmaker) and ClustalX2.0.3 software, respectively. Sequences of duck Class Iα[15] and Class IIβ[16] and whooping crane BG[17] were utilized as outgroups for phylogenetic comparisons. Synonymous and non-synonymous substitutions were identified based on the methods of Nie and Gojobori (102) using SNAP software (103) (http://www.hiv.lanl.gov/content/sequence/SNAP/) .

*RT-PCR and Cloning*

Unfortunately RNA or preserved tissues were unavailable from the turkey (NT WF06) used to generate the CHORI 260 BAC library. Therefore, an outbred commercial female turkey (OCB02) was used for gene expression studies and to verify selected gene annotations. Total splenic RNA was isolated by the Trizol method according to the manufacturer's recommendation (Invitrogen, Inc). RT-PCR was performed using the Qiagen One-Step RT-PCR kit (Qiagen, Inc) with RNA specific primers (C4 Ex 01 F2 TCACACCCCACAACAACTTC, C4 Ex 04 R, GATGTCAGGCAGCACCAG, BTN1F ATTGGGAAGAGGACGTGATG, BTN1R ACTGCCCTTCTGTGAGATCC). Reactions included 0.5 μg RNA as template with an annealing temperature of 60° C and 30 amplification cycles according to manufacturer's protocol. Reaction products were purified with a Qiaquick column (Qiagen, Inc) and directly sequenced or cloned into the p-Drive vector using the Qiagen PCR Cloning kit with transformation into XLI Blue

[15] Genbank AB115246
[16] Genbank DQ490139
[17] Genbank AF033107

electrocompetent *E. coli* cells. Plasmid clones were sequenced using vector-specific primers.

Class I loci (α1-2 domains) were amplified using the AccuScript High Fidelity RT-PCR System (Stratagene) with the following primers: Class1αF CGGGCGAGCCACACTC, Class IαR GGCTCTCTCCTCTCCAGCTC. Sequences were trimmed and analysed using Sequencher software. To match cDNA sequences to their genomic position and identify promoter sequences, long-range, locus-specific PCR reactions were performed with genomic DNA (OCB02) template using Taq Mastermix (Promega) supplemented with 1X Q solution (Qiagen Inc) using the following primers: locus-specific forward, BF1F2 (Iα1) CCATGATGGTGACAGTGAGTG, C4BFR (Iα2) CCCATAGCTATCCCACAACC and an allele-specific reverse primer Iα2a R CTCTGTCCGACCTCCGTCTG, Iα2b R CATCAGCATAAGTTGGATCCG, Iα1a R TTAAACCTCTCAGCAATGCTCTC, Iα1b R CCCATCATCCTCCCATCTTC. Reaction products were purified and directly sequenced. An internal primer, C4BFF GAAGTACCGCAGGGAGTGTG, was required to resolve Iα2 products.

To better assess relative expression of Class I loci, semi-quantitative PCR was performed. Standard curves were generated by *in vitro* transcribed cRNA from *Spe* I linearized plasmid clones using T7 RNA polymerase (NEB) according to the manufacturer's recommendations. Total spleen RNA and 0.5 µg *in vitro* transcribed products were treated with DNAse I (NEB) to remove DNA and purified using the RNeasy kit (Qiagen Inc). Removal of DNA in spleen RNA and cRNA preps was confirmed by PCR using turkey microsatellite and vector-transcript primers, respectively. A total of 10 µg spleen RNA and 0.5 µg of each cRNA served as template for cDNA

synthesis with Affinityscript reverse transcriptase (Stratagene) using random primers according to manufacturer's protocol.

Semi-quantitative PCR on the expressed Class I loci was carried out using a universal forward primer and the allele-specific reverse primers used above (BF RTF TCTTCRTGCACTATGACAGCAC, Iα2a R, Iα2b R, Iα1a R, Iα1b R) with 10-fold serially-diluted spleen and locus/allele cDNA product as template. Briefly, 25 μL reactions contained cDNA, gDNA, or water template, 15 pmol each primer, 1X Q solution (Qiagen Inc) 12.5 μL Taq Mastermix (Promega) and were amplified for 30 cycles with 58° annealing temperature and 30s extension time. Products were resolved on 2% agarose gels stained with ethidium bromide.

Class IIβ loci were amplified and cloned with Qiagen One-Step RT-PCR and Qiagen PCR Cloning kits as described above using promiscuous primers (BlB-E1F CGTGCTGGTGGCACTGCTG, BlB-E4-5R ACGGGGCGCCCTTTCTGAC). To verify assembly and match the resulting cDNA sequences to their genomic position, long-range, locus-specific PCR reactions were performed with genomic DNA (NT WF06 and OCB 02) template using Taq Mastermix (Promega) supplemented with 1X Q solution (Qiagen Inc) and the following primers: a universal forward IIβF GTGCTGGTGGCACTGCTG and locus-specific reverse primers IIβ1R ACGCTCTCCACCACTCCTAC, IIβ2R TTGTCCCATGTCACCTTCAC and IIβ3R CGCTTCTCCACGAAGTTTTC. Amplifications were performed for 30 cycles with 58° anneal temperature and 30s extension time. Products were cloned as above and/or directly sequenced with an internal primer (IIβE3R GTAGAAGCCCGTCACGTAGC).

*Southern Blot Analysis*

For Southern analysis, 500 ng of BAC DNA (97E05) and 20 µg of genomic DNA from NT WF06 were digested with *Eco*RI, *Bam*HI, and *Hind*III overnight with supplied buffer (NEB, Inc), electrophoresed for 16 hr in a 0.8% agarose gel, blotted onto a Nytran N membrane, and UV cross-linked. Plasmid DNA from cDNA clones was digested with *Eco*RI, electrophoresed in 1% agarose, and inserts were extracted using the Qiaquick Gel Extraction kit (Qiagen, Inc). Random $^{32}P$ dATP -labeled probe was prepared from 25 ng cDNA insert DNA using the Prime-a-Gene labeling kit (Promega, Inc). The probe was denatured and hybridized to the blot for 24 hr at 60 °C with RapidHyb buffer (Amersham, Inc). The blot was washed twice with 5X SSC 0.1% SDS at room temperature and twice at 60 °C with 0.5X SSC 0.1% SDS. The blot was exposed to ISC Bioexpress-Blue basic autorad double-emulsion film for 30 min at room temperature.

## RESULTS

*Sequencing and Assembly of the B-locus*

Subcloning and sequencing resulted in 3084 groomed reads with a median length of 525 bp. Primer walking on specific subclones assisted to fill minor gaps and/or obtain higher quality sequence in the assembly. Removal of pTARBAC2.1 vector sequence left a 172,697 bp insert of nearly 8x coverage (Figure 1A). As previously reported, this insert terminated in the sixth (last) intron in the tripartite motif (TRIM) 7.2 gene and the first intron of the TAP1 gene (96).

The remaining portion of the turkey core B- locus sequence was generated by PCR using genomic DNA (NT WF06) as template. This sequence spanned the region

between TAP1 and CenpA and overlaps the end of the BAC clone for a combined total of 197 kb of contiguous, turkey sequence. All sequence generated by PCR was invariant, e.g., no polymorphism (single nucleotide or insertion/deletion) suggesting NT WF06 is monomorphic at the MHC-B locus.

The turkey MHC-B region has a high overall GC content of ~53.6% (Figure 1B) similar to the 55.5% GC of the chicken. Several repetitive DNA types were identified in the B-locus (Figure1C) including 12 CR1/LTR repeats, 30 simple sequence repeats (repeat motif $\geq 5$) and three complex repeats. A large (~300 bp) C/T pentameric repeat (assigned microsatellite locus MNT-482) is located at 61.5 kb. This repeat is completely absent in the chicken. Twenty two tRNA sequences were identified by tRNAScan (Figure 1D). Comparatively, only 4 tRNA sequences were originally annotated in the chicken (30), however reanalysis of the chicken sequence aligned with the turkey indicates presence of the same tRNA sequences in the same syntenic order.

*Gene Identification and Annotation*

Based on BLAST homologies, FGENESH, and EST analysis, the 197 kb region contained 34 predicted genes (Figure 1E) compared to 31 genes in the homologous chicken sequence (~170 kb, Figure 1F). These genes include seven TRIM-like (TRIM 7.1, TRIM7.2, TRIM 39.1, TRIM 39.2, TRIM 27.1, TRIM27.2, TRIM 41), two zinc finger-like (Bzfp1, Bzfp2), 44G24.1, L-amino acid oxidase-like (LAAO), Hep21, guanidine nucleotide binding-like (GNBP), two butyrophilin-like (BTN1, BTN2), three BG-zipper-like (BG1-3), two C-type lectin-like (B-NK and Blec), three MHC Class IIβ loci, Tapasin, RING3, DMA, 2 DMB genes, two MHC Class I loci, TAP1 and 2,

compliment protein C4, and a small portion of the histone gene CenpA. Putative functions and MHC origins of these genes has been discussed by Shiina et al. (30). The final assembly was annotated and submitted to Genbank[18] (Figure 2).

*Interspecies comparisons*

As shown in Figure 2, MHC-B similarity is nearly linear between the turkey and chicken. The loci show high sequence homology with nearly perfect syntenic gene order (Figure 3). The turkey sequence contained genes homologous to all of those identified in the chicken haplotypes sequenced to date (28,30,36). Of note are the replicated blocks representing the additional BG loci in the turkey (chicken possesses one in the syntenic position), the syntenic and inverted homologies at the two Class I genes, the difference in the number of Class IIβ (two located between TAPBP and BRD2 in the turkey compared to one in chicken), and the inversion of the TAPBP gene. The TAPBP gene is in opposite orientation with respect to the chicken (verified by PCR, data not shown).

For each gene identified, the position, predicted coding sequence, and resulting amino acid sequences were determined (Table 2). The predicted coding sequences between the turkey and the B21-like haplotype present in the chicken whole-genome sequence (30) share homologies ranging between 85 and 96.5 percent, with an average 95% nucleotide identity. Aligned amino acid sequences between the two haplotypes are between 73 and 100 percent identical, with an average similarity of 96% (Table 2).

All splice donor and acceptor sequences are the canonical GT/AG. However, analysis of the aligned turkey and chicken sequences suggests alternative predicted protein coding sequences in 11 of the 31 CDS identified by Shiina et al. (30). Two

[18] Genbank DQ993255

instances, BTN1 and C4 were verified through RT-PCR. CpG islands are clearly conserved between these species (vertical arrows, Figure 1 E & F). Islands are present within the Class I and Class IIβ genes as well as TAP1, DMA, and BRD2. CpG islands are also present near TRIM41, the 3` end containing TRIM7.2, and the two zinc finger genes, Bzfp1 and Bzfp2.

Evidence of expression (ESTs) in poultry is available for 32 B-locus genes. Turkey ESTs were identified for 9 (TRIM7.2, GNB2L1, BTN1, BG1-3, RING3, and MHC Class Iα1-2). Two of the 34 genes (TRIM 39.1 and BTN2) lack EST/mRNA-based evidence for transcription in any avian species; however, the levels of similarity between turkey and chicken sequences suggest functional genes, with perhaps unique temporal and/or tissue-specific expression patterns. For the remaining genes, a majority—if not all—exons are represented in the EST/mRNA databases.

Examination of the nucleotide substitutions between the two species provides further evidence of gene function. Yang and Swanson (104) define a dN/dS ratio >1 , =1, and <1 as evidence for positive (diversifying), neutral, and purifying selection, respectively. Axelsson et al. (86) evaluated the substitution rates between turkey and chicken and identified a dN/dS ratio varying from 0.185 to 0.094 depending on chromosome size (macro and micro, respectively). However, most genes in the B region show higher dN/dS ratios (average of 0.259) than genes from other microchromosomes. Ratios for most MHC genes suggested purifying selection, those with very low dN/dS include TRIM 7.2, LAAO, TRIM7.1, TRIM41, and BRD2 (Table 2).

As one might predict, the highly polymorphic Class I and Class IIβ genes showed the largest dN/dS values. Given the close relation of the two species, even the classical

MHC genes failed to show strong evidence of positive selection when the complete coding sequences were analysed (Table 2).

While the greatest level of non-synonymous substitutions occur within the Class I α1α2 and Class IIβ β1 peptide binding domains, the C-type lectin-like NK cell receptor (Blec2) showed a comparably higher ratio of non-synonymous to synonymous substitutions between turkey and chicken throughout the coding sequence. Likewise, the level of variation identified among 14 chicken haplotypes was similarly skewed to non-synonymous substitutions (36). Similar levels of non-synonymous polymorphism has been identified in the NK cell receptors of mammals (105).

Phylogenetic analysis of the coding sequence from MHC antigen genes (Class I/BF, Class IIβ/BL, BG (IG to transmembrane domains to obtain alignments)) of the turkey, quail, and two chicken haplotypes (CB/B12, RJF/B21) was performed using ClustalX (Figure 3C). The results suggest each gene is monophyletic, originating from a single ancestral locus. Genes of each antigen sort together within species, rather than between species as loci (Figure 4). An alternate explanation is that the genes have co-evolved together to optimally function with interacting genes, e.g., $\beta_2$M, TAP, and DM, within each species. Interestingly, a separate phylogenetic examination of the Class IIβ PBR from multiple loci, species, and haplotypes showed an alternate level of clustering (Figure 5). The loci remained separated mostly by species; however some overlap occurred between turkey and quail sequences. This may suggest regions of this gene have species-specific co-evolved regions, as well as regions positively selected by pathogens.

Three BG loci are located between BTN2 and Blec2 in turkey where only a single locus is present in the chicken. An interesting feature of these BG genes is the

organization of tandem repeated 21 bp exons comprising the intracellular coil-coil domains, the significance of which has not been fully identified. BG1, BG2, BG3 each have a predicted 9, 24, and 14 of these exons, respectively. Sequence comparison shows species-specific delineation rather than locus segregation. This suggests these genes may interact with species-specific ligands/receptors in lieu of a molecule of pathogen origin.

*Class I Genes*

A single Class Iα locus flanked by DMB2 and TAP1was identified in the turkey BAC clone. A second locus was identified in the PCR-amplified region located between TAP2 and C4 (Figure 1). Both loci were located in the same orientation and position as in the chicken. Class Iα genes in turkey are comprised of 8 exons encoding a signal peptide, α1-3, transmembrane, and cytoplasmic domains similar to those originally identified in the chicken (106). The Class Iα1 locus lacked a recognizable poly (A) signal.

In order to examine relative expression between loci, universal primers were used to amplify partial transcript sequences of Class I (α1-α2 domains) from total spleen RNA from a commercial bird (OCB02). PCR using genomic DNA as template with one locus-specific and one allele-specific primer verified the presence and locus origin of the alleles. RT-PCR and cloning showed slightly skewed allele/locus representation. Of the 34 Class I clones sequenced, 20 represented a single allele (Iα2b), with the other loci/alleles represented by 5, 4, and 1 clone (Iα1a, Iα2a and Iα1b), respectively. Five clones were found to be chimeras of the four loci. Semi-quantitative PCR with allele-specific primers did not confirm this skewed distribution (Figure 6). Both alleles from

each locus appeared to be expressed at approximately the same level. Class Iα2 loci showed only a slight higher degree of expression compared to Class Iα1 loci.

In contrast to the single dominantly expressed Class I locus (BF2) in chickens, expression levels between the two turkey Class Iα loci were very similar. In order to investigate this difference, gene promoter regions for both Class I loci were studied. Comparing the upstream gene sequence from the two loci identified in NT WF06 and the four loci present in OCB02 found functional elements described for chickens and humans (92,107) in both Class Iα1 and Iα2 loci (Figure 7). Promoter sequences of turkey Class I genes showed considerable variation within and between loci. Twenty-one SNPs were identified between the three turkey Class Iα1 promoters (288 bp), five of which occurred in functional elements of the NT WF06 Class Iα1 promoter. Comparably, only nine SNPs were identified in the turkey Class Iα2 promoters (294 bp), only one of which lay in a functional element- the 3` end of the X1X2 Box.

The promoters of Class Iα genes are similar to those in the chicken, sharing conserved S-X-Y boxes, interferon response elements, and NFκB binding sites. The turkey Class Iα1 promoters are highly similar to chicken BF1 (B21 haplotype) promoters with one exception. While both species share a shortened proximal promoter sequence compared to the Class Iα2/BF2 locus, the promoter of the turkey Class Iα1 locus retains an NFκB/enhancer A sequence nearly identical to the Class Iα2/BF2 locus. Class Iα2 promoters are highly similar between turkey and chicken. The sequences of the functional elements (NFκB, ISRE, S, X1X2, and Y Box) are nearly identical. However, promoters of the turkey Class Iα2 are 15 bp shorter at the 3` end compared to chicken BF2 promoters at the proximal promoter region.

*Class II Genes*

Three Class IIβ loci were identified in the sequenced BAC clone (Figure 1). One Class IIβ locus is located between Blec1 and TAPBL similar to chicken. Two Class IIβ loci, positioned in the same transcriptional orientation, are flanked by TAPBL and BRD2, in contrast to the single locus observed in the chicken (Figure 1). Southern hybridization, locus-specific PCR of genomic DNA, and subsequent sequencing confirmed the presence of three loci in NT WF06 as observed in the BAC clone (Figure 8).

Comparative analysis of this sequence suggests the origin of the turkey Class IIβ2 locus is the result of a non-reciprocal homologous recombination/ gene conversion-inversion event in an ancestral chromosome. Based on the flanking sequence, this event included the 5`UTR of the ClassIβ2 locus through the ClassIIβ1 3` UTR, resulting in the observed inversion of the TAPBP gene (Figure 9). The turkey Class IIβ1 locus is homologous to the chicken BLB1 locus except it retains a large portion of the 5` UTR of the BLB2 locus (-300 to TAPBP). The ClassIIβ2 locus is fully homologous to the chicken BLB1 locus from 5` to 3` UTR. The ClassIIβ3 locus is homologous to the chicken BLB2 locus, however, a large portion of the 5` UTR (-150 to ~ -2500) has significantly diverged from the chicken with no homology to any other known genome sequence.

Similar to Class I loci, Class IIβ gene expression was examined using total spleen RNA from a commercial turkey (OCB 02). Long range PCR and cloning from genomic DNA identified 2 alleles for each Class IIβ locus. Analysis of the peptide binding region (exon 2) showed allele/loci shared 78-100 % identities (Figure 10, Table 3). Interestingly,

one allele from the OCB02 IIβ1 locus (IIβ1b) was found to be identical to an allele of the IIβ3 locus (IIβ3b) at exon 2. Only 2 nucleotides varied between the two transcripts resulting in one synonymous mutation and one non-conservative (valine to glutamate) substitution. It is undetermined whether the two alleles belong to the same haplotype.

To evaluate relative expression of the Class IIβ loci, heterogeneous OCB02 Class IIβ RT-PCR products were cloned and sequenced. Of the 104 clones sequenced, only loci corresponding to alleles of the Class IIβ1 and IIβ3 loci were identified, no clones were identified for either allele at the Class IIβ2 locus (Table 3). Distribution of RT-PCR clones suggest that the turkey Class IIβ1 locus is transcribed at a rate twice that of the Class IIβ3 locus. The high sequence identities between the alleles of the Class IIβ loci in these haplotypes prevented further quantitative PCR assays.

Disparity between expression of Class IIβ loci has also been reported in chicken and quail (38,91,108). Mammalian MHC Class II gene expression is thought to predominantly be under the control of a highly conserved SXY module, consisting of the S, X, X2, and Y boxes (109) which bind constitutively transcribed transcription factors that act as scaffold for the highly regulated Class II transactivator (CIITA) protein. This promoter sequence has also been identified as functionally important in the chicken (110). Upstream promoter regions of the variably expressed Class IIβ genes were examined through direct sequencing. Comparative analysis of the sequenced turkey genes and 14 chicken haplotype sequences (36) reveals a single nucleotide polymorphism in the S box that may correlate with the reduction/elimination of expression in both turkey and chicken (Fig 11). The substitution found in the less expressed chicken BLB1 locus is a G-A transition. The substitution in the Class IIβ2 gene promoter is a G-C transversion.

Comparisons of the peptide binding regions (PBRs) along with additional haplotypes may more accurately identify the selective pressures on these genes (104). Pairwise dN/dS values for the Class IIβ exon 2 sequences identified in this study and 14 chicken haplotypes (36) varied between 0.25 and 19.34 and averaged 1.42 (Table 4). Several alleles were invariant and thus would have an undefined dN/dS value. Overall, turkey PBRs showed lower average dN/dS values within the species, both among and between loci, as compared to chicken. Comparisons of the unexpressed turkey Class IIβ2 locus showed the highest predicted level of selection (based on dN/dS) in that species, whereas the less expressed chicken BLB1 locus showed similar levels of selection as to the predominantly expressed BLB2 locus in chickens.

## DISCUSSION

The avian MHC is of significant scientific interest. In the turkey and chicken (and likely quail) it is divided into two distinct regions with the classical MHC genes in the B locus, and non-classical MHC loci in the Y locus. The B-locus is tightly compact and very gene dense with the classical Class I and Class IIβ loci encompassed within a distance of less than 50 kb. In contrast, the genetically unlinked MHC-Y locus is much less defined. To date, only a limited amount of sequence data is available for the Y-locus, and only from the chicken. This region possesses non-classical MHC loci, lectin-like loci, and additional MHC paralogous loci yet to be defined. Although it lacks classical MHC loci the chicken Y-locus does have an effect on the host response to pathogens (44,89).

The remarkable similarity between the turkey and the chicken MHC-B loci is unexpected. Studies in some avian species have identified greater numbers of Class I and Class IIβ alleles within individuals compared those seen in the turkey and chicken, suggesting presence of additional loci (111-114). Other Galliformes that have been studied appear to have similar numbers of Class IIβ loci (115,116). Despite similar phylogenetic distances from turkey and chicken, the quail contains an expanded set of genes in the region. The sequenced quail haplotype contains 10 Class IIβ, 7 Class Iα, 8 BG-like, 4 NK, and 6 Blec-like genes (37). The duck (*Anas platyrhyncho*) has at least five Class I loci located distally from TAP2. Although initial comparisons between chicken and quail suggested the B-locus to be rapidly diverging and subjected to extensive selection, the similarity of the turkey locus with that of the chicken seems to contradict this observation. The overall MHC-B region appears largely stable between turkey and chicken, with tremendous conservation of gene content and order. Shiina et al. suggested the rapid divergence between chicken and quail might be due to the increased pathogens that quail might be exposed to as a result of its migratory behavior (37). In contrast, neither turkey nor chicken migrate, potentially reducing diversity of pathogen exposure.

The ability to compare the turkey and chicken sequences has resulted in improved gene identification. The annotations of eleven genes in the chicken should be amended based on the comparative alignment. Two of those predictions were confirmed by RT PCR. The BTN1 coding sequence is one third larger than previously thought, containing the PRY/SPRY domains associated with the TRIM genes as well as a RecF/RecN/SMC domain commonly involved in chromosome maintenance and recombination (117). Less

dramatic is the discrepancy in the complement protein C4 where the coding sequence was found to be ~150 bp larger, resolving a previously erroneous exon prediction. Other genes with minor annotation differences were not verified in this study, however EST datasets provide added confirmation.

A characteristic of MHC genes is their significant level of variation both within and between species. However, several genes with diverse functions (TRIM 7.2, LAAO, TRIM7.1, TRIM41, and BRD2) were highly conserved between turkey and chicken. While little is known of the avian paralogs, mammalian TRIM genes are members of a large family of genes, some of which have been found to possess immune functions with involvement in disease resistance (33,118,119). LAAO is a metabolic enzyme and BRD2 is suggested to be a regulator of transcription (120,121).

BG genes are unique to avian lineages and appear to be numerous and spread throughout the avian MHC. To date, little is known of their biological function. These genes have highly conserved immunoglobin-variable-like and transmembrane domains, yet show considerable divergence in their intracellular regions. Three BG loci were identified in this study. Additional BG loci are located upstream of the BF-BL region in the chicken and there is evidence for additional loci upstream in the turkey as well based on BAC end sequences. The actual number of additional loci however is currently not known in either species. Due to polymorphism and multiple genome copies of BG genes, no expression data was available to confirm the annotations provided in this work.

The putative NK cell receptor Blec2 shows some of the most significant divergence between turkey and chicken. Indeed, even within chicken haplotypes it is under the highest level of selection (36). This level of selection suggests this gene may

have a significant role in immunology. Perhaps an allele of this locus in chickens is responsible for the genetic resistance to Marek's disease long associated with the B21 haplotype (122). The minimal number of avian MHC Class I genes and their proximal location within the MHC-B could constrain this gene to co-evolve with the dominantly expressed Class I gene for proper NK cell surveillance. However, using a reporter gene construct, Viertlboeck et al (123) found this receptor to be unresponsive in co-cultures with unstimulated chicken cells and specific transfectants harboring the Class I and Blec genes. Stimulated splenocytes (ConA or PMA) possessed a ligand that the NK cell receptor recognized, but it was not determined if the ligand was allele-specific for each NK cell receptor, or if activated cells of an alternative haplotype could also stimulate splenocytes. Further evidence by Rogers and Kaufman suggest the ligand for B-NK is not an MHC molecule, thus confounding the role of Blec2 (124).

Two Class I (BF) loci were identified in the turkey occupying the same position and orientation in the MHC-B as the chicken. The two turkey loci had only minor variation at the level of expression in contrast to a dominantly expressed BF2 locus and a less (or un-) expressed BF1 in chicken. Similar locus effects on expression have also been found in the duck (125). Class Iα expression/promoter analysis in the turkey suggests the expression disparity observed between the two BF loci (in certain haplotypes) in chicken is primarily due to the disrupted NFκB site with the proximal promoter deletions having little effect. The turkey Class Iα1 locus promoter has a single T to G transversion in the NFκB binding site, whereas the chicken BF1 locus has three nucleotide substitutions. Additionally, there may be a minor transcriptional enhancing effect of the Class Iα2/BF2 promoters due to their close proximity to the C4 promoter as

there is less than 1 kb between the Class Iα1 and the C4 start codons. Additional studies with different turkey MHC-B haplotypes are needed to confirm the lack of Class I expression disparities as compared to those seen between loci in the chicken and duck. It is possible the two haplotypes surveyed in this study are the exception, rather than the rule.

Three Class IIβ loci were identified in the turkey, whereas the chicken possesses two. The quail has remarkable flexibility with regards to Class IIβ loci, with between 1 and 3 loci occupying the location between TAPBL and BRD2 (38). In a previous study of Class IIβ β1sequences in turkeys, Ahmed et al. used PCR to identify three separate domains in genomic DNA from several members of a possibly closed flock (62). PCR-RFLP found up to three alleles present within a given individual, suggesting turkeys may be polymorphic in both Class IIβ alleles and loci. However, based on the core turkey MHC-B locus, it is likely that two haplotypes sharing alleles (as defined by *Hinf I* digestion) were present in this population. A second contributing factor may have been null amplifications leading to an under estimation of total allele numbers. For example, in the Class IIβ loci identified in the present study, up to three base substitutions occurred at the primer-binding site of the sense primer used by Ahmed et al., supporting the possibility of null amplifications.

The Class IIβ1 locus of the turkey showed the highest level of expression. The Class IIβ3 locus, positioned proximal to the BRD2 gene, was expressed at about half the level of Class IIβ1, and expression of the Class IIβ2 locus could not be detected. The lack of expression of the Class IIβ2 loci is not likely the result of null or failed PCR amplification. A single nucleotide substitution (C/A at the 6$^{th}$ nucleotide position) was

present in the reverse primer binding site of Class IIβ2a allele. This was not the case for the alternate Class IIβ2 allele where no primer binding site substitutions were present. In chickens the Class IIβ locus, positioned between the TAPBP and BRD2 (Class IIβ2), is dominantly expressed, whereas the locus between Blec1 and TAPBP (Class IIβ1) is less well expressed.

Control of Class II gene expression is thought to be regulated by the CIITA gene product and its interaction with the highly conserved SXY module in the class IIβ promoter region. Comparisons of the SXY module between the turkey and chicken loci found several nucleotide substitutions in the S-box potentially responsible for expression differences. Perhaps the more dramatic transversion in the turkey Class IIβ2 locus ablates transcription whereas the more conservative transition in the chicken BLB1 S-Box permits at least partial transcription. Other polymorphisms were identified between the promoters of the chicken loci, nevertheless they were not unique and likely not responsible for the reduced expression observed between loci.

This explanation of expression control does not consider the differences observed between the turkey Class IIβ1 and β3 loci and may be contradicted by results in the quail. Whereas the SXY module is nearly identical between turkey and chicken, the quail module is clearly more divergent with a much greater level of polymorphism occurring between loci. No consensus SXY module can be clearly identified that segregates the expressed and the un/under expressed loci. Of the nine sequenced quail promoters there is a single SNP (T→C) present in the X box of the two untranscribed loci.

Elements in addition to the SXY module most likely effect the overall transcription of avian Class IIβ loci. While the turkey Class IIβ1 locus and the

untranscribed Class IIβ2 locus share significant homology at over 350 bp in their 5`UTR regions due to their duplicate origin, only 150 bp of the 5` UTR (the SXY module) of the Class IIβ3 locus is similar to the other two loci. Conceivably upstream of the SXY box lay Class II enhancer/repressor(s) either unique in the avian lineage and/or not yet identified in other species. Alternatively the function of the SXY box and/or CIITA differs significantly in birds.

In addition to simple promoter element effects, Class IIβ expression may be effected by gene order and position. It is possible the juxtapositioned TAPBP promoter (although in opposite orientation) enhances expression of the adjacent Class IIβ locus with less than 1.2kb separating the start codons of Class IIβ1 and the TAPBP (Class IIβ1 or BLB2 in turkey and chicken, respectively). This orientation appears to be unique to the turkey in that the gene orientation in the quail was similar to that of the chicken (37).

Similarity at the MHC-B locus between the turkey and chicken should be viewed in light of the limited data. Only one turkey haplotype has been thoroughly examined and the extent of intraspecies variation is not known. Locus-specific PCR on heterogeneous turkey DNA at least confirmed the presence and orientation of classical MHC loci in an additional bird. Sequence-level characterization of additional haplotypes is needed to verify the general gene content and to identify the overall level of variation within this region. In the 14 chicken haplotypes that have been re-sequenced, overall sequence variability was high (1 SNP/25bp), however no variation in gene number or orientation was found (30). Further, recombination within the chicken MHC-B locus has rarely been observed (29).

Haplotypes of the chicken MHC-B have profound influence on resistance or susceptibility to numerous pathogens. Infections of Rous Sarcoma Virus (RSV) or Marek's Disease Virus (MDV) show two of the most dramatic influences of MHC haplotypes on disease (53,126). In both instances different haplotypes have strong influences on the progression or regression of disease. However, while strong hypotheses suggest a role for the Class I locus (53,127) in neither case is it precisely known which particular gene/allele is responsible for the effects. The effect of MHC haplotypes on disease susceptibility in turkeys is currently unknown.

This study examined the genes present in the core turkey MHC-B locus, the region described as the minimal essential MHC in chickens. Little is known however of the distal MHC regions. Examination of overlapping BAC clones has provided the best assembly of the extended chicken B-locus (30). Five loci (Bzfp3, Blec3, KIFC, BG2 and BG3) were identified in the chicken upstream of the region included in this study and four loci (TNXB, LTB-4R1, CD1A2, and CD1A1) can be placed downstream (30). The chicken used in whole genome sequencing (RJF 256 from line UCD 001) is an MHC recombinant and is heterozygous upstream of the Bzfp1 gene (L.D. Chaves, unpublished). Heterozygosity in addition to the highly similar, polymorphic, and multiple-copy extended BG gene family have made assembly of this region of the chicken genome difficult. Less than 4 kb of overlapping and extending whole genome sequence can be aligned to the haplotype sequenced by Shiina et al. (L. D. Chaves, unpublished results). Further sequencing of the chicken MHC-B will invariably require additional, overlapping, haplotype-specific BAC/cosmid clones.

Overlapping BAC clones extending the 5` end of the turkey B-locus have been identified in the CHORI 260 BAC library. However, no clones were found overlapping the 3` end. Based on the results of the present study in addition to unpublished results (L.D. Chaves unpublished), the turkey selected for BAC library construction and whole genome sequencing (NT WF06) appears monomorphic within the MHC-B locus. If true this will aid the assembly of the turkey extended B-locus and may ultimately improve the ordering and assembly of the homologous chicken sequence.

**Table 1:** PCR and internal sequencing primers and amplicon size of the 3` sequenced turkey MHC-B region.

| | Fwd Primer | | Rev Primer | Size (bp) |
|---|---|---|---|---|
| Tap1F | GTTCCAAACCCACACCATTC | Tap2Ex2R | GCTCTTCTCCAGCCTGGTG | 2004 |
| Tap2Ex1F2 | CCTACATTCTGCGCCTGTC | Tap2Ex3R | GAAAGTCTATGACACCCGGC | 1385 |
| Tap2Ex3F | TCAACGTCATGCTGAGGAAC | Tap2Ex6R | CTGCAAGACGTCACCTTTGA | 1092 |
| Tap2Ex6F | CTGGCGTATTCCTATGGTGAC | Tap2Ex9R | GTCTGCCTTCTCCAGCATC | 948 |
| Tap2int8F | GCAGTGTGTAGGGTGGAAGG | BFFalt1 | CATAAGTTGGATCCGCCTCC | 2697 |
| Tap2_BFF | GAAACCAGGAGAGAGCTGGAG | | | Internal Seq |
| | | Tap2_BFR | GCTCTGCCCTCATCACTCAC | Internal Seq |
| BFDNAF | GTTGAAGCGTTCCTGCAGTG | BFDNAR | CCTCTCCAGCTCTGCCTTC | 536 |
| BFRalt1 | TCCAGGTAATGCTTCCGTTG | C4Ex2R | CTGAGCACTGTGGGGCTCT | 2383 |
| | | C4BFR | CCCATAGCTATCCCACAACC | Internal Seq |
| C4BFF | GAAGTACCGCAGGGAGTGTG | | | Internal Seq |
| C4EX1F2 | TCACACCCCACAACAACTTC | C4Ex4R | GATGTCAGGCAGCACCAG | 1020 |
| C4Int3F | CTACAAGGCTGGGGTCTCAC | C4Int6R | CCTGATGTCCTTTGGGGTC | 1123 |
| C4Ex6F | TGGCTTCATTGTGCTGAGTG | C4Ex10R2 | CACCACCCCTCATAGAGCTG | 1486 |
| C4Ex9F | CCCTACACCCTCCTGGTGAG | C4Int12R | CAGACCCCAAATCCATCATA | 1687 |
| C4Ex12F2 | GGACCCACTGAAGGTGACAG | C4Ex16R2 | GCTTCTTCCCTGAGAGCTGG | 1459 |
| C4Ex12F2 | GGACCCACTGAAGGTGACAG | C4Int17R | CCGTAGTGATGAGGTCCTGG | 1788 |
| C4Ex17F | CTGCTCCCTGACTCCATCAC | C4Ex21R | GATGGGGACTTCAGCATGAG | 1500 |
| C4Int20F | CTGTGGTGTGGTGTCCCTAC | C4Ex24R | CTTCTTGGAGCACAGGGCT | 1416 |
| | | C4Ex22R | GCAGTCCCTGATGTCAATGG | Internal Seq |
| C4Ex24F | CTGTCCCGGCCCTATTTG | C4Ex26R | CACATCCTATGCATTGGCAC | 1634 |
| C4Int25F | TTCCTCCATCTCTCATTCACC | | | Internal Seq |
| C4Ex26F | GGAGCAGTTGGGGACCTATG | C4Ex27R | TTCTGGCACTGCTGCAGA | 1223 |
| C4Ex27F | ACCTTCTGGCCATCTGGTG | C4Ex30R | CTGTCACCTCCTAACGCCA | 842 |
| C4Ex29F | CAGTGACAGTGCAGGTGGAG | C4Ex33R | GAAGGAAACCAGTTGGTGCT | 1114 |
| C4Int32F | GTTGGGATGGGGACGTATC | C4Ex35R | CACCATCTTCTACAACGCCC | 489 |
| C4bassayR (Int34F) | CCTAAATGGGTGGGGTCTTG | MG3'BLR (Ex38R) | CAGCTCTTCCTCACCAGCAG | 1236 |
| C4int37F | TACGGGGAAAGTTTGGTGAG | CenPA Ex2R | GAGCGGTGCTGCTCTGATAG | 1101 |

**Table 2:** Coding sequence comparisons of homologous turkey and chicken MHC-B loci. [A] Comparative coding sequence deviates from Shiina et al. (30), [B] Based on ESTs, 2 transcripts may exist, [C] No EST/mRNA Evidence, [D] Partial comparative analysis 543 NTs, [E] Compared with BLB1, [F] No direct comparisons made.

| Gene | Strand | Position | Exons | Turkey | | Chicken | | Nucleotide Identitiy | dN/dS | AA Sub | Amino Acid Identity |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Bp | AA | Bp | AA | | | | |
| TRIM 7.2 (partial) | - | 1354-5458 | 6 | 991 | 330 | 1518 | 505 | 0.943 | 0.027 | 5 | 0.985 |
| Bzfp2[A] | + | 7834-11105 | 3 | 1920 | 639 | 1821 | 606 | 0.876 | 0.324 | 118 | 0.815 |
| Bzfp1a[B] | - | 12433-19212 | 3 | 1824 | 607 | 1812 | 603 | 0.884 | 0.260 | 95 | 0.843 |
| Bzfp1b[B] | - | 13745-15038 | 3 | 165 | 54 | 240 | 79 | 0.939 | 0.594 | 5 | 0.907 |
| 44G24.1[A] | - | 20405-20815 | 1 | 411 | 136 | 411 | 136 | 0.886 | 0.200 | 22 | 0.838 |
| L-amino acid oxidase | + | 24031-29854 | 7 | 1563 | 520 | 1572 | 523 | 0.935 | 0.097 | 28 | 0.946 |
| TRIM7.1[A] | - | 34958-44455 | 7 | 1434 | 477 | 1434 | 477 | 0.958 | 0.027 | 6 | 0.987 |
| Hep21 protein | - | 47355-48150 | 3 | 318 | 105 | 324 | 107 | 0.912 | 0.144 | 8 | 0.924 |
| TRIM39.2 | - | 54835-58681 | 6 | 1392 | 463 | 1392 | 463 | 0.935 | 0.111 | 27 | 0.942 |
| TRIM27.2 | + | 61026-64684 | 7 | 1431 | 476 | 1431 | 476 | 0.944 | 0.261 | 38 | 0.920 |
| TRIM39.1[A,C] | - | 65730-69573 | 5 | 789 | 262 | 801 | 266 | 0.934 | 0.317 | 25 | 0.905 |
| TRIM27.1 | - | 70729-74368 | 7 | 1488 | 495 | 1488 | 495 | 0.935 | 0.158 | 34 | 0.931 |
| TRIM41[A] | + | 76006-80803 | 7 | 1884 | 627 | 1878 | 625 | 0.962 | 0.036 | 9 | 0.986 |
| GNBPL | - | 82371-85788 | 8 | 954 | 317 | 954 | 317 | 0.965 | NA | 0 | 1.000 |
| BTN 1[A] | + | 88919-99310 | 38 | 1284 | 427 | 1284 | 427 | 0.949 | 0.168 | 23 | 0.946 |
| BTN 2[A,C] | + | 102542-105958 | 8 | 957 | 318 | 975 | 324 | 0.918 | 0.368 | 43 | 0.865 |
| BG 1[D] | - | 107389-111044 | 15 | 948 | 315 | 1023 | 340 | 0.897 | | 31 | 0.829 |
| BG 2[F] | - | 112639-117813 | 28 | 1080 | 359 | | | | | | |
| BG 3[F] | - | 119380-123213 | 19 | 870 | 289 | | | | | | |
| B-NK[A] | - | 125503-128012 | 6 | 696 | 231 | 693 | 230 | 0.861 | 0.510 | 57 | 0.753 |
| B-Lec | + | 129859-131819 | 5 | 567 | 188 | 567 | 188 | 0.928 | 0.257 | 18 | 0.904 |
| MG-ClassIIβ1 | - | 132673-133999 | 6 | 792 | 263 | 792 | 263 | 0.932 | 0.629 | 28 | 0.894 |
| TAPASIN | + | 135154-139324 | 8 | 1293 | 430 | 1293 | 430 | 0.920 | 0.295 | 45 | 0.895 |
| MG-ClassIIβ2[E] | + | 140312-141634 | 6 | 792 | 263 | 792 | 263 | 0.937 | 0.826 | 26 | 0.901 |
| MG-ClassIIβ3 | + | 144442-145824 | 6 | 792 | 263 | 792 | 263 | 0.927 | 1.002 | 31 | 0.882 |
| BRD2 | - | 147230-152871 | 12 | 2340 | 779 | 2340 | 779 | 0.961 | 0.012 | 4 | 0.995 |
| DMA | + | 156695-159022 | 4 | 792 | 263 | 792 | 263 | 0.912 | 0.335 | 30 | 0.886 |
| DMB 1[A] | + | 159246-161497 | 5 | 987 | 328 | 933 | 310 | 0.914 | 0.447 | 44 | 0.866 |
| DMB 2 | + | 161954-164765 | 6 | 777 | 258 | 777 | 258 | 0.927 | 0.205 | 22 | 0.915 |
| MG-ClassIα1 | + | 165569-167586 | 8 | 1083 | 360 | 1083 | 360 | 0.860 | 0.547 | 84 | 0.767 |
| TAP1 | - | 168635-172948 | 11 | 1764 | 587 | 1752 | 583 | 0.944 | 0.197 | 32 | 0.945 |
| TAP2 | + | 173510-176702 | 9 | 2100 | 699 | 2106 | 701 | 0.929 | 0.122 | 34 | 0.951 |
| MG-ClassIα2 | - | 177983-180000 | 8 | 1083 | 360 | 1068 | 355 | 0.849 | 0.628 | 97 | 0.731 |
| C4[A] | + | 180995-196378 | 40 | 5100 | 1699 | 5082 | 1693 | 0.939 | 0.150 | 92 | 0.946 |
| Histone (Partial) | + | 196827-197022 | 2 | 111 | 36 | 396 | 131 | 0.964 | NA | 0 | 1.000 |

[A] Comparative coding sequence deviates from Shiina et al. (30), [B] Based on ESTs, 2 transcripts may exist, [C] No EST/mRNA Evidence, [D] Partial comparative analysis 543 NTs, [E] Compared with BLB1, [F] No direct comparisons made.

**Table 3:** Selective pressures identified through pairwise dN/dS comparisons of Class IIβ loci (PBR) varies between species and loci.

| | Comparisons | Identical Alleles | Min | Max | Mean |
|---|---|---|---|---|---|
| **Turkey** | 35 | 1 | 0.47 | 2.35 | 0.77 |
| Class IIb1 | 3 | 0 | 0.47 | 0.97 | 0.68 |
| Class IIb2 | 3 | 0 | 1.02 | 2.35 | 1.48 |
| Class IIb3 | 3 | 0 | 0.49 | 1.02 | 0.77 |
| 1v2 | 9 | 0 | 0.49 | 0.89 | 0.68 |
| 1v3 | 8 | 1 | 0.47 | 0.81 | 0.66 |
| 2v3 | 9 | 0 | 0.49 | 0.96 | 0.75 |
| **Chicken** | 360 | 18 | 0.26 | 19.34 | 1.88 |
| Blb 1 | 84 | 7 | 0.26 | 14.86 | 2.04 |
| Blb 2 | 80 | 11 | 0.80 | 5.65 | 2.02 |
| 1v2 | 196 | 0 | 0.51 | 19.34 | 1.76 |
| **Overall** | 647 | 19 | 0.25 | 19.34 | 1.42 |

**Table 4:** Turkey Class IIβ β1 domain and expression variation. Clones representing alleles are in the diagonally redundant cells.

| | Exon 2 nucleotide similarities | | | | | |
|---|---|---|---|---|---|---|
| | 1A | 1B | 2A | 2B | 3A | 3B |
| 1A | 37 | 0.90 | 0.87 | 0.84 | 0.91 | 0.90 |
| 1B | 0.86 | 34 | 0.81 | 0.78 | 0.85 | 1.00 |
| 2A | 0.77 | 0.67 | 0 | 0.93 | 0.86 | 0.81 |
| 2B | 0.78 | 0.70 | 0.84 | 0 | 0.84 | 0.78 |
| 3A | 0.83 | 0.78 | 0.71 | 0.73 | 18 | 0.85 |
| 3B | 0.86 | 1.00 | 0.67 | 0.70 | 0.78 | 15 |
| | Exon 2 amino acid similarity | | | | | |

**Figure 1:** Sequence features of the turkey MHC-B region. A. BACs identified containing the MHC-B region of the turkey and the additional region amplified by PCR. B. GC content of the turkey MHC calculated by continuous 100 bp windows. C. Repetitive elements identified in the region. D. tRNAs present in the turkey MHC. E. Genes and orientation predicted within the turkey MHC-B. Black arrows denote CpG islands. F. The homologous chicken sequence is provided for comparison.

A

BG BL BF

01H21
97E05
38O12
PCR
56J17

B

80%
50%
20%

C

D

E

TRIM 7.2 Bzfp2 Bzfp1 44G24.1 LAAO TRIM 7.1 Hep21 TRIM 39.2 TRIM 27.2 TRIM 39.1 TRIM 27.1 TRIM 41 GNB2L1 BTN1 BTN2 BG1 BG2 BG3 Blec2 Blec1 IIβ1 TAPBP IIβ2 IIβ3 BRD2 DMA DMB1 DMB2 Iα1 TAP1 TAP2 Iα2 C4 CenpA

0 50 100 150 197 (kb)

F

TRIM 7.2 Bzfp2 Bzfp1 44G24.1 LAAO TRIM 7.1 Hep21 TRIM 39.2 TRIM 27.2 TRIM 39.1 TRIM 27.1 TRIM 41 GNB2L1 BTN1 BTN2 BG1 Blec2 Blec1 BLB1 TAPBP BLB2 BRD2 DMA DMB1 DMB2 BF1 TAP1 TAP2 BF2 C4 CenpA

**Figure 2:** Comparative idenitity matrix plot of the turkey and chicken MHC-B loci.

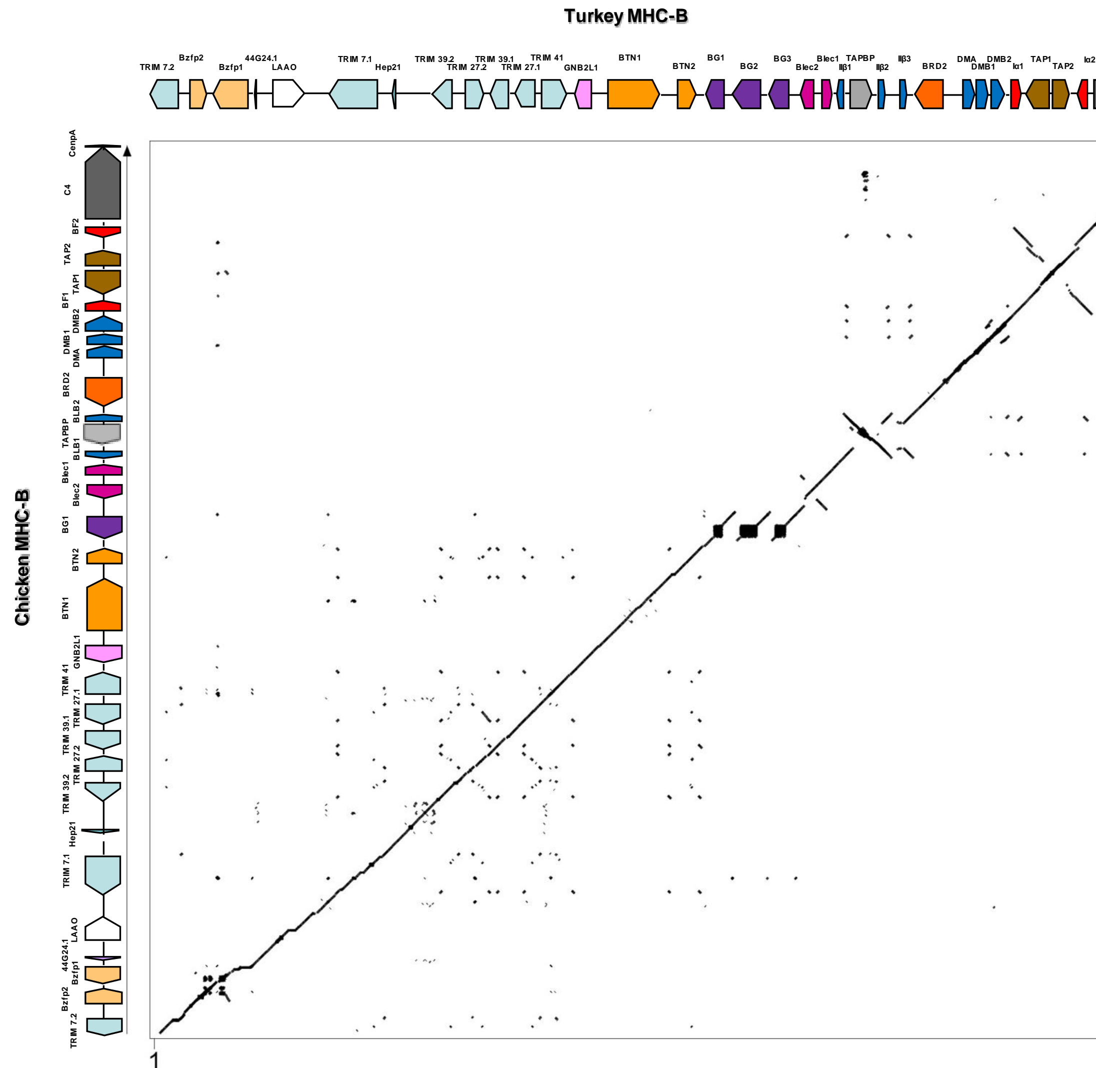


**Figure 3:** Genbank accession file of the turkey MHC-B region.

```
LOCUS       DQ993255              197022 bp    DNA     linear   VRT 10-MAR-2008
DEFINITION  Meleagris gallopavo major histocompatibility complex (MHC) B locus,
            partial sequence.
ACCESSION   DQ993255
VERSION     DQ993255.2  GI:169730347
KEYWORDS    .
SOURCE      Meleagris gallopavo (turkey)
  ORGANISM  Meleagris gallopavo
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria;
            Aves; Neognathae; Galliformes; Phasianidae; Meleagridinae;
            Meleagris.
REFERENCE   1  (bases 1 to 197022)
  AUTHORS   Chaves,L.D., Krueth,S.B. and Reed,K.M.
  TITLE     Defining the Turkey MHC: Sequence and Genes of the B-Locus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 197022)
  AUTHORS   Chaves,L.D. and Reed,K.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2006) Veterinary and Biomedical Sciences,
            University of Minnesota, 295 ASVM, 1988 Fitch Ave, St. Paul, MN
            55108, USA
REFERENCE   3  (bases 1 to 197022)
  AUTHORS   Chaves,L.D., Krueth,S.B. and Reed,K.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2008) Veterinary and Biomedical Sciences,
            University of Minnesota, 295 ASVM, 1988 Fitch Ave, St. Paul, MN
            55108, USA
  REMARK    Sequence update by submitter
COMMENT     On Mar 10, 2008 this sequence version replaced gi:148250079.
FEATURES             Location/Qualifiers
     source          1..197022
                     /organism="Meleagris gallopavo"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9103"
     misc_feature    complement(<1..>197022)
                     /note="major histocompatibility complex (MHC) B-locus"
     misc_feature    1..172679
                     /note="derived from BAC clone CH260-97E5"
     gene            complement(<1354..>5458)
                     /gene="TRIM7.2"
     mRNA            complement(join(<1354..1386,1486..1601,1759..1781,
                     2156..2374,2488..2574,4946..>5458))
                     /gene="TRIM7.2"
                     /product="tripartite motif protein 7.2"
     CDS             complement(join(<1354..1386,1486..1601,1759..1781,
                     2156..2374,2488..2574,4946..5458))
                     /gene="TRIM7.2"
                     /codon_start=1
                     /product="tripartite motif protein 7.2"
                     /protein_id="ACA64749.1"
                     /db_xref="GI:169730348"
                     /translation="MEAPPAPPAVPPAPCSAARSLQDELTCPVCLEYFTDPVLVAECG
                     HNFCRACVTQCWEDSARRLCCPQCREPVPQRLFRPNRSLGNIVHIVRQLGLPPGPADP
                     PPGPPAPAPPLPAAVPPVPPGPPRCPRHGEPLRLFCVQDRRAVCVVCHLSREHRTHTV
                     LPAEEAAHAAEEVPQEHLSSLRKGREEAKAERERQSEDLLKQTEVERQKIVAECKELR
```

```
                     GFLEEKEQLLLSRLEELDRDIVKRRDESVSRLSEEIAQLDKLLGEHGAENGPGNTAVT
                     TAGSSFESWMFCKPEPGFAELEKKLKSFSQKSAVLKEVLLEFKENLRFELENDT"
     gene            <7834..>11105
                     /gene="Bzfp2"
     mRNA            join(<7834..7921,8327..8416,9364..>11105)
                     /gene="Bzfp2"
                     /product="B-locus zinc finger protein 2"
     CDS             join(7834..7921,8327..8416,9364..11105)
                     /gene="Bzfp2"
                     /codon_start=1
                     /product="B-locus zinc finger protein 2"
                     /protein_id="ACA64750.1"
                     /db_xref="GI:169730349"
                     /translation="MEPYVLLDPRQRALYRDVMQESYETLMSLEFPISKHNLLSHRDE
                     PTALDLHVPKETLLGDNGAAQEPPQEDPKATKPTEEGPPCSPTRQKSGEPPHLHQTQP
                     RTVCAECGKSFSHKSALVKHQKIHSGDRPHECPQCGKAFIQRSDLTIHQRVHTGERPY
                     ACPDCGRRFSVSSSLLTHQRTHVPAGEKPNRCPECGRSFADSAALDRHRKSHLGGKPY
                     ECGVCGKAFAWSSHLERHRRIHTGEKPFQCGQCGRAFAWSSHLDRHMRTHVMVEEEEE
                     EEKEEGGEDEDLPLLPQKCADCGKRLNHQTDPQRFKHKGTQTSPTRDEDGPSVTPTSS
                     NGDGPSVTPMSPVGDEDGSSLTPIFPTEDEDGPSITPASTEDEDRSSITPTSATGNED
                     GSPITPIFPTGDEDGSPITPIFPSGDEDRPSMAPYPRKQCSTSLTQSSNPKHHCIHPT
                     EQPHTRPHCKHRFQRSSASTKHQQTHPKPRTEEPTTKPFPCRACGKTFAWLSHLERHR
                     RVHTGEKPFHCGECGRAFAVSSHLERHRRLHRGENPYQCGLCGKGFAWSSHYDRHRLT
                     HTGEKPFSCAQCGKCFGRSSHRNRHQRAHMKHPSHPKWGTQPSPLHGVRGRNPASLLP
                     SSPAPQQPWVPPPSS"
     gene            complement(<12433..>19212)
                     /gene="Bzfp1a"
     mRNA            complement(join(<12433..13778,14418..14532,18850..>19212))
                     /gene="Bzfp1a"
                     /product="B-locus zinc finger protein 1a"
     CDS             complement(join(12433..13778,14418..14532,18850..19212))
                     /gene="Bzfp1a"
                     /codon_start=1
                     /product="B-locus zinc finger protein 1a"
                     /protein_id="ACA64751.1"
                     /db_xref="GI:169730350"
                     /translation="MATPSERQSQQQRKTRRRRQKFHRVQWRCLSRRQRWKRSQKASS
                     NGMRTEPHRRNLSHQSELLTPASRRMSITARKMMCSFITSFSRGLVSDADQRRRERRG
                     SFIGTLELRAAVDAAMEGGGFPFPVGDALPAMEPCLLLDKALYRDVMQESYETLMSIA
                     QGLVNHKAADEAVVGLEEDAVGSEEPPMLPPHNPAQNRSLGSNRRKAKRPDKAVPIGD
                     TQTPHTTPTPKPNGVKLSQVRDRPFSCPDCGKSFPWASHLERHRRVHTGERPYSCPEC
                     GESYSQSSHLLQHRRTHSSDRPHKCQHCGKRFAGATELIAHSRGHAAEKPHKCGDCGK
                     GFVWASHLERHRRVHTGEKPYECPECGEAFSQGSHLSKHRRSHTGERPHRCSACGKCF
                     GQSSELSRHRRSHVGKKPQRCADCGKAFRATSALLRHQRCHRRERSHRCGDCGKGFAW
                     ASHLQRHRRVHTGERPFPCGLCGESFSQKAHLLQHGKTHRPERPYKCGDCGKRFEDAP
                     PFLAHRRSHTALKSFSCGDCGKGFAWASHLERHRRVHTGEKPYECPECGEAFSQGSHL
                     TKHRRSHGPKVALVLVEEGVEEGETLREAPLSYGAKESLRR"
     gene            complement(<13745..>15038)
                     /gene="Bzfp1b"
     mRNA            complement(join(<13745..13778,14418..14532,15023..>15038))
                     /gene="Bzfp1b"
                     /product="B-locus zinc finger protein 1b"
     CDS             complement(join(13745..13778,14418..14532,15023..15038))
                     /gene="Bzfp1b"
                     /codon_start=1
                     /product="B-locus zinc finger protein 1b"
                     /protein_id="ACA64752.1"
                     /db_xref="GI:169730351"
                     /translation="MDLIPGVSLPCRGCPACHGALPAAGQGAVPGCDAGKLRDADVHR
                     SGLGESQSSR"
     tRNA            join(19452..19489,19511..19556)
                     /product="tRNA-Leu"
     tRNA            19860..19931
```

```
                    /product="tRNA-Glu"
mRNA                complement(<20405..>20815)
                    /product="unknown"
CDS                 complement(20405..20815)
                    /note="44G24.1"
                    /codon_start=1
                    /product="unknown"
                    /protein_id="ACA64753.1"
                    /db_xref="GI:169730352"
                    /translation="MVSAGGRKKRPQKGKRPQPKRGPQKRRRQKTASRRPPAKRVKFN
                    GKTKPKPKRQQKKRSSSPFFDPSKMLRQDGRLLSREAKGLMMAFLRDVCRRVEREAER
                    LRKESQRPMVGPAHVRAALHHIMPKDWKRADGGR"
repeat_region       complement(21269..21382)
                    /rpt_family="GGLTR8B"
                    /rpt_type=dispersed
gene                <24031..>29854
                    /gene="LAO"
mRNA                join(<24031..24043,25409..25623,27204..27310,27945..28146,
                    28346..28414,28640..28776,29035..>29854)
                    /gene="LAO"
                    /product="L-amino acid oxidase"
CDS                 join(24031..24043,25409..25623,27204..27310,27945..28146,
                    28346..28414,28640..28776,29035..29854)
                    /gene="LAO"
                    /codon_start=1
                    /product="L-amino acid oxidase"
                    /protein_id="ACA64754.1"
                    /db_xref="GI:169730353"
                    /translation="MAAMVLFQILLLLGLASAKRFPCFPEYCLHDKDYEELLSIVKNG
                    LEPTTRPSKVVIVGAGISGLTAAKLLRDAGHKVTILETSNRVGGRIRTYREKDWYVEL
                    GAMRLPSKHRLVREFIRQFDLKLNPFHQTDDNAWYFVNGIRARVAEVNRNPDILNYTL
                    RASERGKSASQLYREALQSAFKEFQTMNCKEYLAKHDSFSTKEYLIKVGNLSRGAVQM
                    IGDLLNEDSGFYLSFLASLWDFDIFSQGESFDEITGGFDQLPKAFHKALPGVIKLNCT
                    VEKIMTKGGKVRVFYRAPDTLAPSIETADYVLVASSAKATRHIQFFPPLSPTKNDALR
                    SIHYASASKIALVCTEKFWEKDGIRGGQSITDRPSRFIYYPSHNFSSGVGVILASYTW
                    NNDADFLMPLADEKYLDLVLQDLADIHQVSKSYLQYTCDRYVIQKWQLDRHSLGAFAA
                    FTPYQFTDYSQALFEHEGRVHFAGEHAAQPHAWIDTSIKAAVRAASNIHDDSETQNLN
                    LNLEGFVEKGEL"
tRNA                complement(31766..31839)
                    /product="tRNA-Asn"
tRNA                complement(32129..32201)
                    /product="tRNA-Thr"
tRNA                32525..32606
                    /product="tRNA-Leu"
tRNA                complement(32974..33046)
                    /product="tRNA-Val"
tRNA                complement(33547..33618)
                    /product="tRNA-Pro"
tRNA                complement(33777..33849)
                    /product="tRNA-Thr"
gene                complement(<34958..>44455)
                    /gene="TRIM7.1"
mRNA                complement(join(<34958..35472,37874..37909,38903..39018,
                    39410..39432,42083..42313,42748..42843,44039..>44455))
                    /gene="TRIM7.1"
                    /product="tripartite motif protein 7.1"
CDS                 complement(join(34958..35472,37874..37909,38903..39018,
                    39410..39432,42083..42313,42748..42843,44039..44455))
                    /gene="TRIM7.1"
                    /codon_start=1
                    /product="tripartite motif protein 7.1"
                    /protein_id="ACA64755.1"
                    /db_xref="GI:169730354"
```

```
                     /translation="MAAVFLPGNLQDEATCSVCLEFFKDPVSIECGHNFCRACIVKSW
                     KDLEMDFPCPQCREIFQQKSLRPNRQLANMSEIISQFALRGAKGAEEDGLCMKHREAL
                     KLYCKDDRKSICVVCDRSREHRPHAVVPVDEASEEYKEKIQGCLDFLKKERQELLEFK
                     VNDDKKTQELLKTIETERQKLLVEFEGLRQFLHDQEHILLGQLDKMEKSIAKRQNENI
                     TDLSKEITLLNKLITELEEKIQQPVLEFLKDVMSIISRSDEVKCQKPVPVCTDMKMHV
                     CNFSLKMVVLEKVLKRFKENLRDELGRGEKEDLTLDPDSANHLLILSADLKSVRMGCR
                     KQELPDNPKRFDTNSRVLATTGFKSGRHYWEVEVGASDGWAFGVAKESVRRKGLTQFS
                     PEEGIWAVQQNGRRYWAVTSPQRTPLCLGQKLSKVRVYLDYEGEEVSFYNAENMQHIF
                     TFNVAFREKVFPLFSVCSTVTYIKLCS"
     gene            complement(<47335..>48150)
                     /gene="HEP21"
     mRNA            complement(join(<47335..47483,47711..47830,48102..>48150))
                     /gene="HEP21"
                     /product="Hep21"
     CDS             complement(join(47335..47483,47711..47830,48102..48150))
                     /gene="HEP21"
                     /codon_start=1
                     /product="Hep21"
                     /protein_id="ACA64756.1"
                     /db_xref="GI:169730355"
                     /translation="MKLLFVGLAFVLCVGVVEALQCKVCKYKIPYFGCYVGANETTCE
                     RREKCAIIKTSLGKMTLYYQQGCTSALNCGRERASDAESRLTSRYSCCETDLCNEEWD
                     EDP"
     repeat_region   complement(48003..48049)
                     /mobile_element="LINE:CR1-E"
     tRNA            complement(48286..48358)
                     /product="tRNA-Lys"
     tRNA            complement(49927..49999)
                     /product="tRNA-Lys"
     tRNA            complement(50296..50367)
                     /product="tRNA-Glu"
     tRNA            complement(51552..51624)
                     /product="tRNA-Val"
     tRNA            complement(51814..51885)
                     /product="tRNA-Glu"
     tRNA            52156..52227
                     /product="tRNA-Glu"
     tRNA            complement(52428..52500)
                     /product="tRNA-Val"
     tRNA            complement(join(52959..53006,53027..53064))
                     /product="tRNA-Leu"
     tRNA            complement(53209..53281)
                     /product="tRNA-Lys"
     tRNA            53515..53587
                     /product="tRNA-Val"
     tRNA            complement(53781..53852)
                     /product="tRNA-Glu"
     repeat_region   53991..54051
                     /mobile_element="LINE:CR1-D2"
     gene            complement(<54835..>58681)
                     /gene="TRIM39.2"
     mRNA            complement(join(<54835..55349,55867..55982,56331..56353,
                     56866..57096,57354..57449,58271..>58681))
                     /gene="TRIM39.2"
                     /product="tripartite motif protein 39.2"
     CDS             complement(join(54835..55349,55867..55982,56331..56353,
                     56866..57096,57354..57449,58271..58681))
                     /gene="TRIM39.2"
                     /codon_start=1
                     /product="tripartite motif protein 39.2"
                     /protein_id="ACA64757.1"
                     /db_xref="GI:169730356"
                     /translation="MDEDDPAESFQDEASCSLCLGFFQDPVSIHCGHNFCRACITRCW
```

```
                     EEEEETFPCPRCKATATQPNLRPNRELAKIIEIAKRLSLRAPNAGERLCERHREVLKL
                     FCEEDQAPICLVCRESQAHRLHAAVPIEEAVEEQKEKIQAHMQILKEKKEKLQGLKEA
                     EEGKSLEFLEKVQQERQKVVLDIKELQKFVEEQERLLLGRLEKLDQEIVRRKEENLAK
                     ILEEISSISEQIRELEEKCQQPPCEFLQDSRNILSRLENEKSQKVAEILPGAEENPTS
                     PAQKNTSLKETLMQFRENLTLDPDTAHPRLVLSEDQKCVRWEEARHPIPDNPKRFDSS
                     RCILGCQGFNSGRHYWEVEVGDGEAWAVGVAKESVERKGRISVKPEVGIWAVGQCGAQ
                     YQALTSPTSPITLLSAPKVIGIYLDYEAGRVAFFDANNEVPIFAYPPTSFGGERVLPL
                     LCLGRGCCFTLCP"
     repeat_region   complement(56459..56523)
                     /mobile_element="LINE:CR1-F"
     repeat_region   complement(56519..56622)
                     /mobile_element="LINE:CR1-F2"
     gene            <61026..>64684
                     /gene="TRIM27.2"
     mRNA            join(<61026..61445,62351..62446,62960..63190,63338..63360,
                     63505..63620,63779..63808,64170..>64684)
                     /gene="TRIM27.2"
                     /product="tripartite motif protein 27.2"
     CDS             join(61026..61445,62351..62446,62960..63190,63338..63360,
                     63505..63620,63779..63808,64170..64684)
                     /gene="TRIM27.2"
                     /codon_start=1
                     /product="tripartite motif protein 27.2"
                     /protein_id="ACA64758.1"
                     /db_xref="GI:169730357"
                     /translation="MAAPNPTAHLPSEASCPICLEYFRDPVSIHCGHNFCRDCITRSW
                     EWCTGDFCCPQCKETAAERILYPNRELARMLEIARRLSLQAAHRDAAGQEGCKKHREP
                     LSIYCKDDEAFICLICRESRLHRAHVMLPVQDAVQEYKEQIQSHLQALKEGRDKLLEF
                     REAELRRNWEYLEKTSTERQKVLGVFEGLRLFLEEQEHHLLAQLENTERDVEKTQEEN
                     VTILTKEISHLDAIIQEMEEKCQQPASNFLQDIRSTLSRLGKEIFQQPTLLLPDLERN
                     LNHFREKNNALEEILKNFKEILMFELPEKMSVTLDPSTAHPQLAVSADGSSVRWEEAQ
                     RDAANEEFGTDPFVLGHEGITSGRCCWEVEVTPKGSWAVGVAKESLKRREESGASPEV
                     ELWSMGFCEGQFWALSSFERMTLSQIQVPKRVRVTLDYESGQVAFFDADKRALIFIFP
                     SASFKGESIHPWFLVWGEGSQITLCP"
     gene            complement(<65730..>69573)
                     /gene="TRIM39.1"
     mRNA            complement(join(<65730..66250,67263..67366,67872..67894,
                     68587..68697,69544..>69573))
                     /gene="TRIM39.1"
                     /product="tripartite motif protein 39.1"
     CDS             complement(join(65730..66250,67263..67366,67872..67894,
                     68587..68697,69544..69573))
                     /gene="TRIM39.1"
                     /codon_start=1
                     /product="tripartite motif protein 39.1"
                     /protein_id="ACA64759.1"
                     /db_xref="GI:169730358"
                     /translation="MFPWELRARILEDLKNEIEKRQKSTARLSEEISQQETKSQGEKS
                     QHSDVKSPTTRYGKVTFQLPAGPEERLCHFSWRNGALKEHLKKLQASITLDPDTAHPD
                     LVLSEDRKSVKRGAGQQDLPDNPERFAYWPFVLGHQSFSSGRHSWEVEVGDEGDWAIG
                     VARESIPRKGQLSLCPKGGIWGVEKWGGQIRALTTHKVTLLALRWVPRRVSIHLDYAG
                     GTVAFFDADEGGIIFVFSHASFAGEGVRPWLWVVGVRSKLRLCP"
     gene            complement(<70729..>74368)
                     /gene="TRIM27.1"
     mRNA            complement(join(<70729..71324,71831..71860,72015..72130,
                     72319..72341,72544..72774,73321..73413,73970..>74368))
                     /gene="TRIM27.1"
                     /product="tripartite motif protein 27.1"
     CDS             complement(join(70729..71324,71831..71860,72015..72130,
                     72319..72341,72544..72774,73321..73413,73970..74368))
                     /gene="TRIM27.1"
                     /codon_start=1
                     /product="tripartite motif protein 27.1"
```

```
                     /protein_id="ACA64760.1"
                     /db_xref="GI:169730359"
                     /translation="MATQSPLDSLQSEASCSICLGYFQDPVSIPCGHNFCRDCIARCW
                     EGLEGNFPCPQCRKTTSHKNFRPSRELANIASIARQLNLRSSTWCKQHCEALKLFCKE
                     DQEPICAQCDASQAHRSHTVLPAHQAAQEYKEEIQARLQASKEEREKYLESRKTGGRM
                     SLYLEKTKKEGKRIVCEFEQLQRFLKEKERFLLVQLADLDRAITKVQEDAVTKVLEEM
                     SHLDTLIWEMEGKFQQPDSKFLLDIRRLLNSCEMMKFSPPAEIYPTLERRLEDFLQKN
                     VLVRCTLRKCQDSLMFQLQEPANITMDPTTAHPNLQLSEDRKQARGQLMPQELPENPE
                     RFNFEPCVLGCQGFTSGRHFWEVEVGQGGVWALGVARASMRRKGPMSFSPKEGVWALE
                     AYHSLTSPRANLRLNALPKRIRVSLDYEGSRVAFFSSDDDTPILVYTKASFNGEKVLP
                     WFKIGMGARLQEITQNSPSEEQSTSRQLMSHLDWVGFRSPLKICP"
     tRNA            complement(74924..74996)
                     /product="tRNA-Lys"
     tRNA            complement(75309..75381)
                     /product="tRNA-Val"
     gene            <76006..>80803
                     /gene="TRIM41"
     mRNA            join(<76006..76378,76913..77427,77847..77942,78451..78681,
                     79068..79090,79648..79775,80286..>80803)
                     /gene="TRIM41"
                     /product="tripartite motif protein 41"
     CDS             join(76006..76378,76913..77427,77847..77942,78451..78681,
                     79068..79090,79648..79775,80286..80803)
                     /gene="TRIM41"
                     /codon_start=1
                     /product="tripartite motif protein 41"
                     /protein_id="ACA64761.1"
                     /db_xref="GI:169730360"
                     /translation="MEAVGDDGASSGRLNPVETLQEEAICAICLDYFVEPVSIGCGHN
                     FCRVCIAQLWGGGGEAELEESGGAAALEEEEEEELEEEEEDDLGEEELDVEQEEEEED
                     GGGEEEEDDMWSEEEEDGELWEDPVEEELWDGGVQGELYFGDDDYDEDVMEEDVEEEE
                     EEDDEAQSPPPPVLPARPRRLQTFTCPQCRKTFFQRNFRPNLQLANMVQIIRQLHPHP
                     QRLAPPAGPSASGGPGGNPGILVATGGRGCPNLCEKHQEPLKLFCEVDEQAICVVCRE
                     SRSHKHHSVVPLEEVVQDYKYKLQSHLEPLKKKLDAVLKQKSNEQEKITELREKMKLE
                     IKEFESDFELLHQFLIGEHVLLLHQLEERYESLLARQSSNISQLEEQSAALSRLIAEA
                     EDKSKQDGLQLLKDIKGTFIRCENIKFQEPEMVLVDVGKKYRNYFLQDVVMRKMEKAF
                     SKVPQADITLDPDTAHPRLSLSLDRRSVKLAERRQELPDNPKRFDSDYCVLGSQGFTT
                     GRHYWEVEVGGRKGWAVGAARETARRKEKTMGPHQKREIWCVGTNGKKYQALTATEQM
                     ALSPSERPRRFGVYLDYERGQLCFYNAESMTHIHTFNASFHERIFPFFRILAKGTRIK
                     ICT"
     repeat_region   76398..76844
                     /mobile_element="LINE:CR1-Y3"
     repeat_region   76843..76920
                     /mobile_element="LINE:CR1-Y3"
     repeat_region   complement(77967..78048)
                     /mobile_element="LINE:CR1-F2"
     gene            complement(<82371..>85788)
                     /gene="GNB2L1"
     mRNA            complement(join(<82371..82436,82588..82698,83063..83203,
                     83413..83523,83656..83751,84037..84184,84967..85138,
                     85680..>85788))
                     /gene="GNB2L1"
                     /product="guanine nucleotide-binding protein beta subunit"
     CDS             complement(join(82371..82436,82588..82698,83063..83203,
                     83413..83523,83656..83751,84037..84184,84967..85138,
                     85680..85788))
                     /gene="GNB2L1"
                     /codon_start=1
                     /product="guanine nucleotide-binding protein beta subunit"
                     /protein_id="ACA64762.1"
                     /db_xref="GI:169730361"
                     /translation="MTEQMTLRGTLKGHNGWVTQIATTPQFPDMILSASRDKTIIMWK
                     LTRDETNYGIPQRALRGHSHFVSDVVISSDGQFALSGSWDGTLRLWDLTTGTTTRRFV
```

```
                     GHTKDVLSVAFSSDNRQIVSGSRDKTIKLWNTLGVCKYTVQDESHSEWVSCVRFSPNS
                     SNPIIVSCGWDKLVKVWNLANCKLKTNHIGHTGYLNTVTVSPDGSLCASGGKDGQAML
                     WDLNEGKHLYTLDGGDIINALCFSPNRYWLCAATGPSIKIWDLEGKIIVDELKQEVIS
                     TSSKAEPPQCTSLAWSADGQTLFAGYTDNLVRVWQVTIGTR"
     gene            <88914..>99310
                     /gene="BTN1"
     mRNA            join(<88914..88926,89013..89033,89329..89349,89432..89452,
                     89550..89570,89786..89806,89909..89929,90171..90191,
                     90439..90459,90585..90605,91059..91079,91358..91378,
                     91473..91493,91623..91643,91858..91878,92312..92332,
                     92475..92495,92719..92739,93516..93536,93708..93728,
                     93807..93827,94284..94304,94632..94652,94888..94908,
                     95400..95420,95588..95608,95862..95882,96405..96425,
                     96664..96687,96775..96795,96984..97004,97412..97432,
                     97565..97585,97699..97719,97922..97942,98403..98423,
                     98516..98536,98799..>99310)
                     /gene="BTN1"
                     /product="B-butyrophilin 1"
     CDS             join(88914..88926,89013..89033,89329..89349,89432..89452,
                     89550..89570,89786..89806,89909..89929,90171..90191,
                     90439..90459,90585..90605,91059..91079,91358..91378,
                     91473..91493,91623..91643,91858..91878,92312..92332,
                     92475..92495,92719..92739,93516..93536,93708..93728,
                     93807..93827,94284..94304,94632..94652,94888..94908,
                     95400..95420,95588..95608,95862..95882,96405..96425,
                     96664..96687,96775..96795,96984..97004,97412..97432,
                     97565..97585,97699..97719,97922..97942,98403..98423,
                     98516..98536,98799..99310)
                     /gene="BTN1"
                     /codon_start=1
                     /product="B-butyrophilin 1"
                     /protein_id="ACA64763.1"
                     /db_xref="GI:169730362"
                     /translation="MMASGKRNADLEEWDAKIRKLTEDFEESDTELEECDTEDGELMA
                     EIDNLTAELEERDRKIRKLTSDLGDVDSKLKERDRKITKLSAEIRRLTAVLGDRDSKL
                     RKLTAELAKCDATIRLFTVELGERHTKIGELTAEVGDYSRKLRKHAAELEERDLKIRE
                     QEAEIRRLTELLEDRDSDAREQDVLIRKLSEELEELRGQEVEESEPEREEHDAKIVEL
                     SAEVEKQERRIDELTAELEHYKAKARKCFEEHRRSEEKLVNVTLDPETAHPRLILSED
                     QKSVRWEYSLQESPDGPERFDADPCVLGCETFTSGRHCWVVDLTEGQYCAVGVSRESV
                     PRKGATSFNPDEGIWAVQQWGFKNRALTSPPTPLNLPRVPKKIRISLDYEWGEVAFFD
                     AENQIPIFTFPLTSFGGERLRPWFWVELGSLSLPR"
     gene            <102542..>105958
                     /gene="BTN2"
     mRNA            join(<102542..102892,103039..103134,103213..103443,
                     103553..103575,103925..104040,104311..104370,
                     104617..104681,105944..>105958)
                     /gene="BTN2"
                     /product="B-butyrophilin 2"
     CDS             join(102542..102892,103039..103134,103213..103443,
                     103553..103575,103925..104040,104311..104370,
                     104617..104681,105944..105958)
                     /gene="BTN2"
                     /codon_start=1
                     /product="B-butyrophilin 2"
                     /protein_id="ACA64764.1"
                     /db_xref="GI:169730363"
                     /translation="MASAASRAKEKGEEGSHHPIPEHRMDDGCGGGTRPERGRLGDPE
                     VISHRNHLKKGNFQTKLHLEHLAEKLKLLGLDGDRGEKEKLCLWHKRTFLSNGPRDDG
                     GPTGAPIKESVQEDGEQICCNLENLKKQKKELLELKMSGEKRCQDFLTRTEAERQKIV
                     SEFRQLRHFLKEKEVMLLAQLGELDRAVLRRQEEEEAKVEGDISLLSILICEMEEKLK
                     RPTREFLQDARSTLDRWDMGRTQRTMENFADVERRLRVISQQNKILKETLGRFQDLLP
                     SELEKEEGASVGEEGKAFVTLDPNTATAGLVLSRDRRGRLEP"
     tRNA            complement(105909..105990)
```

```
                     /product="tRNA-Leu"
     gene            complement(<107389..>111044)
                     /gene="BG1"
     mRNA            complement(join(<107389..107411,107563..107604,
                     107707..107853,107937..107957,108062..108082,
                     108317..108337,108451..108471,108625..108645,
                     108750..108770,108881..108901,109122..109142,
                     109446..109466,109781..109885,110249..110590,
                     110945..>111044))
                     /gene="BG1"
                     /product="BG-like antigen 1"
     CDS             complement(join(107389..107411,107563..107604,
                     107707..107853,107937..107957,108062..108082,
                     108317..108337,108451..108471,108625..108645,
                     108750..108770,108881..108901,109122..109142,
                     109446..109466,109781..109885,110249..110590,
                     110945..111044))
                     /gene="BG1"
                     /codon_start=1
                     /product="BG-like antigen 1"
                     /protein_id="ACA64765.1"
                     /db_xref="GI:169730364"
                     /translation="MCFLSGCNHPSFAFQWRTLLAHLVALHLLHLGSAQLTVVAPNHR
                     VTAIVGQDVVLRCHLSACKDARSSDIRWIQHRSSGFVHHYQNGEDLEQMEEYKGRTEL
                     LRDGLPDGNLDLCITAVSSSDSGSYSCAVQDGDGYAEAVVNLEVSDPFSEIVHPWKVA
                     LAVVITLLLGSFVIIAFLHRKQAARIRELKRKYAELERSDAKLEKLAGKLEQQTEAVE
                     RSDAKLETLAENLKKLAGKLEQQTEAVEKGNSQWSKPQSLNPRESEVIPKDYGRHEKS
                     LHTSFSTPTNPEKRHKAMERSHLKHPQYPNCRGYQNCDN"
     gene            complement(<112639..>117813)
                     /gene="BG2"
     mRNA            complement(join(<112639..112664,112758..112778,
                     112871..112891,112994..113014,113122..113142,
                     113243..113263,113387..113407,113510..113530,
                     113641..113661,113763..113783,113878..113898,
                     113996..114016,114113..114133,114229..114249,
                     114608..114628,114730..114750,114861..114881,
                     115002..115022,115129..115149,115256..115276,
                     115378..115398,115509..115529,115632..115652,
                     115764..115784,116152..116172,116491..116598,
                     117018..117359,117714..>117813))
                     /gene="BG2"
                     /product="BG-like antigen 2"
     CDS             complement(join(112639..112664,112758..112778,
                     112871..112891,112994..113014,113122..113142,
                     113243..113263,113387..113407,113510..113530,
                     113641..113661,113763..113783,113878..113898,
                     113996..114016,114113..114133,114229..114249,
                     114608..114628,114730..114750,114861..114881,
                     115002..115022,115129..115149,115256..115276,
                     115378..115398,115509..115529,115632..115652,
                     115764..115784,116152..116172,116491..116598,
                     117018..117359,117714..117813))
                     /gene="BG2"
                     /codon_start=1
                     /product="BG-like antigen 2"
                     /protein_id="ACA64766.1"
                     /db_xref="GI:169730365"
                     /translation="MNLLLGCNHPSFTLPWRTLLAHLVALHLLHLGSAQLRVVAPSLP
                     VTAVVGQDVVLRCHLCPCKNAWNSDIRWVQHRSSGLVHHYQNGVDLEQMEEYKGRTEL
                     LRNGLSDGNLDLRITAVSSSDSGSYSCAVQDGDGYAEAVVNLEVSDPFSEIVHPWKVA
                     LAVVITILLGSFAVIIVFLHRKQVALSRELKRTDAKLAEQAAVLNFTAGSLKKQAAEL
                     VEQTKEVEKKNLVLKKSSENIDSTARDLKKQAAELEKLKEKMLKQTNKLEEQTEKLVK
                     ENEEVEKQNSELKKNCEKMDFTAGSLKILAAKMNFTAGSLMKQTEAVQNQNSELKKDC
```

```
                EEMGEQTEALVVETEESGKPSEEKD"
gene            complement(<119380..>123213)
                /gene="BG3"
mRNA            complement(join(<119380..119387,119484..119504,
                119608..119628,119739..119759,119861..119881,
                119977..119997,120238..120258,120363..120383,
                120486..120506,120614..120634,120731..120751,
                120856..120876,120986..121006,121135..121155,
                121266..121286,121615..121635,121950..122054,
                122419..122760,123114..>123213))
                /gene="BG3"
                /product="BG-like antigen 3"
CDS             complement(join(119380..119387,119484..119504,
                119608..119628,119739..119759,119861..119881,
                119977..119997,120238..120258,120363..120383,
                120486..120506,120614..120634,120731..120751,
                120856..120876,120986..121006,121135..121155,
                121266..121286,121615..121635,121950..122054,
                122419..122760,123114..123213))
                /gene="BG3"
                /codon_start=1
                /product="BG-like antigen 3"
                /protein_id="ACA64767.1"
                /db_xref="GI:169730366"
                /translation="MRFSLGCNHPSFAFQWRTLLAHLVALHLLHLGSAQLRVVAPSLP
                VTAIVGQDVVLRCHLCPCKDAWSSDIRWIQHRSSGLVHHYQNGEDLEQMEEYKGRTEL
                LRNGLPDGNLDLRITAVSSSDSGSYICAVQDGDGYAEEMVNLEVSDPFSEIVHPWKVA
                LAVVITLLLGSFVIIAFLHRKQAAQIRELKSKDAKLDEQDDILYSSAAGLKKLAEKLV
                EQIQVVEKKQSDMQKIYKNTDFGVADLKPLAAELEKHCEKLEKHCKKLERGGAKLEKQ
                DAELGERTEELEE"
repeat_region   complement(124145..124185)
                /mobile_element="LINE:CR1-C2"
repeat_region   complement(125035..125119)
                /mobile_element="LINE:CR1-F"
gene            complement(<125503..>128012)
                /gene="Blec2"
mRNA            complement(join(<125503..125629,125703..125812,
                126341..126495,126572..126706,127098..127193,
                127940..>128012))
                /gene="Blec2"
                /product="lectin-like natural killer cell surface protein"
CDS             complement(join(125503..125629,125703..125812,
                126341..126495,126572..126706,127098..127193,
                127940..128012))
                /gene="Blec2"
                /codon_start=1
                /product="lectin-like natural killer cell surface protein"
                /protein_id="ACA64768.1"
                /db_xref="GI:169730367"
                /translation="MDEEIIYADLRQPTGSLSPAQQQHVWAFPVPWDFQKVGLLGVLL
                LLLLLLVGLSVWVFQKTTPPPSIAQHCPETNGTKGMELSGSSSIERYFCPSTPKSSAA
                PAACLLCPQFWRLLGDRCYRLSTGKGIWIQAKRECESLQSQLAVLQKKAEKDFLKEIA
                GAVNQPLWIGLEVSKYQWKWVDNSSFNVAEFGNLSLVENRCGTFKNTKVEDDDCSGEH
                EWVCQKEPLHLHP"
gene            <129859..>131819
                /gene="BLec1"
mRNA            join(<129859..129916,130298..130384,130818..130990,
                131311..131417,131678..>131819)
                /gene="BLec1"
                /product="c-type lectin-like receptor"
CDS             join(129859..129916,130298..130384,130818..130990,
                131311..131417,131678..131819)
                /gene="BLec1"
```

```
                     /codon_start=1
                     /product="c-type lectin-like receptor"
                     /protein_id="ACA64769.1"
                     /db_xref="GI:169730368"
                     /translation="MEEMREYPNFLGTERSRREGPSHGACITFQLTMAAVFTVLLITA
                     VAFAVQAFQPHPQPCAQCPFDWVGFRGKCYYFSEDESNWTSSQNNCSALGASLAVFDS
                     TEDLSFTMRHKGSSPHWVGLSREGEEHPWEWMSRSPSSHPFQVQGDGLCAYLGDLGLT
                     SSHCSAHRNWVCTKPALQNPRKNFCIST"
     mRNA            complement(join(<132673..132686,132772..132795,
                     132868..132978,133061..133342,133429..133698,
                     133909..>133999))
                     /product="MHC class II antigen beta chain 1"
     CDS             complement(join(132673..132686,132772..132795,
                     132868..132978,133061..133342,133429..133698,
                     133909..133999))
                     /codon_start=1
                     /product="MHC class II antigen beta chain 1"
                     /protein_id="ACA64770.1"
                     /db_xref="GI:169730369"
                     /translation="MGSGRVPAAGAVLVALLALGARPAAGTRPSAFFLYGTTGDCHFL
                     NGTERVWYVDRYFYNRQQFTHFDSDVGKFVADTPLGEPTAEYWNSNTEFMEIKRNEVD
                     TFCRHNYGVFESFTVQRSVEPKVRVSALQSGSLPETDRLACYVTGFYPPEIEVKWFLN
                     GREETERVVSTDVMQNGDWTYQVLVVLETVPRRGDSYVCRVEHASLRQPVSQPWEPPA
                     DAGRSKLLTGVGGFVLGLVFLALGLVVFLRGQKGRPVAAAPGMLN"
     repeat_region   134476..134702
                     /mobile_element="LINE:CR1-C4"
     gene            <135154..>139324
                     /gene="TAPBP"
     mRNA            join(<135154..135184,135268..135477,136049..136288,
                     137130..137474,138285..138623,138912..139001,
                     139201..139226,139313..>139324)
                     /gene="TAPBP"
                     /product="tapasin"
     CDS             join(135154..135184,135268..135477,136049..136288,
                     137130..137474,138285..138623,138912..139001,
                     139201..139226,139313..139324)
                     /gene="TAPBP"
                     /codon_start=1
                     /product="tapasin"
                     /protein_id="ACA64771.1"
                     /db_xref="GI:169730370"
                     /translation="MAAGLRLLLAGFFWSLFRVEVTASPPPPPAPVRCALLEGEGRGR
                     GLPGGGKARPALLRFGGDAENPPEPGSEPEVTFSVSDPWGALTPLGVPPRTPPSCELN
                     PTNPQTVSAPWSRPLHPDARSPPTAGGQWWVAAVGTPQYSATALLQGGTGTEGTITAA
                     VSLAVLTNTPTLRVRVGSPTLLHCAFAAPLSPFILEWRHQSRGAGRILLAYDSSTARA
                     SHTTPGAELLLGTRDGDGVTEVTLRLARPSPGDEGTYICSVFLPHGHTQTVLQLDVFE
                     PPKVTLSPKNLVVTPGTSAELQCHVSGFYPLDVTVTWQRRAGGSGPSRSPGGTVMDSW
                     TSGHRQAADGTYSRTAAARLIPARPQHHGDVYSCVVTHTALAKPMRVSVRLLLAGTEG
                     PHLEDITGLFLVAFILCGLIRWLYPTAARPEEETKKSQ"
     mRNA            join(<140312..140402,140613..140882,140969..141250,
                     141340..141450,141523..141546,141621..>141634)
                     /product="MHC class II antigen beta chain 2"
     CDS             join(140312..140402,140613..140882,140969..141250,
                     141340..141450,141523..141546,141621..141634)
                     /codon_start=1
                     /product="MHC class II antigen beta chain 2"
                     /protein_id="ACA64772.1"
                     /db_xref="GI:169730371"
                     /translation="MGSGRVPAAGAVLVALLALGARPAAGTRPSALFQFSASSECHFL
                     NGTERVRFVQRYIYNRQQYAHFDSDMGKFVADTPLGEPSAENWNSNTEILEDSQGAVD
                     TYCRHNYGVGESFTVQRSVEPKVRVSALQSGSLPETDRLACYVTGFYPPEIEVKWFLN
                     GREETERVVSTDVMQNGDWTYQVLVVLETVPRRGDSYVCRVEHASLRQPVSQPWEPPA
                     DAGRSKLLTGVGGFVLGLVFLALGLVVFLRGQKGRPVAAAPGMLN"
```

```
     mRNA            join(<144442..144532,144744..145013,145100..145381,
                     145470..145580,145654..145677,145811..>145824)
                     /product="MHC class II antigen beta chain 3"
     CDS             join(144442..144532,144744..145013,145100..145381,
                     145470..145580,145654..145677,145811..145824)
                     /codon_start=1
                     /product="MHC class II antigen beta chain 3"
                     /protein_id="ACA64773.1"
                     /db_xref="GI:169730372"
                     /translation="MGSGRVPTAGAVLVALLALGARPAAGTRPSALFQFSASSECHFL
                     NGTERVRLVQRYIYNRQQFTHFDSDVGKFVADTPLGEPSAEYLNSNTEFIEYLRGAVD
                     TACRHNYGVIESFTVQRSVEPKVRVSALQSGSLPETDRLACYVTGFYPPEIEVKWFLN
                     GREETERVVSTDVMQNGDWTYQVLVVLETVPRRGDSYVCRVEHASLRQPVSQPWEPPA
                     DAGRSKLLTGVGGFVLGLVFLALGLVVFLRGQKGRPVAAAPGMLN"
     gene            complement(<147230..>152871)
                     /gene="BRD2"
     mRNA            complement(join(<147230..147360,147434..147556,
                     147641..147951,148041..148240,148319..148573,
                     148650..148778,148872..149243,149332..149540,
                     149622..149763,149841..149978,150813..151113,
                     152843..>152871))
                     /gene="BRD2"
                     /product="bromodomain-containing protein 2"
     CDS             complement(join(147230..147360,147434..147556,
                     147641..147951,148041..148240,148319..148573,
                     148650..148778,148872..149243,149332..149540,
                     149622..149763,149841..149978,150813..151113,
                     152843..152871))
                     /gene="BRD2"
                     /codon_start=1
                     /product="bromodomain-containing protein 2"
                     /protein_id="ACA64774.1"
                     /db_xref="GI:169730373"
                     /translation="MLQNVNPQSKILGEGNAGLMGLATESTPGKRIRKPSLLYEGFES
                     PTMASVPALQTPQANPPPPEVSNPKKPGRVTNQLQYLHKVVMKALWKHQFAWPFRQPV
                     DAVKLGLPDYHKIIKQPMDMGTIKRRLENNYYWGAAECMQDFNTMFTNCYIYNKPTDD
                     IVLMAQTLEKIFLQKVAQMPPEEQEIVVPVAKNSHKKGASRAAALLAGLTAAQQVPAV
                     SSVSHTAVYTPSPDIATTIVNIPHPSVISAPLLKSLHSAAPAVLAAPAPTQPVAKKKG
                     VKRKADTTTPTTTAIIATSGESSPSATLLEAKAAKIPARRESGRPIKPPKKDLPDSQQ
                     HQTSKKGKLSEQLKYCNGILKELLSKKHAAYAWPFYKPVDASALGLHDYHEIIKHPMD
                     LSTIKRKMENRDYHDAQEFAADVRLMFSNCYKYNPPDHDVVAMARKLQDVFEFSYAKM
                     PDEPQDASPPSVSAPLLGALSKSSSEESSSDEDDEDEDDEDDDEDESSSESSSESEES
                     SDSEEERANRLAELQEQLRAVHEQLAALSQGPVSKPKKKREKKKKKSEKHKGRGGDE
                     ESRARQAQLRKAKKGGGGGSSGGSSKNSKKAAKATLPPPPALYDSEEEEESKPMTYDE
                     KRQLSLDINKLPGEKLGRVVHIIQSREPSLRDSNPEEIEIDFETLKPSTLRELERYVL
                     SCLRKKPRKPYSETMKKPVGKTKEELALEKKRELEKRLQDVSGQLNSAKKPPKKANEK
                     PESAQQVAVSRLSASSSSSDSSSSSSSSSSSDTSDSDSG"
     gene            <156695..>159022
                     /gene="DMA"
     mRNA            join(<156695..156767,157414..157680,158491..158772,
                     158853..>159022)
                     /gene="DMA"
                     /product="MHC class II antigen M alpha chain"
     CDS             join(156695..156767,157414..157680,158491..158772,
                     158853..159022)
                     /gene="DMA"
                     /codon_start=1
                     /product="MHC class II antigen M alpha chain"
                     /protein_id="ACA64775.1"
                     /db_xref="GI:169730374"
                     /translation="MGAASGRGALVAALLGAALGSVRAEPSLHTLSEVLFCQPDTPLL
                     GLSVAFDLEQLFSFDVPNSQWLPQLPDGPAWPAGTEQPQELQHDTILCRKLHDELTSL
                     ATGPRPMPEAKGIPVADVFLQQPLHLGYPNTLVCMVGNIFPPAITISWQRDGVPITDG
```

```
                     VTHLTYIPMEDLGFMRFSYLAVTPHSGDIYSCIVTRERDNMSVVAYWVPQDPIPSDVL
                     AMAVCGAVTALGILLALLGLGLLLSARRHSTWGQRGQWGHSPRTH"
     gene            <159246..>161497
                     /gene="DMB1"
     mRNA            join(<159246..159330,159819..160100,160721..160996,
                     161068..161181,161268..>161497)
                     /gene="DMB1"
                     /product="MHC class II antigen M beta chain 1"
     CDS             join(159246..159330,159819..160100,160721..160996,
                     161068..161181,161268..161497)
                     /gene="DMB1"
                     /codon_start=1
                     /product="MHC class II antigen M beta chain 1"
                     /protein_id="ACA64776.1"
                     /db_xref="GI:169730375"
                     /translation="MKAAMLPMGPTGFRVLGAMGLALSCGTAGAFVVHMASSCLLTAN
                     GSLGGFDLIVAFNKNPLVCYDPDAYIFYPCDGWLLHDVATLVAIFLNNNATWVQRAEA
                     RRQACTELAAQFWAHTALRRTPPQVRIVPIPISNDPDSVRLICHVWGFFPPAVTIQWL
                     QNGLVVASEDTKLLPNGDWTYRTQTTLRASTAAGSTYTCSVWHSSLEQPLQVDWSPNL
                     SLGMMAQVAVAVMALTLGLVAFSAGVFSFCQRPRGPGAGPSPHPDAGSDSNPGPQSDP
                     RPSSGPHPSPGPHSGSDPHPGPGPHPSPDPHSVSDPHPGPGPGPGFASGPYP"
     gene            <161954..>164765
                     /gene="DMB2"
     mRNA            join(<161954..161996,162777..163058,163850..164134,
                     164209..164328,164637..164672,164755..>164765)
                     /gene="DMB2"
                     /product="MHC class II antigen M beta chain 2"
     CDS             join(161954..161996,162777..163058,163850..164134,
                     164209..164328,164637..164672,164755..164765)
                     /gene="DMB2"
                     /codon_start=1
                     /product="MHC class II antigen M beta chain 2"
                     /protein_id="ACA64777.1"
                     /db_xref="GI:169730376"
                     /translation="MLVLLGLLLGARGAGAFMVHMASSCPLAANGSLLGFDLTVAFNK
                     NPLVCYDPDVRRFNACDWRLLHDVAEWIASSLNNDSTWVQRAEARRRTCSKLAAQFWA
                     QTALRRTQPQVRIIPAQTGNPSVPIRLTCHVWGFYPPEVTIIWLHNGDIVGPGDHSPM
                     SAIPNGNWTYQTQVALSVAPEVGDTYTCSVQHASLEEPLLEDWSPGLTLEVTLMVAVA
                     TVAMVLGLGLLLIGVFCWRAQPPAPGYTPLPGHNYPSGSI"
     repeat_region   complement(164343..164552)
                     /rpt_type=tandem
                     /rpt_unit_seq="acccataggg"
                     /satellite="microsatellite"
     repeat_region   complement(164459..164581)
                     /rpt_type=tandem
                     /rpt_unit_seq="acccatagag"
                     /satellite="microsatellite"
     repeat_region   164990..165096
                     /mobile_element="LINE:CR1-X1"
     mRNA            join(<165569..165635,165745..166014,166249..166524,
                     166614..166886,166959..167066,167190..167222,
                     167381..167413,167564..>167586)
                     /product="MHC class I antigen alpha chain 1"
     CDS             join(165569..165635,165745..166014,166249..166524,
                     166614..166886,166959..167066,167190..167222,
                     167381..167413,167564..167586)
                     /codon_start=1
                     /product="MHC class I antigen alpha chain 1"
                     /protein_id="ACA64778.1"
                     /db_xref="GI:169730377"
                     /translation="MGSWGILGWGLLLLRGVCGAAGEPHSLRYFHTAMTDPGSGMPWF
                     VDVGYVDGEIFVHYNSTARKVVPRTEWMKAPGAVDPEYWERSTRIAQSNERVYRVSLG
                     TLQERFNHSGGSHTVQRKIGCDILEDGTTRGYHQVAYDGRDFIAFDKDTKTFTATVPE
```

```
                     AVLTKRRWEDDGSVAERFKHYLEETCVEWLRTYMERGKAELERREPPEVRVWGKKADG
                     ILTLSCRAHGFYPRPIDVSWMKDGVAQNQDTQSGGIAPNSDGTYHTLVTIDARPEDWD
                     KYQCRVEHASLQQPGFYSWEPSQPNLVPIVVGVIVAIVAIAIVGGVGFIIYRSHAGKK
                     GKGYNIAPAQDGGSNSSGTGSDRPVI"
     gene            complement(<168635..>172948)
                     /gene="TAP1"
     mRNA            complement(join(<168635..168874,169012..169148,
                     169220..169379,169631..169804,170146..170331,
                     170424..170552,171014..171211,171656..171861,
                     171933..172063,172137..172239,172849..>172948))
                     /gene="TAP1"
                     /product="transporter associated with antigen processing
                     1"
     CDS             complement(join(168635..168874,169012..169148,
                     169220..169379,169631..169804,170146..170331,
                     170424..170552,171014..171211,171656..171861,
                     171933..172063,172137..172239,172849..172948))
                     /gene="TAP1"
                     /codon_start=1
                     /product="transporter associated with antigen processing
                     1"
                     /protein_id="ACA64779.1"
                     /db_xref="GI:169730378"
                     /translation="MGAMGAGRRLLLSLSPERRRCAAVIGLMAASALGEMAVPYYTGR
                     ASDWVAREDELAAILPMVLLGLSSAVTELVCDVTFVGTLSRTQSRLQRRVFAAVLRQD
                     IAELRADGAGDVATRVTRDAEDVREALGEALSLLLWYLARGVCLFATMAWLSPRMALL
                     TALALPLLLALPRAVGHFQQALAPQMQKAQARASEVAVETFQAMATVRSFANEDGAAA
                     HYRQRLQQSHRLEKKDVALYTVSAWTSGFSALALKVGILYYGGQLVAAGTVSTGDLVT
                     FLLYQMQFTDVVEVLLRYYPTLTKAVGSSEKIFEFLDREPQVAPSGTMAPSELQGHLQ
                     LEDVWFSYPGRQEPVLKGVSLELRPGEVLALLGPPGAGKSTLVALVSRLHQPTAGRLL
                     LDGHPLPAYQHSYLCHQVAVVPQEPLLFARSLHANISYGLGCCSRAQVTAAARKVGAH
                     NFITCLPQGYDTEVGELGGQLSGGQRQAVAIARALLRDPRILILDEHSSALDTECQQQ
                     VEQEILAARGSGRAVLMVTGRAALAARAQRVAVLEGGEVRQQGPPHEVLRPGSRTWQL
                     LQDWGQQGAPEEGDRGSGGEG"
     gene            <173510..>176702
                     /gene="TAP2"
     mRNA            join(<173510..174105,174501..174631,174701..174906,
                     175117..175314,175389..175517,175653..176015,
                     176084..176243,176310..176446,176523..>176702)
                     /gene="TAP2"
                     /product="transporter associated with antigen processing
                     2"
     CDS             join(173510..174105,174501..174631,174701..174906,
                     175117..175314,175389..175517,175653..176015,
                     176084..176243,176310..176446,176523..176702)
                     /gene="TAP2"
                     /codon_start=1
                     /product="transporter associated with antigen processing
                     2"
                     /protein_id="ACA64780.1"
                     /db_xref="GI:169730379"
                     /translation="MQPHILRLSCTLLLADLALMLALAHFFPALAQLGRVGSWLEAGL
                     RLLVLGGAGRLLVPRGPHGAAVLLSLGPAVFLTLRGYVGLPGAAPVLLATATPSWLVL
                     THGTVVVALLTWSLLVPTVAVGAKEAEAWVPLRRLLALAWPEWPFLGCAFLFLVLAAL
                     GETSVPYCTGRALDVLRQGDGLAAFTAAVGLMCLASASSSLFAGCRGGLFTFIRFRFV
                     LRTRDQLFSSLVYRDLAFFQKTTAAELASRLTTDVMLASNVLALNINVMLRNLGQVLG
                     LCAFMLGLSPRLTMLALLEVPLAVAARKVYDTRHQMLQRAVLDAAADTGAAVQESISS
                     IETVRVFNGEEEEEHRYSQALDKTLQLRDQRDTEKAIFLLIRRALQLAVQALVLYCGH
                     QQLREGTLTAGSLIAFILYQTKAGSCVQALAYSYGDLLSNAAAACKVFDYLDQERAVG
                     TGGTYVPTKLRGHVTFHQVSFAYPTRPERLVLQDVTFELRPGEVTALAGLNGSGKSTC
                     VALLERFYEPGAGEVLLDGVPLRDYEHRYLHRQVALVGQEPVLFSGSIRDNIAYGMED
                     CKEEEIIAAARAAGAFGFISALEQGFSTDVGERGGQLSAGQKQRIAIARALVRRPTVL
                     ILDEATSALDGDSDAALQQWVRNGGDRTVLLITHQPRMLEKADRVVVLEHGTVAEMGT
```

```
                     PNELRTLGGPYSRLLQR"
     mRNA            complement(join(<177983..178005,178156..178188,
                     178348..178380,178504..178611,178684..178956,
                     179045..179320,179555..179824,179934..>180000))
                     /product="MHC class I antigen alpha chain 2"
     CDS             complement(join(177983..178005,178156..178188,
                     178348..178380,178504..178611,178684..178956,
                     179045..179320,179555..179824,179934..180000))
                     /codon_start=1
                     /product="MHC class I antigen alpha chain 2"
                     /protein_id="ACA64781.1"
                     /db_xref="GI:169730380"
                     /translation="MGSCGALGWGLLLLGAVCGAAGEPHSLRYFDTAMTDPGTGLPWF
                     VSVGYVDGEIFVHYDSTARRYVPHTEWVKAPGAVDPEYWERCTQILQRTEQVYRVSLG
                     ILQERFNQSGGSHTVQRMYGCDILGNSTTRGYNQYAYDGRDFIAFDKDKKTFTAAVPE
                     AVLTKRKWEADPTYVDQRKHYLEETCVQWLRRHMEHGKAELERREPPEVRVWGKKTDG
                     ILTLSCRAHGFYPRPIDVSWMKDGVAQNQDTQSGGITPNSDGTYHTLVTIDARPEDGD
                     KYQCRVEHASLQQPGLYSWEPSQPNLVPIVVGVIVAIVAIAIVGGVGFIIYRSHAGKK
                     EKGYNIAPAQDGGSTSSGTGSDRPVI"
     gene            <180995..>196378
                     /gene="C4"
     mRNA            join(<180995..181228,181483..181660,181968..182038,
                     182124..182212,182752..182834,182902..182998,
                     183378..183486,183641..183764,184073..184185,
                     184326..184496,184639..184824,185235..185420,
                     185666..185821,186082..186208,186360..186431,
                     186753..186947,187089..187203,187290..187360,
                     187501..187640,188102..188302,188597..188648,
                     188724..188813,189285..189482,189744..189819,
                     189982..190138,190248..190364,191553..191718,
                     192736..192902,193010..193180,193380..193439,
                     193549..193642,193732..193957,194043..194133,
                     194448..194522,194659..194761,194865..194954,
                     195153..195242,195564..195647,195966..196098,
                     196237..>196378)
                     /gene="C4"
                     /product="complement C4"
     CDS             join(180995..181228,181483..181660,181968..182038,
                     182124..182212,182752..182834,182902..182998,
                     183378..183486,183641..183764,184073..184185,
                     184326..184496,184639..184824,185235..185420,
                     185666..185821,186082..186208,186360..186431,
                     186753..186947,187089..187203,187290..187360,
                     187501..187640,188102..188302,188597..188648,
                     188724..188813,189285..189482,189744..189819,
                     189982..190138,190248..190364,191553..191718,
                     192736..192902,193010..193180,193380..193439,
                     193549..193642,193732..193957,194043..194133,
                     194448..194522,194659..194761,194865..194954,
                     195153..195242,195564..195647,195966..196098,
                     196237..196378)
                     /gene="C4"
                     /codon_start=1
                     /product="complement C4"
                     /protein_id="ACA64782.1"
                     /db_xref="GI:169730381"
                     /translation="MEQLWGWWLKQCSLRLVLVAPRRVALGTPMGLLVAAVGPVTGTV
                     TAWAEGDRGAGPCALPVPFALTPHNNFNQLLHIEVTPVQAEHCGALWGRGLLLEAHSP
                     HLPPPSTRSLHVALGGPRGHLIVQTDKPLYAPRQTVRFRVFSMDPDLQPNPEPILVTI
                     MNPMGARVREVQRVPLNTVLSDQLVLPDIALPGTWHIRAQLAASPNTNGSTAFEVRKY
                     VLPGFEVRIRPERGFVVLSDPEPAPLRIHLHVQFPDGAPVWGRAQLRVGLREAGGRGG
                     RFLRGLEQQCQVTEGHASLEVSPVGVAKAVGVAMADLQGALLRLAVGVVESAGGELVE
                     RELLVPLVRSQWVLQLQRSARFFVPGAPYTLLGRVLQAGGGTAPGVPVRVTVGVTGAP
```

```
                     APPLKELRADNSGDITVPINVPKGATRMELSVEARQPGVPPARAQLLVTPMAAVSGRF
                     LVLGGPRGSLRPGEELRLQLHHMGPPPAPQRFHLLAMARGRVVVAHTVASQVTLTEVA
                     LPVTADMAPYLHVVAFFLSGGNVVAASWGGAVRGGCDEQVQVKVPPRGTTLRPQDPLT
                     VTVTTETPVTVAIGAIDTAVLSLEPRHRLNAAKVEAALGSSDLGCSPGGGPDAVGIFG
                     AAGLVLGLRGGSPTVQAAPHCPPAPTRQRRSLELLKLLEEKAGPWRDNAVMWRCCRDG
                     ATALPIRVTCQQRGQRVTTAGGCRDAFLQCCEVAQNLRRRGQRGGLARVLEAQLAEQL
                     LDDDEDVPTRSFFPESWLWQRVRVVGTARLSVLLPDSITTWEIQAVAIVPGHGLCVAE
                     PQRVTVTQDVRVALRLPPSIRPREQLQLQPLIHSRLPRSINVTVTLSAVEGLCAALDG
                     VPQMLELPPGRAVAAPLTLVALHPGDIPITVTARGPWGLGDRVTRVLHVEPEGELHLE
                     ETTYILDTDDKRSRSLKLPGDVPAEIVPDGDFSMSVRVGGRVPGWALQGALGVGDSLL
                     RSPRGCGEQSLMSMAPTAAALRFLDESEGWGQLPPGHRQRGLRTLQQGFERVQSFRKS
                     DGSYGAWLHRDSSTWLTALVLRVLALSRPYLPVAASGPAESLRWVLEQQRSDGAFLEH
                     RAVVHREMQGGVADPGPEATVLLTAFVVVALHGARTLLPPDSPELPLLDKSLSRASTF
                     LRGRVEQLGTYGTAITSYALALVDTTPPGPHPAVERLRGMARSAHGGRATFWPSGGPA
                     ATVEATGYALLALLQSRDIAGAGRAARWLRQQSNYGGGFHSTQDTVVALEALAQMWLQ
                     WGRGSTSGLNLGLSWPGGARGRAGGTRMTLKPGLEPLEQELKVPLGSPVTVQVEGHGE
                     GTLTVLRQFRLLSPPNATCQALHLEVAITGPILYHDEDYEDYENYEEAEPKEGEEPTE
                     GEMPVEGAGPVEGEGPADGPAPLSPMSLWDARKRRRRSTHSPAHEVAFLVCFRRSPGV
                     ALTGMAVVEISLLSGFSPHRTDLDKLRDLVDHWISHYELDGNQLVLYLDEVPPERQCL
                     SFGATQDVAVGHMQPAMAAIYDYYEPGQRCTIFYNAPQRSSTIATLCSPKICECAQGG
                     CPHAQRRTQEVTADDRHDFACYSPRVDYALVVRVLSQSEIGAFVAFETEIKEVLLGGQ
                     DTTAAPGERRRLLVRKSCPLRLQPHDIYLVMGGSGGTQDPEGRPQFLLGPHSWVEEVP
                     SPGRCRATRLRSHCAQLQEFRTRLSQLGCQL"
     gene            <196827..>197022
                     /gene="CenpA"
     mRNA            join(<196827..196911,196997..>197022)
                     /gene="CenpA"
                     /product="centromere protein A"
     CDS             join(196827..196911,196997..>197022)
                     /gene="CenpA"
                     /codon_start=1
                     /product="centromere protein A"
                     /protein_id="ACA64783.1"
                     /db_xref="GI:169730382"
                     /translation="MPRPKPRSPRRRGRPPPAAPPPPPARPRARRYRPGQR"
```

*Complete sequence is available from Genbank*

**Figure 4:** Phylogenetic analysis of the major avian MHC genes (A-Class I, B-Class IIβ, C-BG loci). Clustal analysis for BG genes was performed on the IG to transmembrane domains to obtain alignments.

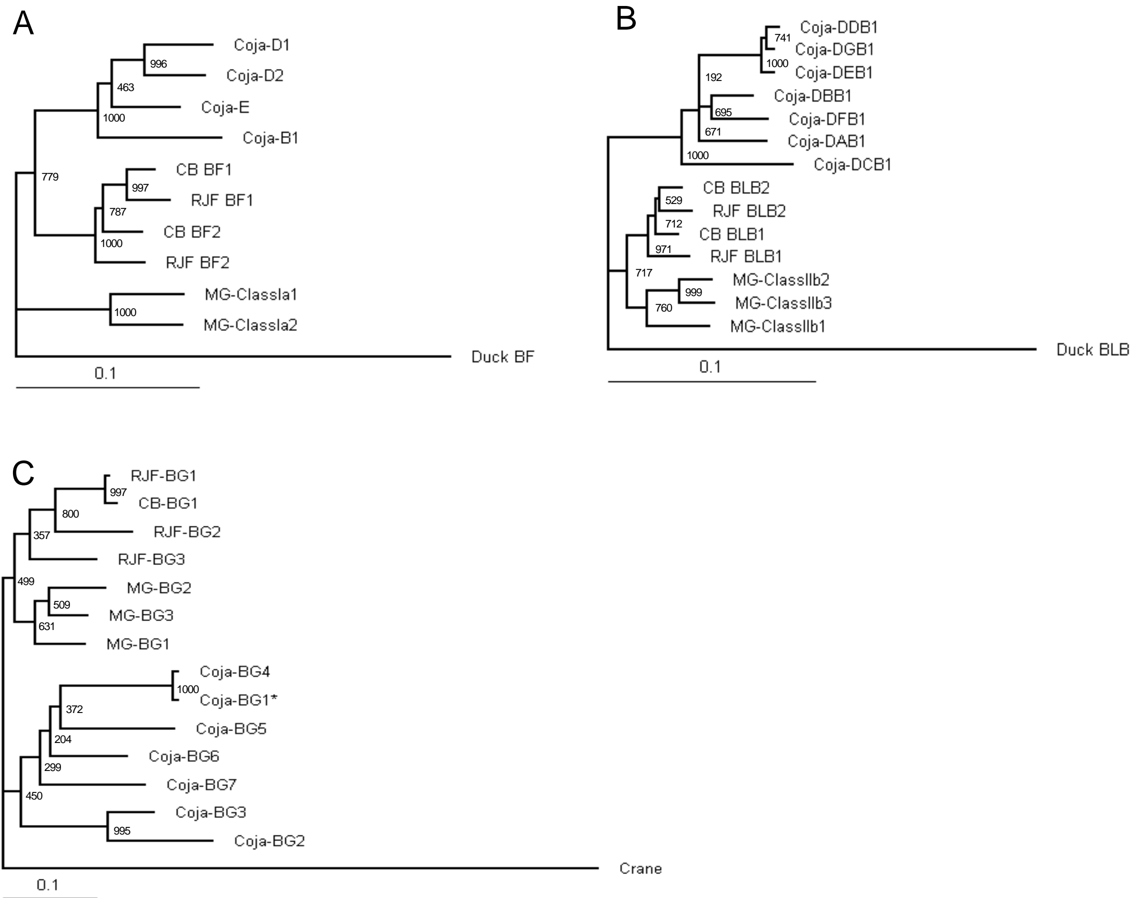


**Figure 5:** Phylogenetic analysis of Class IIβ β1 sequence from turkey, chicken and quail.

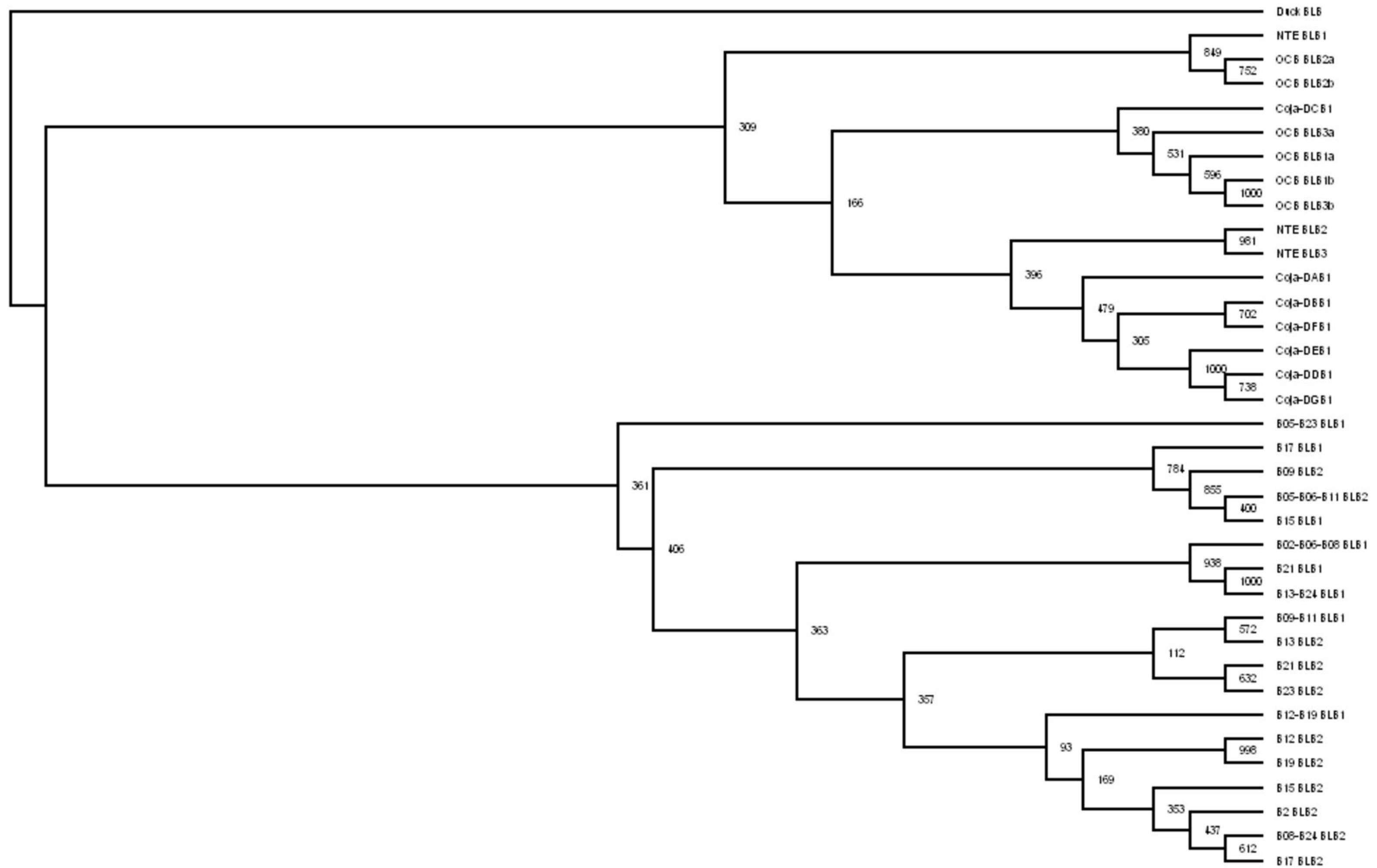


**Figure 6:** Semi-quantitative PCR analysis of turkey MHC Class I loci identified in the OCB02 bird. qPCR suggest only modest variation in expression levels between the two Class I loci.

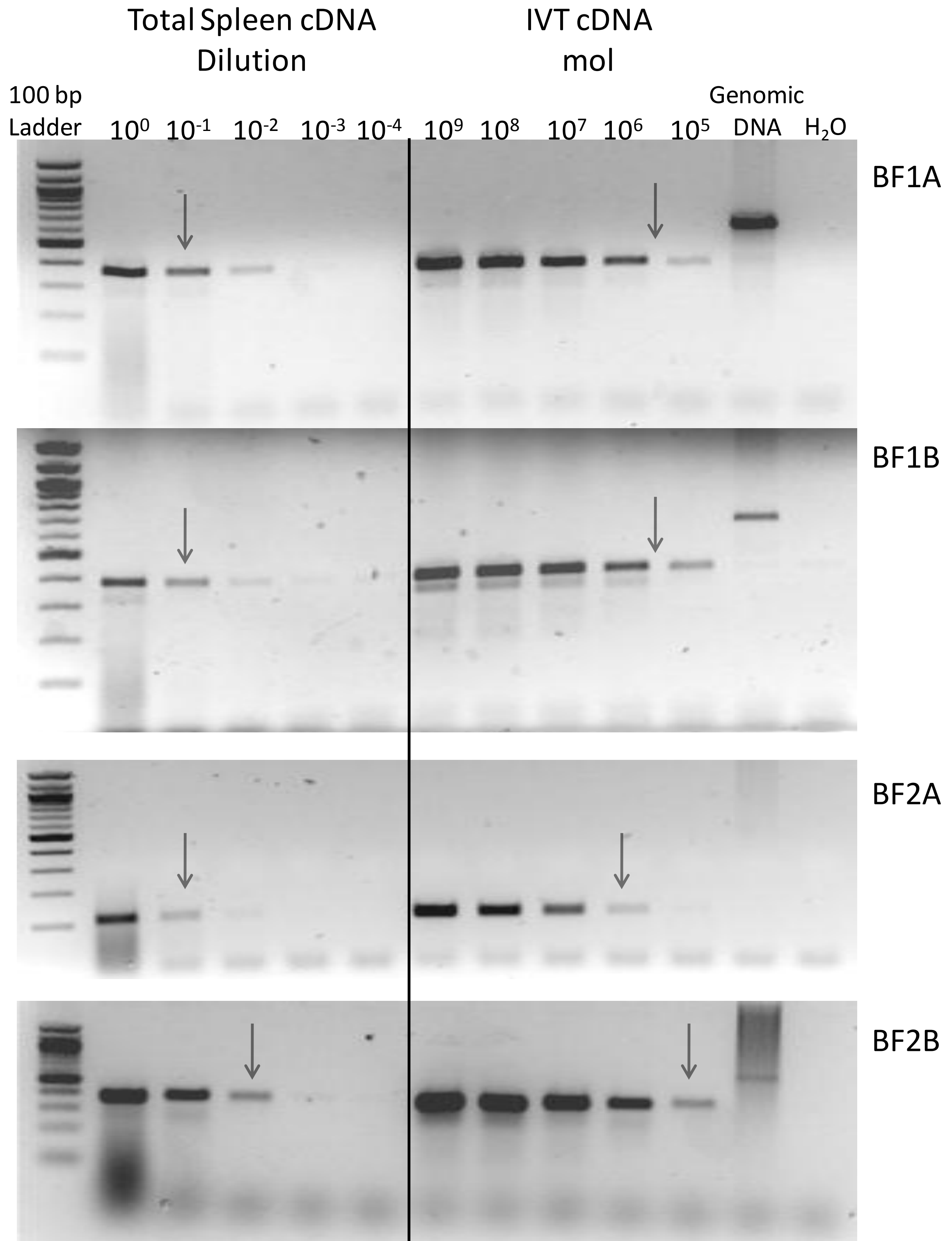


**Figure 7:** Class I gene promoters. Turkey MHC Class I promoters compared to chicken B21 loci identify conserved regulatory elements between species and loci. Promoter elements are underlined, dashes indicate identical bases, and colons indicate gaps.

```
                                                                                                           NFκB
B21 BF2        -------C-- -CA-C--GGT -AA---A--- ----CGT--- ---------- -G---G---- --CG---T-- A--GG---:- ----------
NT WF06 Ia2    TCGGCCCTCT CTCCTGGCTC AGGGACGCGG :CCCGCAGCG GACACAGCCC GAGT:TCCCC CCGCTGCCCG GAGACCC:GG GGGGTTCCCA
OCB02 Ia2a     ---------- ---------- ---------- ---------- ---G------ ---------- ---------- ---------- ----------
OCB02 Ia2b     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
NT WF06 Ia1    ------ACT- -C-------- ------A--- ----CGTTG- ---------- TG--GG--G- T--G--TG-- -T-GG----- ----G-----
OCB02 Ia1a     ------ACT- -C-------- ------A--- ----CGTTG- ---------- TG--GG--G- T--G--TG-- -T-GG----- ----G-----
OCB02 1a1b     ------ACT- -C-------- ------A--- ----CGTTG- ---------- TG--GG--G- T--G--TG-- -T-GG----- ----G-----
B21 BF1        -----G-C-- -C-----G-T ---A--A--- C----GG-G- ---------- ----GC--TT T--G--T--- -GAGG-AG-- A---GA---C

                                                      ISRE            S Box                         X1X2 Box
B21 BF2        ---------- A--------- ---------- ---------- ---------- -----A---- --G------- ---------- ----------
NT WF06 Ia2    CACCGCGCCC TTC:CCCTCC CCCGCTCCGC GCTTTCGCTT TCGCTTCACA ACCTGCGGGA GCACATTCTG CCTGGCGCCC GATGACGTCA
OCB02 Ia2a     ---------- ---------- ---------- ---------- ---------- ---------- A--------- ---------- ----------
OCB02 Ia2b     ---------- C--------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
NT WF06 Ia1    ------T--- G--AT--G-- --------C- ---------- ---A--TC-- GT-CA----- -AG------- ---------- -G--------
OCB02 Ia1a     ------T--- G--AT--G-- --------C- ---------- ---------- --------A- ---------- ---------- ----------
OCB02 1a1b     ------T--- G--AT--G-- --------C- ---------- ---------- --------A- ---------- ---------- ----------
B21 BF1        ---------- G------G-- ---------- ---------- ---------- -----A---- --G------- ---------- ---------T

                                    Y Box
B21 BF2        -AT-AGACT- -AA--A---- -----GA-A- -C-A----G- -G-------- ---------- --G--G---- --TAGG---- --T-------
NT WF06 Ia2    CGCACGCTCC CGGCTGCCAT TGGCGAGGCG GAGGCGGAAG AACCAATGGG GGCGCGGGGC GGTGCTGAGG AGCCCAAAAA GCCGAAGGAG
OCB02 Ia2a     G--------- ---------- -----G---- ----G----- ---------- ---------- ---------- --T------- ----------
OCB02 Ia2b     G--------- ---------- -----G---- ----G----- T--------- ---------- ---------- --T------- ----------
NT WF06 Ia1    --AG------ --T-C----- -----G---- --------G- ---------- A----A-T-- --GA-A---C G-TGTCCCGG AA-CTG-AG-
OCB02 Ia1a     ----A----G --T-C----- -----GA--- --------G- ---------- -----A-T-- --GA-A---C G-GGTCCCGG AA-CTGAAG-
OCB02 1a1b     ----A----G --T-C----- -----G---- --------G- ---------- -----A-T-- --GA-A---C G-GGTCCCGG AA-CTG-AG-
B21 BF1        ---G------ ---TC----- -----G---- -CAA----G- ---------- -------T-- --GA-G-G-A CTGGTCCCAG AA--TG-AG-

                                                                 Start
B21 BF2        -T---CT--- T----CGGAC TTGAGAGTGC ---------A G------
NT WF06 Ia2    CGGCGTGGGG AGCGG::::: :::::::::: AGCGGTGCGC AGCGATG
OCB02 Ia2a     -----C---- ---------- ---------- ---------- -------
OCB02 Ia2b     -----C---- ---------- ---------- ---------- -------
NT WF06 Ia1    :::::::::: :::::::::: ::::C--C-G --T------- -------
OCB02 Ia1a     :::::::::: :::::::::: ::::C--C-G ---------- -------
OCB02 1a1b     :::::::::: :::::::::: ::::C--C-G ----A----- -------
B21 BF1        :::::::::: :::::::::: :::::AGCGG T------C-G G------
```

**Figure 8:** Southern hybridization of BAC and genomic DNA with a Class IIβ probe confirms orientation and presence of three Class IIβ loci in the turkey based on *in silico* and *in Vitro* restriction analysis.

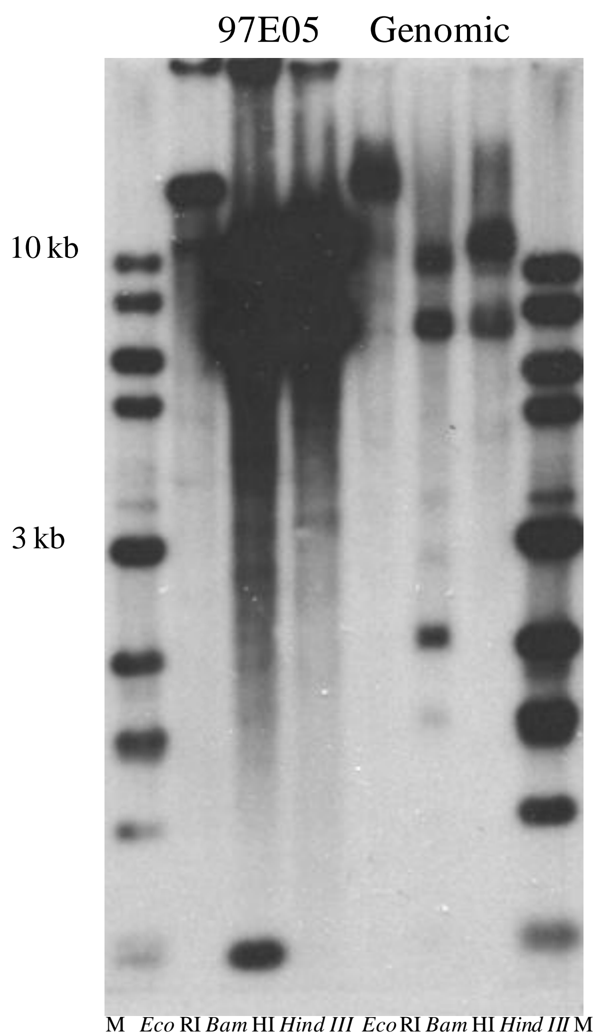


**Figure 9:** Class IIβ loci in turkey were generated from a hypothesized primordial rearrangement of a chicken-like MHC-B locus. A. Two chicken-like loci recombine in opposite orientation leading to locus duplication and promoter switching. The dominantly expressed BLB2 locus is red, the less expressed BLB1 locus is blue. B. Representaion of the turkey MHC-B Class IIβ loci indicating two loci originating from the BLB2 locus and one from the BLB1 locus through gene conversion.

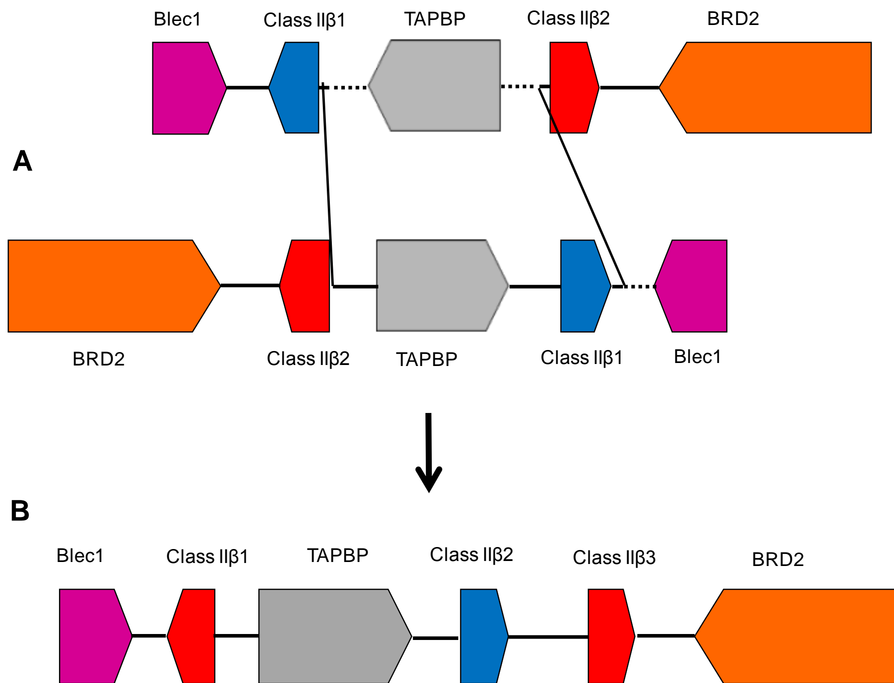


**Figure 10:** Predicted amino acid sequenced of Class IIβ in OCB02. The amino acid sequence of the Class IIβ loci show considerable levels of variation, especially in the peptide binding region encoded by exon 2 (underlined) and nearly complete identity between locus IIβ1b and IIβ3b.

```
Class IIb 1a MGSGRVPAAGAVLVALLALGARPAAGTRPSAFFLYGVISECQFQNGTERVRFVERQIYNRQQFTHFDSDVGKFVADTPLGEPQA
Class IIb 1b --------------------------------------F----L---------------Q--TS--------Y---S--D-Y-V
Class IIb 2a ------------------------------------TT-D---F-------L-D-YF---------------Y---------P-
Class IIb 2b ------------------------------------TT-D---F-------L-D-YF-------------------------T-
Class IIb 3a ----------------------------------S---F--H-V---QQ--H---DF-----LM--------Y-----------
Class IIb 3b --------------------------------------F----L---------------Q--TS--------Y---S--D-Y-V

Class IIb 1a EYWNSDTQYMEYKRGQVDRYCRHNYGVIESFTVQRSVEPKVRVSALQSGSLPETDRLACYVTGFYPPEIEVKWFLNGREETERV
Class IIb 1b -L----------R--E--------------------------------------------------------------------
Class IIb 2a --L--N-EFI--L-SA-----------F--------G----Q------------------------------------------
Class IIb 2b ------AEIL-DAQ-A--TF-------G----------------------------------------------------A---
Class IIb 3a ---------------S---F-------F--------------------------------------------------------
Class IIb 3b -L----------R--E--------------------------------------------------------------------

Class IIb 1a VSTDVMQNGDWTYQVLVVLETVPRRGDSYVCRVEHASLRQPVSQPWELPADAGRSKLLTGVGGFVLGLVFLALGLVVFLRGQKG
Class IIb 1b ----------------------------------------------VP------------------------------------
Class IIb 2a -----------------------------------------------P------------------------------------
Class IIb 2b -----------------------------------------------P------------------------------------
Class IIb 3a -----------------------------------------------P------------------------------------
Class IIb 3b -----------------------------------------------P------------------------------------

Class IIb 1a RPVAAAPGMLN
Class IIb 1b -----------
Class IIb 2a -----------
Class IIb 2b -----------
Class IIb 3a -----------
Class IIb 3b -----------
```

**Figure 11:** Class IIβ promoter regions share significant homology between species and loci. Consensus avian SXY sequence motifs were adapted from Zoorob et al. (128).Three turkey and 14 chicken haplotypes were aligned and compared at the CIITA-binding module. While most polymorphisms were shared between loci and/or species, two SNPs within the S Box (underlined) belong to loci with reduced/eliminated expression. A single C→T substitution (italics) was identified in a single chicken haplotype (B9) of which the functional significance is unknown.

```
            -95                                                                                  -60
ClassIIβ1     YGAGAGGGAGCCCCGCGGMGCMGAACTCTGCCCGGAGACGGGTGACGCCGCCCAGMGCCGCCGCCGCTCATTGGCCCTCCCCGCC
ClassIIβ3     TCGA--------------A--A---------------------------------M-----------------------------
ClassIIβ2     T--------C----A---C--C-------------------------------R-M-----------------------------
GG BLB1       TCR----A------T---C-YC----------Y---R----------------G-C-----------------------------
GG BLB2       TC---R-----Y--K---Y--MK-----Y---Y--------------------G-C-----------------------S-----
                    GGAGCC                  CCCGGAGACGGGTGACGCCGCCC             CTCATTGGCC
                     S Box                          X-X2 Box                       Y Box
```

# Chapter 4

## Initial Sequence Analysis of the Turkey MHC-Y Locus Supports Ancient Origins of Immunological Gene Clusters

# Initial Sequence Analysis of the Turkey MHC-Y locus Supports Ancient Origins of Immunological Gene Clusters

**Chaves L.D., S.B. Krueth, and K.M Reed**

Department of Veterinary and Biomedical Sciences,

College of Veterinary Medicine,

University of Minnesota, St Paul, MN 55108

## INTRODUCTION

A key component of the immune response found in vertebrate genomes is the major histocompatibility complex (MHC) which encodes many genes, including loci responsible for antigen presentation to T-cells. In placental mammals, this locus is typically divided into three functionally related regions, the class I and class II regions, separated by the class III region. MHC Class Iα genes, encoded in the class I region, present mostly endogenously derived peptides and are expressed on all nucleated cells. MHC Class II molecules, located in the class II region, present mostly exogenously derived peptides and are primarily expressed on antigen presenting cells (dendritic cells, B-cells, and macrophages). The class III region encodes genes not typically associated with antigen presentation including TNFs complement proteins, and additional, non-immune related loci.

In the turkey and chicken two genetically unlinked clusters, the MHC-B and -Y loci, located on the same microchromosome (GGA16 in chicken) define the MHC (39,47-49,96). In turkeys, a BAC clone from the MHC-B region was physically mapped by fluorescent *in situ* hybridization (FISH) to the distal p-arm, whereas a MHC-Y locus BAC was positioned near the centromere. An interstitial telomere array was located between the two regions (96).

In poultry species examined to date, the MHC-B regions contain both the classical Class I and Class IIβ genes, however, in reduced number and size compared to mammals. This significant constriction in size and gene content was first identified in the chicken and described as a “minimal essential MHC”. This minimal essential MHC is thus far exclusive to the turkey and chicken.

The MHC-Y locus (Y, *Rfp-Y*) was initially identified in the chicken through conflicting results in restriction mapping and serological typing (39). Studies of the MHC-Y show evidence for genetic effects of the non-classical MHC genes on disease susceptibility as well as transplant immunity (44,89,129).  Mapping efforts suggested two Class I and two Class II-like loci were located within MHC-Y (40). Two Class I Y loci were identified and cloned from a single cosmid clone representing a portion of the Y locus. At least one of these Class I Y genes is polymorphic and expressed in blood and spleen tissues as well as alloreactive within the species (42,130), while the other locus is an apparent pseudogene. An additional sequenced and assembled cosmid clone containing a portion of the chicken MHC-Y region has also been identified (43). This sequence contains two lectin-like loci and an additional Class I locus (YF) flanked by a CR-1 repeat element. This class I locus may also be expressed as many EST sequences matching the 3`UTR can be identified in the public databases. Together, three Class I-like loci have been identified in the chicken Y region, with varying levels of expression.

The core turkey MHC-B region has been sequenced and annotated (Chapter 3) and there is remarkable homology and synteny between turkey and chicken. However, little is known of the gene content, arrangement, or the evolutionary origin of the MHC-Y region in avian species. This work was undertaken to describe a mapped region of the MHC-Y locus in the turkey, to elucidate gene content and possible function, and to provide insights into the evolutionary origins of the MHC. Results of this study provide genomic resources for the studying of the turkey MHC in disease susceptibility and resistance.

## MATERIALS AND METHODS

*Sequencing and Assembly*

A bacterial artificial chromosome (BAC) clone from the CHORI-260 library (159B18) containing a portion of the MHC-Y region was previously identified by overgo hybridization (96). BAC DNA was purified using the Qiagen Large Construct kit, shotgun subcloned and sequenced as previously described (Chapter 3). Approximately 1200 subclones with an average insert of 3 kb were paired-end sequenced. Primer walking on selected subclones and additional PCR reactions based on overlapping EST sequences were performed to fill gaps and join contigs. Sequences were manually edited, aligned, and assembled using Sequencher software (Gene Codes, Corp.).

Southern hybridizations were performed to estimate repeat content, gap size, and verify assembly. One µg of BAC DNA was digested with the appropriate enzyme(s) for 4 hours with the supplied buffer (NEB, Inc), electrophoresed for 4 hr in a 0.8% agarose gel, blotted onto a Nytran N membrane, and UV cross-linked. Random $^{32}$P dCTP-labeled plasmid probes were prepared from using the NEBlot kit (NEB, Inc). Probes were denatured and hybridized for 2 hr at 65 °C with RapidHyb buffer (Amersham, Inc). The blots were washed twice with 5X SSC 0.1% SDS at room temp and twice at 65 °C with 0.5X SSC 0.1% SDS and exposed to ISC Kodak X-Omat autorad film for 30 min at room temp.

*Gene Identification and Annotation*

Sequences were analyzed with the basic local alignment search tool (BLAST) GENESCAN (http://genes.mit.edu/GENSCAN.html), and Softberry FGENESH (http://linux1.softberry.com/all.htm) to identify putative transcripts and homologies to

known genes. Comparisons between predicted gene sequences and available expressed sequence tags (ESTs) from poultry species were performed using Sequencher software. Repetitive elements were identified using REPEATMASKER and Tandem Repeats Finder (100) (http://tandem.bu.edu/trf/trf.basic.submit.html). Sequences were screened for tRNA elements using tRNAScan (101), CpG islands were elicited with Softberry CpGfinder (http://linux1.softberry.com/all.htm), and GC content analysis was performed with 100 bp windows using Isochore (http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html). Homologous chicken whole genome sequence was assembled where available.

## RESULTS

### *Sequencing and Assembly of the Y-locus*

Three non-repetitive supercontigs were assembled from approximately 1150 subclones. The largest supercontig (SC 1) was 67.6 kb and included 708 reads in 6 contigs (Figure 1). Four known gaps (totaling 4.6 kb) in regions unable to be sequenced and/or assembled were verified through PCR and Southern hybridizations. A second supercontig (SC 2) was assembled from three smaller contigs of 40, 11, and 3 reads, based on potential gene content (Figure 1). Orientation of SC 2 with respect to SC 1 within the BAC clone could not be definitively established. An additional 12.7 kb supercontig (SC 3) containing the BAC vector (12.7 kb) and limited flanking sequence (previously described through BAC end sequencing[19]) was assembled from 131

[19] Genbank DX922429 and DX922430

sequencing reads and contains a single estimated gap of 400 bp. Assembled supercontigs 1 and 2 have been annotated and submitted to Genbank[20] (Figure 2).

The genomic DNA insert contained in 159B18 is dominated by repetitive elements. Within SC1, repetitive elements (CR1, LTR, satellites, etc.) encompass ~22 kb (~33%) of the supercontig (Figure 1B). Moreover, the overall assembled sequence contains repetitive sequence not easily categorized (Figure 3). In addition to assembled contigs, many of the BAC subclones contain repetitive sequence with various motifs impeding assembly. Nearly half of all sequence reads contained at least a portion of the 535 bp turkey satellite repetitive element (MgaSat2). This element appears to be represented mostly in a single tandem array likely comprising the 5`flanking region of the clone (Figure 1, Figure 4). Southern analysis with a MgaSat2 probe (535 *Eco* RI repeat unit) identified only a single strong hybridization band of significant molecular weight in either *Hind* III or *Bam* HI digested BAC DNA. However, the large bands completely disappear when double digest with *Eco* RI, suggesting the large fragments contain significant numbers of tandem repeats.

Overall, the GC content of SC 1 was 52% with localized regions of 80% (Figure 1C, Figure 5). Additional repetitive sequence with high GC content generated short sequencing reads that could not be assembled. Included among these are a tandemly repeated, 77 bp *Pvu* II element (66% GC, 65 reads), a 47 bp *Msp* I repeat (61% GC, 14 reads), and a 159 bp *Alu* I element (67% GC). Another less well defined repeat element contributes to a large assembly gap in SC 1. Interestingly, the *Pvu* II element showed high sequence similarity to a chicken subtelomeric repeat[21].

---

[20] FJ601914
[21] AF124926

*Gene Identification and annotation*

Within the assembled contigs, genes homologous to loci found in the class I and class III region of mammals were present. Included are five loci: IαY1 (MHC Class I-like, YF), STK19, BAT1, G4-like, and BAT3, as well as two additional predicted loci (Predicted transcript 1 and 2) not previously associated with the MHC of vertebrates (Table 1). Unlike the portion of the sequenced chicken MHC-Y, this BAC clone did not contain lectin-like loci. Additionally, despite being mapped to Y in the chicken, no Class IIβ loci could be identified.

A single Class I-like locus (IαY1) within SC 1 was predicted *in silico*. The predicted gene was similar to, yet significantly different (85% nucleotide identity) from the class I-like sequence (henceforth referred to as $I\alpha Y^{p}$)[22] originally identified in the genome of the bird used to generate the turkey BAC library (96). For a comparative analysis to the whole genome sequence of the chicken, the unique 3` UTR of turkey IαY1 locus was used to identify and manually assemble RJF trace files most similar to IαY1. The IαY1 locus is also considerably different in gene structure from cloned chicken YF genes, YF transcripts (ESTs), and *de novo* assembled Red Jungle Fowl (RJF) whole genome sequence, suggesting this locus is a pseudogene (Figure 6). The predicted coding sequence of IαY1 and its RJF homolog are only 83% similar. IαY1 has several insertion/deletions residing in the coding sequence that would alter the reading frame. In addition, the sequence lacks a canonical splice acceptor site preceding one predicted exon. If transcribed and translated this locus would encode a significantly truncated protein. The partial turkey sequence previously identified ($I\alpha Y^{p}$) maintains a similar

[22] DQ836349

reading frame with the chicken loci, lacking the frameshift altering insertions and splice acceptor mutation, however, the overlapping predicted coding sequence (469 bp) is only 86% similar to the RJF.

Although its specific function is not fully understood, in two-hybrid assays the serine/theronine kinase (STK19) gene product localizes in the nucleus and may have a role in pre-mRNA (131). Coding sequence for STK19 could not be elucidated from the chicken whole genome sequence, however, comparison of the predicted transcript to the consensus coding sequence deduced from 6 chicken ESTs[23] identified 59 nucleotide substitutions resulting in 14 amino acid differences.

The coding sequence of BAT1 in both turkey and chicken[24] is 1287 bp (428 AA). Comparison between the two species found 61 synonymous nucleotide substitutions. Comparison of turkey and human[25] amino acid sequences indicates only eight amino acid substitutions. BAT1 encodes a factor that promotes pre-spliceosome assembly essential for pre-mRNA splicing (132) and has been suggested to have a role in nuclear export (133). As such, this gene would be expected to be under purifying selection.

Limited functional information is available for the single-exon open reading frame of G4. Yeast two-hybrid analysis suggests G4 (C6orf47 in human) may have a role in fibroblast growth factor receptor 3 signaling (131). The turkey G4 open reading frame encodes a membrane-bound protein that is weakly similar to a predicted zebrafish G4 locus. Neither corresponding ESTs in Genbank nor a homologous sequence in the chicken whole-genome sequencing reads could be identified.

[23] BU227616, BU240983, BU422929, BU439218, BU372428, CN224015
[24] AJ720994
[25] BAD96632

BAT3 has been implicated to play an important role in apoptosis in response to DNA damage (134). Only a portion of turkey and chicken BAT3 coding sequence could be defined by this BAC assembly, EST sequences, or the chicken whole genome sequence, preventing direct sequence comparisons.

Finally, two predicted loci present in the turkey MHC-Y encode novel genes not associated with other characterized MHCs. Predicted transcript 1 is weakly similar to CHIR-B3, a member of a cluster of immunoglobin-like receptors found in a region paralogous to the MHC on chicken microchromosome 31 (135). Gene predictions do not suggest a transmembrane protein like the CHIR-B genes, however, like CHIR-B loci, predicted transcript 1 possesses immunoreceptor tyrosine-based inhibitory motifs near the C-terminus. This locus is predicted without biological evidence of expression, and may be a pseudogene or possess a gene structure different than that predicted *in silico*. Predicted transcript 2 encodes a soluble protein however, no homologous sequence in any other species could be identified.

.

**DISCUSSION**

This report describes a preliminary sequencing effort of the turkey MHC-Y locus. Repetitive elements appear to dominate this region of the turkey (and chicken) genome, hindering sequencing and assembly. The region has amassed numerous transposon-like elements in addition to subtelomeric and potentially centromeric repeats and likely is the product of primordial chromosomal rearrangements. Confounding the chicken assembly of this region may be the heterozygosity of RJF #256 at chromosome 16, the bird from which the whole genome was sequenced. The best chicken assembly of chromosome 16

was published by Shiina et al. and encompasses 246 kb of the MHC-B region (30). The whole genome sequence assembly (version 2.1) for the MHC-Y consists of a large, unordered contig containing considerable gaps and repetitive elements.

The repeats and heterozygosity within the region will require alternative efforts to fully sequence and assemble the turkey MHC-Y. In addition to large insert, paired-end clone sequencing (e.g., BAC and cosmid), next-generation, non-Sanger based techniques may aid sequencing and assembly in these difficult regions. Indeed, closure of one gap located in SC 1 (at 54.5 kb) was assisted by this new sequencing technology and results from a preliminary 2x sequencing project (unpublished data, International Turkey Genome Sequencing Consortium). The anticipated full 7X genome sequence may fill even more gaps.

Along with satellites and transposable elements, the region of the turkey MHC contained in 159B18 has many less defined repeats. Previous mapping efforts physically placed this BAC clone centromeric to the MHC-B core. The large MgaSat 1 array as well as some of the other repetitive elements may indeed represent a portion of this chromosome's centromere, as centromeres typically have several classes of repeats (136). Further supporting this hypothesis is the presence of MgaSat 1 repeats located near the centromere of other turkey chromosomes (96). Additionally, centromeric retroviral and satellite elements have been found to be bound by centromere proteins and even express novel RNAs in *Drosophila* (137). Interestingly, the presence of a subtelomeric repeat located within 159B18, in addition to the previously identified interstitial telomere repeat located on the chromosome between MHC-Y and B, suggests evolution of this MHC chromosome may have involved chromosomal rearrangement and perhaps the fusion of

chromosomal elements. Such chromosomal rearrangements (inversions, *de novo* centromere creation/translocation) are not unexpected as several examples exist in mammalian MHCs (138,139)

The IαY1 gene identified in the BAC clone does not appear to be a functional class I locus. Nonetheless, at least one other Class I-like locus ($I\alpha Y^{p}$) exists in the turkey and is homologous with regards to gene structure to the sequenced YF genes in chicken. Whether there is a 1:1 correspondence between turkey and chicken in the number of Y locus Class I-like loci has yet to be determined. The comparative alignment of an RJF homolog of IαY1, the sequenced chicken cosmid containing a portion of the MHC-Y (43), and two additional YF loci sequenced (42) suggests there are at least three YF loci present in chicken (Figure 6). However, as cited in Miller et al. (29), Y haplotypes may vary greatly in Class I-like loci, with 16 YF alleles and/or loci identified in the RJF.

The IαY1 locus and homologous RJF sequence assembled herein are more similar to the chicken YF gene present in the cosmid clone sequence of Rogers et al. (43). However, the RJF locus assembly terminated approximately 1 kb upstream in the 5` UTR with a CR1 element, so a contiguous upstream flanking sequence could not be assembled to determine if this is indeed a locus independent from the cosmid sequence or if this locus was tightly linked to YLec loci as well. The total gene number, the repetitive nature of the region, and potential heterozygosity within Y will likely require concerted efforts to fully understand the Y-locus.

The sequenced turkey BAC contains genes homologous to the class I/III region of mammalian MHCs including the Class I locus, BAT1 and 3, STK19, and the G4-like locus. Two additional novel genes are predicted to exist within the turkey MHC-Y

region. The first locus, Predicted transcript 1, is weakly similar to an immunoglobin-like receptor found in chickens. Over 60 of these genes cluster on a single microchromosome, in a region orthologous to the human leukocyte receptor complex. Also identified in the MHC-B and Y loci of turkey and chicken are other lectin-like leukocyte receptors not found in other MHCs. The second novel gene (Predicted transcript 2) had no known homolog in any other species. While G4 is weakly similar to an MHC locus, both Predicted transcript 2 and G4 may be *in silico* artifacts, as neither are present in the current chicken genome build nor represented in EST datasets.

Lesser vertebrates (fish and frogs) contain many MHC gene clusters in conserved synteny with mammals; however such clusters are small and may be individually located on different chromosomes (140-142). It appears likely that during the emergence of the reptiles there was a consolidation of many of the MHC genes into an immune supercomplex as previously suggested (143,144). Unfortunately, limited data is available regarding the structure of the reptilian MHC, other than poultry species. Further consolidation, expansions, and/or deletions of the primordial MHC obviously has occurred during the delineation of birds and reptile-like mammals (Figure 7). For example, proteosome genes are noticeably linked to antigen processing genes in amphibians and mammals, however, they are separated from the MHC-B locus in birds (36). Similarly, TNF, LTA, LTB, and NFKBIL1 located between BAT1 and BAT3 in mammals may have been recently translocated there as they do not occupy the homologous region in birds or frogs. Alternatively, the MHC-B and/or Y homologs identified in poultry are small duplication blocks, and other (larger syntenic) blocks containing these additional loci exist elsewhere in the genome. Previous studies clearly

suggest this is not the case for the MHC-B; however, further efforts are required to make such an argument for the poultry MHC-Y region.

From a phylogenetic perspective it appears that the Class I loci were initially linked to the class II region and components of the antigen processing machinery, with subsequent relocation of Class II loci (140,141). The appearance of tightly linked classical Class I and II genes does not appear until the amphibian lineage (142). However, at least one Class I-like locus has been linked to Class II loci in trout (145). If indeed there are Class II genes/pseudogenes located in the Y locus, the Class II loci may have been translocated simultaneously with Class I loci after the condensing of the Class I and Class II loci subsequent to the divergence from amphibians.

Additional class I-like loci appear linked to some class III loci as far back as fish (140,141). Belov et al suggest the arrangement of placental MHCs, e.g. class I and class II regions divided by the class III region, was derived after the divergence from marsupials (143). However, the presence of Class I-like loci linked to class III genes in both birds, fish and monotremes clearly indicates this arrangement has a much more ancient origin and that the Class I region may have evolved differently in marsupials.

Comparison of the turkey and chicken MHC to mammals is not straight forward. Studies of monotreme, marsupial, and placental MHCs as well as avian loci allow one to hypothesize the development of a primordial MHC in vertebrates. In the homologous region of placental mammals the Class II loci are significantly expanded and the Class I loci (with the exception of few remaining pseudogenes) have been lost. The marsupial MHC, with a combination of characteristics similar to both birds and monotremes, provides a bridge between the primitive reptilian-like MHC of the platypus, the expanded

and evolved MHCs found in higher placental mammals, and indicates an alternate rearrangement hypothesis between monotremes, birds, and therians (Figure 7). Like turkey and chicken, monotremes have a constricted and tightly linked Class I and Class II cluster. In addition, monotreme MHCs also appear to be the product of a fission/fusion event. However, instead of inverting and fusing together like poultry, the monotreme MHC split and fused with two sex chromosomes (146). Many of the genes found in the turkey MHC-Y BAC clone (STK19, G4, BAT1, and BAT3) are located in the class III region (proximal to the class I region) of placental mammal MHCs, whereas the Class I-like IαY1 is similar to loci positioned in the class I region, which phylogenetically contain Class I/Class I-like loci.

The sequenced turkey MHC-Y region is not likely homologous to the 20 kb region[26] previously published for the chicken (43). Neither YLec loci nor the original turkey Class I-like locus ($I\alpha Y^{p}$) amplified and sequenced in NT WF06 (96) are present in this clone. This sequence likely represents a flanking (extended) Y region most homologous to the mamillian class I/class III border region and is physically located within the centromere region of the chromosome. A larger Y region encompassing more Class I-like, YLec-like as well as other loci is expected to flank the sequenced region. Position/orientation of the addition Y locus sequence is currently not known.

[26] AJ277927

**Table 1:** Loci identified in the BAC clone 159B18. For each locus, the BLAST results, closest match, and position of coding sequence within SC1 are provided. Coding sequence and amino acid length between turkey and chicken are based on comparative alignments, where available. Predicted protein domain/ characteristics are indicated.

| | | | | Position in SC 1 | | CDS Length | | Amino Acid Length | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Gene** | **BLAST Results** | **Genbank** | **Strand** | **Start** | **End** | **Turkey** | **Chicken** | **Turkey** | **Chicken** | **Protein Domains** |
| Predicted transcript 1 | IG-like receptor CHIR-B3 | XP_001236733 | - | 19819 | 20442 | 1019 | NA | 339 | NA | soluble, 2 ITIMs |
| IαY1 | Class I-like | NM_001030675 | + | 21461 | 23568 | 600 | 1107 | 199 | 368 | MHC superfamily |
| STK19 | Ser/Thr Kinase | NP_001121922 | + | 44321 | 48168 | 693 | 693 | 230 | 230 | STK19 superfamily |
| BAT1 | HLA-B associated transcript 1 | Q5ZHZ0 | + | 51151 | 67482 | 1287 | 1287 | 428 | 428 | DEADc, HELICc |
| | | | | SC 2 | | | | | | |
| Predicted transcript 2 | No Significant Similarity | | - | NA | | 842+ | NA | 280+ | NA | soluble |
| G4-like | G4-like gene | XP_001344067.2 | + | NA | | 600 | NA | 199 | NA | membrane bound |
| BAT3 | HLA-B associated transcript 3 | CAN59674.1\| | + | NA | | 522+ | NA | 173+ | NA | Ubiquitin family |

**Figure 1:** Contig assemblies of BAC clone 159B18. A. Seven loci identified in SC1 and SC2 representing the turkey region homologous to the class I/class III region of mammals. Exact orientation of SC2 is not known. SC1 is flanked by a large repetitive element (MgaSat2) of unknown total genomic size. Green bars represent gaps of unknown size in SC2 assembly. B. Sequence elements identified in SC1: Satellite (red), CR1 (blue), LTR (pink) and CpG islands (light cyan). Gaps of known size are indicated in yellow. C. GC content of SC1 with continous 100 bp windows.

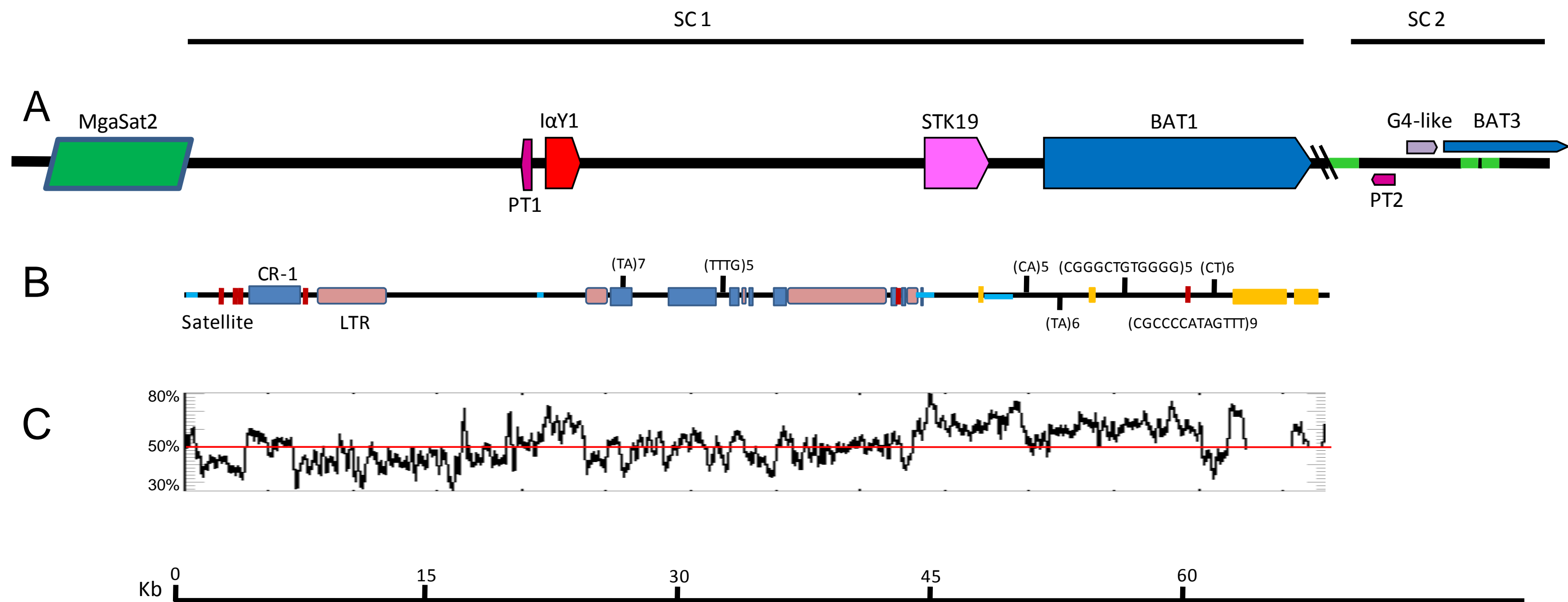


**Figure 2:** Genbank accession of clone 159B18 SC1.

```
LOCUS       FJ601914          67626 bp    DNA     linear   VRT 30-DEC-2008
DEFINITION  Meleagris gallopavo major histocompatibility complex (MHC) Y locus,
            partial sequence.
ACCESSION   1167662
VERSION
KEYWORDS    .
SOURCE      Meleagris gallopavo
  ORGANISM  Meleagris gallopavo
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria;
            Aves; Neognathae; Galliformes; Phasianidae; Meleagridinae;
            Meleagris.
REFERENCE   1  (bases 1 to 67626)
  AUTHORS   Chaves,L.D., Krueth,S.B. and Reed,K.M.
  TITLE     Initial sequence analysis of the turkey MHC-Y locus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 67626)
  AUTHORS   Chaves,L.D., Krueth,S.B. and Reed,K.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-2008) Veterinary Biosciences, University of
            Minnesota, 1988 Fitch Ave. 295 AS/VM, Saint Paul, MN 55108, USA
FEATURES             Location/Qualifiers
     source          1..67626
                     /organism="Meleagris gallopavo"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9103"
     misc_feature    1..67626
                     /note="derived from BAC clone CH260-159B18"
     misc_feature    1..67626
                     /note="major histocompatibility complex (MHC) Y-locus"
     satellite       complement(1927..2083)
                     /note="GGXHOI"
     satellite       complement(2982..3114)
                     /note="GGXHOI"
     satellite       complement(3273..3366)
                     /note="Sat-CR1_GG"
     repeat_region   complement(3641..6534)
                     /note="LINE/CR1"
     satellite       6876..7265
                     /note="Sat-CR1_GG"
     LTR             7836..8058
                     /note="GGLTR11-int"
     LTR             8309..8535
                     /note="GGLTR11-int"
     LTR             8888..9098
                     /note="GGLTR7-int"
     LTR             9099..9119
                     /note="GGLTR7-int"
     LTR             9120..9596
                     /note="GGLTR7-int"
     LTR             9580..9820
                     /note="GGLTR7-int"
     LTR             9823..10146
                     /note="GGERVL18-LTR"
     LTR             10184..10987
                     /note="GGLTR7-int"
     LTR             11072..11809
                     /note="GGLTR7-int"
```

```
LTR             11794..11955
                /note="GGLTR11-int"
gene            complement(15415..20442)
                /gene="Predicted transcript 1"
CDS             complement(join(15415..15548,16536..16671,17863..18021,
                19247..19356,19965..20442))
                /gene="Predicted transcript 1"
                /note="similar to IG-like receptor CHIR-B3"
                /codon_start=1
                /product="Predicted transcript 1"
                /translation="MTETLVHWAWKKGHLWEVAPIFSIDKWRDIRDRMWDSVSNGGKE
                EKLLQTLGPLWRSIINSLEAMQAERQVAAAAIQALAATTNLASEFDASVRLITHPACS
                SAGQVKASVEEAAPDLSNDTAAESSAVLQQVKGAIPDITPDGHCQAQPPAESQSDDGA
                SSPGPDSIRSSHPIPDRGEEKSRPPPSLLLPSALPLKATLVICIQNAQISEKHSSWKK
                RTGLEYLGIQNDWWVVAQYILTLDAGATKLPHHSRRAAVPSRRCPDSSAVQRLSQLET
                AALPLTNERLGAGQHKAVGDNTATDGCQDSCYPNQRSKYGHIHPFSFAKPGLQPPSER
                IVYS"
gene            21461..23568
                /gene="IaY1"
                /pseudo
CDS             join(21461..21569,21688..21948,22194..22473,22569..22839,
                22922..23041,23157..23189,23337..23369,23549..23568)
                /gene="IaY1"
                /pseudo
                /codon_start=1
LTR             complement(23801..24257)
                /note="GGLTR11"
LTR             complement(24256..25035)
                /note="GGLTR11-int"
repeat_region   25250..25763
                /mobile_element="LINE:CR1-G"
repeat_region   25257..25942
                /mobile_element="LINE:CR1-G"
repeat_region   26151..26220
                /mobile_element="LINE:CR1-C2"
repeat_region   complement(28915..29727)
                /mobile_element="LINE:CR1-Y3"
repeat_region   complement(29728..29868)
                /mobile_element="LINE:CR1-Y3"
repeat_region   30191..30316
                /mobile_element="LINE:CR1-Y3"
repeat_region   complement(30555..30827)
                /mobile_element="LINE:CR1-G"
repeat_region   complement(31652..31755)
                /mobile_element="LINE:CR1-F2"
repeat_region   complement(32276..32850)
                /mobile_element="LINE:CR1-F2"
LTR             complement(33285..33533)
                /note="GGLTR11"
repeat_region   complement(33648..33761)
                /mobile_element="LINE:CR1-H"
repeat_region   35072..35787
                /mobile_element="LINE:CR1-Y2"
LTR             35788..36091
                /note="GGERVL18-LTR"
LTR             36141..41410
                /note="GGERVL18-int"
LTR             41411..41656
                /note="GGERVL18-LTR"
repeat_region   41726..42037
                /mobile_element="LINE:CR1-Y2"
satellite       complement(42042..42211)
                /note="CTCTATGGGG"
```

```
     repeat_region   42217..42389
                     /mobile_element="LINE:CR1-Y4"
     LTR             43245..43374
                     /note="GGLTR7A"
     LTR             43576..43910
                     /note="GGLTR7A"
     repeat_region   43911..44073
                     /mobile_element="LINE:CR1-B2"
     gene            44321..48168
                     /gene="STK19"
     CDS             join(44321..44503,44803..44898,45416..45543,47713..47810,
                     47897..48038,48123..48168)
                     /gene="STK19"
                     /note="protein kinase"
                     /codon_start=1
                     /product="Ser/Thr Kinase 19"
                     /translation="MASRRRSGRFGKHRPGPEAVEAALREAAELFPRRSAALPALVLR
                     HQLYGLAGDRTAVDRHLNRLQAEGRIRLLHLGLGPDTVAVVGAEEYRETVLRAVEGSP
                     SAPLVRRFLDGVRPLLSFERGQLEEMGLKDCDVTQLVAAGVLTVRDAGSWWLALPGAG
                     RFIRALLHGRKQLLAAVRRSRHREVLQAELGQRRARPSLGLRYVLLDLLGAELLRSVP
                     TTSGPLLRLAET"
     gene            51151..67464
                     /gene="BAT1"
     CDS             join(51151..51361,52896..53023,54361..54453,56863..57046,
                     58110..58228,59371..59502,61792..61901,65599..65743,
                     65821..65968,67448..67464)
                     /gene="BAT1"
                     /codon_start=1
                     /product="HLA-B associated transcript 1"
                     /translation="MAENDVDNELLDYEEDEVENAAGGDGSEAPPKKDVKGSYVSIHS
                     SGFRDFLLKPELLRAIVDCGFEHPSEVQHECIPQAILGMDVLCQAKSGMGKTAVFVLA
                     TLQQLEPVTGQVSVLVMCHTRELAFQISKEYERFSKYMPSVKVAVFFGGLAVKKDEEV
                     LKKNCPHIVVGTPGRILALARNKSLNLKHIKHFILDECDKMLEQLDMRRDVQEIFRMT
                     PHEKQVMMFSATLSKEIRPVCRKFMQDPMEIFVDDETKLTLHGLQQYYVKLKDNEKNR
                     KLFDLLDVLEFNQVVIFVKSVQRCIALAQLLVEQNFPAIAIHRGMPQEERLSRYQQFK
                     DFQRRILVATNLFGRGMDIERVNIAFNYDMPEDSDTYLHRVARAGRFGTKGLAITFVS
                     DENDAKILNDVQDRFEVNISELPDEIDISSYIEQTR"
     satellite       59511..59717
                     /gene="BAT1"
                     /note="GCCCCACAGCT"
```

*Complete sequence is available from Genbank*

**Figure 3:** A. Dot homology plot of SC1 against itself illustrates local areas of repetitive alignments. B. Example of highly repetitive sequence near the 5` end of SC1.

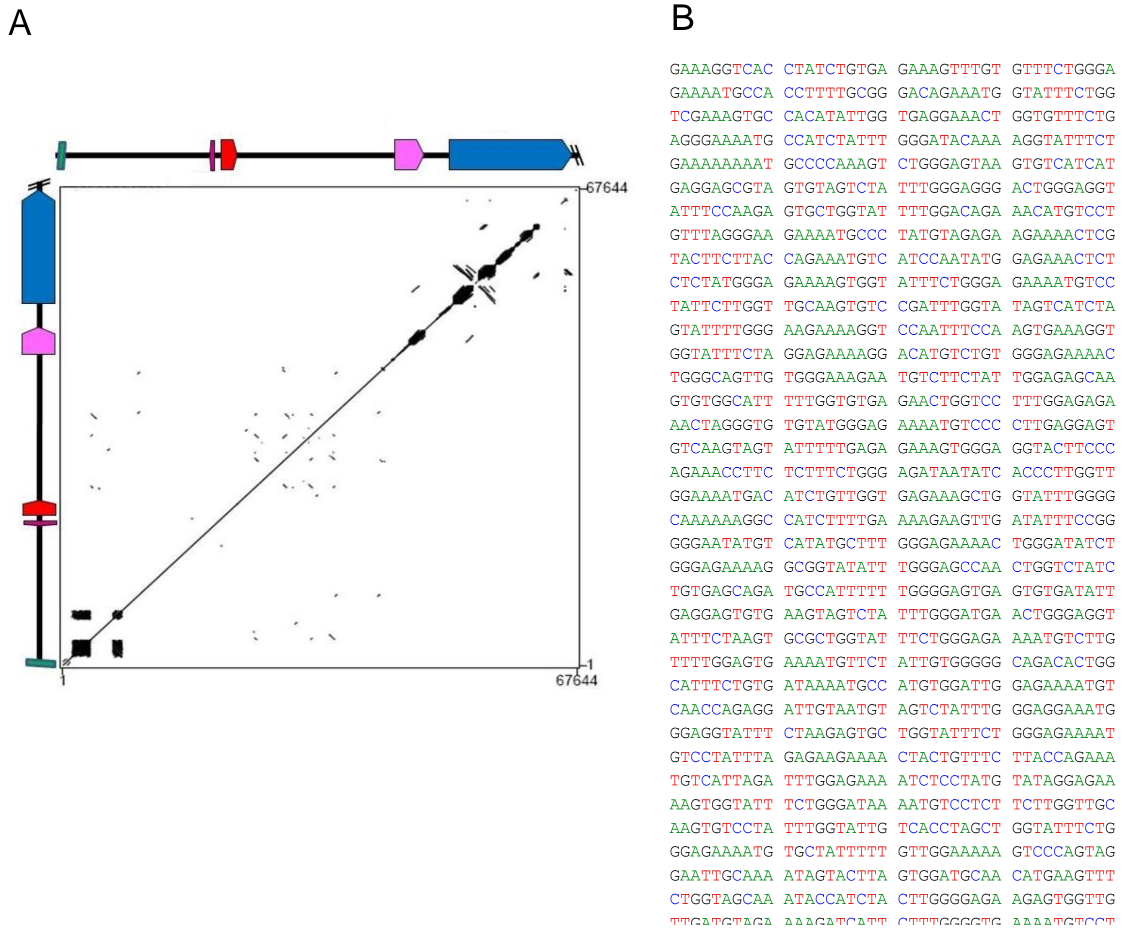


**Figure 4:** Southern hybridization identifies a single repetitive array of MgaSat2 (~535 bp *Eco* RI repeat) in clone 159B18. BAC DNA digested with *Hind* III or *Bam* HI and double digested with *Eco* RI. Lambda *Hind* III digest fragments indicated for comparison.

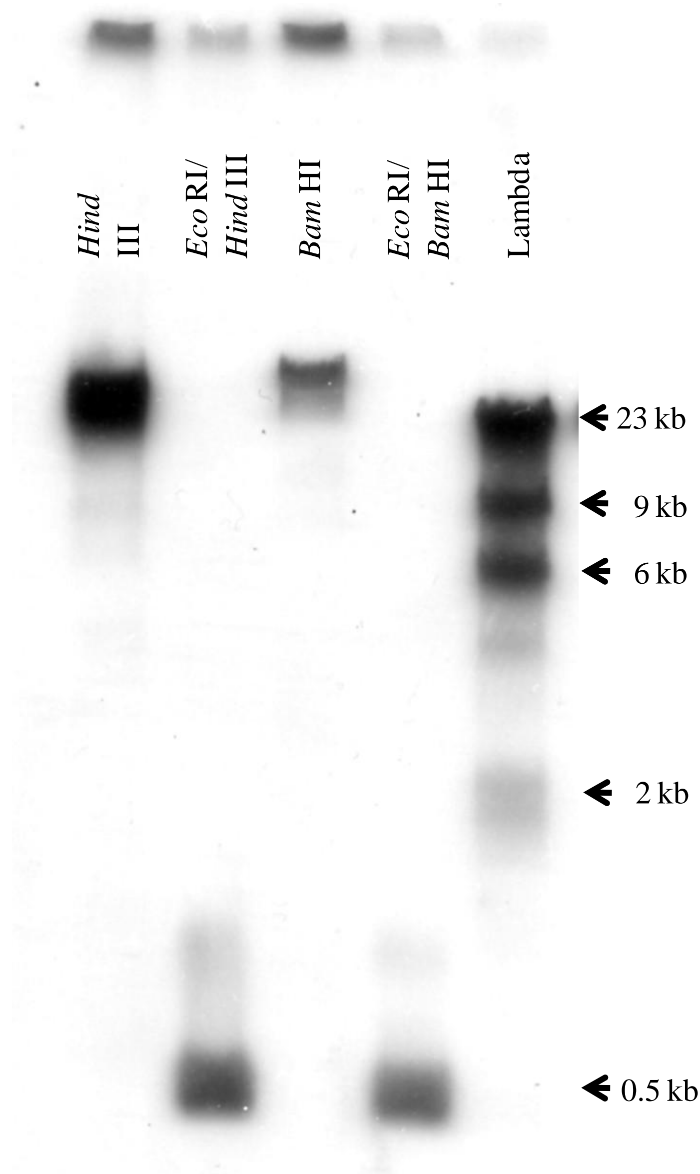


**Figure 5:** Representative example of GC-rich islands within the 3` end of SC1.

```
CTCCCAGAAC CTCCATCTCC CCCCAGTTCC TCCCAGAACC CCCCATCCTC CCACAACCCT CCATTCCCCC CCCCAAATCC
CCATCCCCTT CCCAGAACCC CTCATCCCCC CCCAAATAAC TCCTAGAACC CCCCCAACCC TCCCCAAATT CCCCATCCCC
GCCAATCCCC CCCCAACCCT TCCCACAATC TCCTCCCAAT TCTCTCATAG ACCCTCACTT CCCCCCCCCC AAATCCCTTT
TAGAGTCCCC CAATCTCCCA GAACCTCCCA TCCTCCCCAG ACGCCCCTTG GACCTCTCAA TCCCTCCCAG AACCTCCCCC
TTTCTCCCAC TCCCTCCCCT CACCCCCCAA TTCCTGACCC CCAACCTCTC GCAGAACCCC CCATACCTCC CCATCCCCAT
CAGACTCCCC CCAGTTCCTG CCTCCTGCCC CCCCCCCAGA ACCCCCCATT CCTGCCCCAC AATCCCCCTC CCAAACCCCG
CAATTCCTGC CCCCAACCCC TGCATTTCTC CTTAGCAGCC CCCAACCACC CAGATCCCCC CTCAATTCCT GCCCCCCAAC
TCCTCCCAGA TGCCCCCCAT CTCCCTATTC CCTCCCTCAG AACCCCCAGT CCCCCAGACC TCCCCCTAAT TCCTGCTCTC
ATCCCCCCCC CCCAGAACCC TCCACTGCTT TTCTTCTCAA CCCCCCAATT CCTGCCCCCC AACCCCTCCC AGATTCCACT
ACCCCCTCAC CCCCCACTTC TTGCCCCCCA ACCCCACCCA GAACCCCCAA TTCCTCTCAG ACCCCCACAT CTCCCCCCAT
TTCCTTCCAT CCTCCTCACT CCCCCAACCC TTCCTAAACC CCTCCCAGAT CCCTTCTTAC CCCCCCCGTT TCCCCCCCAC
TCCCCTCACC CCCCCAATTT CTGACCCCCC AACCCCTCCC AAATCCCCCC AAGCCCCCCA TTTCCCATCA GAACTCCCCA
GCGCACCCCC AGTTCCTGCC TTCATCCCCT CCCAGAATCC CTCGCTGCCC CCAAATCCTC CCCTCATCCC CCCACTTCCT
GTCCCTAAAC GCCCCCCAGA GCCCCCCAAC CCCTCCTAGA CTCCCCCTAG AGCCCCCCAA CCCTTCTCAG ATCTCCAGAG
CTGCTCACCT CTCCCATCTC CCTTCAGAAT CTCCCAACTC CTCCAGAGTC TGCCACTACC TCTCCCCTCA ACCCCCAGTT
CTAGGCCTCC AACCCCCCCC ATACCCCCCC CAGTTCCTGC TCTCATCCCC CCGAACCCCC AGTGCCCTTT CCCTCACCTC
CCCCCAATTC CTGACCCCCA ACTTCTCCAG ACTCCCAGAT CTGCTCACCT TCCCCATTTC CCTTCAGAAG CCCCCAACCC
CCCAAGGCTC CCCCACTACC TCTCCCCTCA CCCCCCCAAT TCCTGACGTC ATCCCCCCCC ACTTCCTGCC CCCCAAGCCC
CCCATTTCCC CTCAGAACCC TCCAGACCCC CGCAATTCCT GCCCCCAGAA CCTCTCACTG CCCCTCCCCT CATCCCCCCA
CTTCCTGCCC CCAACCCCTG CAGAGCCACC CACCTCCTGT TCCCCTCCCC CCCATTCACG CCCTGCCCCC CGGCCCCCAG
GTGGCGGTGT TTTTCGGAGG GCTGGCGGTG AAGAAGGACG AGGAGGTGCT AAAGAAGAAC TGCCCGCACA TCGTGGTGGG
GACGCCGGGC CGCATCCTGG CCTTGGCCCG GAACAAGAGC CTCAACCTCA AACACATCAA GCACTTCATC CTGGACGAGT
GCGACAAGAT GCTGGAGCAG CTGGGTGGGT TGTGGGGTCG TGGGGTGGGT TGTGGGGTTG TGGGGTCGCG GGGTGTGCTG
AGCACGTGAG ACACTTCGTC CTGGACGAGT GCGACAAGAT GCTGGAGCAG CTGGGTGGGC TGTGGGGTTA TGGGGTGGGC
TATGATGTTA TAGGGCTGAA GGGGGATTAT GGCGTGTAGG GTGTGCTGAG CACATGAGAC GCTTCATCCT GGATGAGTGC
GACAAGATGC TGGAGCAGCT GGGTGGGCTG TGGGGTGTGT TGTAAGGTTA CGGGGTGGGC TGTGGGGTTA GAGGGGGGTA
ATGGGTTTAT GGGGTGTGGG ATCTGCTGAG CATATCAAGG AGTTCATCCT GGGTGAATGC GACAAGATGC TGGAGCAGCT
GGGTGGGTTA TGGGGGCAGT TATGGGGTGG GCTATGGGGT TGTAGGGTGG ACTATGAGGT TATGGGGTAT GGGGTGTGCT
GAGCATATCA AGCACTTCAT ACTGAGTGAG CGTGAGATGC TGGAGCAGCT GGGTGGGCTG TGGGGCGGGC TGTGGGGCGG
GCTGTGGGGC GGGCTGTGGG GCGGGCTGTG GGGCGGGCTG TGGGGTGGGC TGTGGGGTGT GCTGAGCTCA TGAAACACTT
TATCCTGGAT GAGTGCGATG ATGCTGGAGC AGCTGGGTGA GTTATGGGGT GTGGGGTTAT GGGATCTGTT GTGGGGTGTG
TTGTGGGGCA GACTAGGGAG CTATATGGGA TGTGCTATAG GGCACGGTGT GGGGAGCTGG ATGGGATGTG CTATAGGGCA
CGGTGTGGGG AGCTGGATGG GATGTGCTGT AGGACAAGCT ATAGGGCATG TTACGGGGTT CTCTGTGGGT TGTGCTGTAG
GGCACGCTAT GGGGTGTTAT AGGGCACGCT GTGGGTCTGA CAGCCCTGTT CCCCCAGACG TGTAGGGAGG TGCAGGAGGT
GTGCATGGGG TGCCTTATGG GGTGTGATAT GGGATGTGCT GTGGGATGTC CTATAGGGCA CGCTATGGGA TGTGCTATAG
GGCCTGTTGT GGGGCGTTAG AGGGCACTGA CAGCCCTGTT CCCCCAGACA TGCGCCGGGA CGTGCAGGAG ATCTTCCGCA
TGACGCCCCA CGAGAAGCAG GTGATGATGT TCAGTGCCAC CCTGAGCAAG GAGATCCGGC CCGTCTGCAG GAAGTTCATG
```

**Figure 6:** Comparison of two turkey (96) and four chicken (42,43) MHC-Y class I loci indicate the single Class I locus identified in 159B18 is a pseudogene. Exons and introns are indicated by capital and lower case letters, respectively. Three mutations (boxed) include two reading frame deletions and a splice acceptor mutation. Identical nucleotides indicated by periods and dashes indicate deleations.

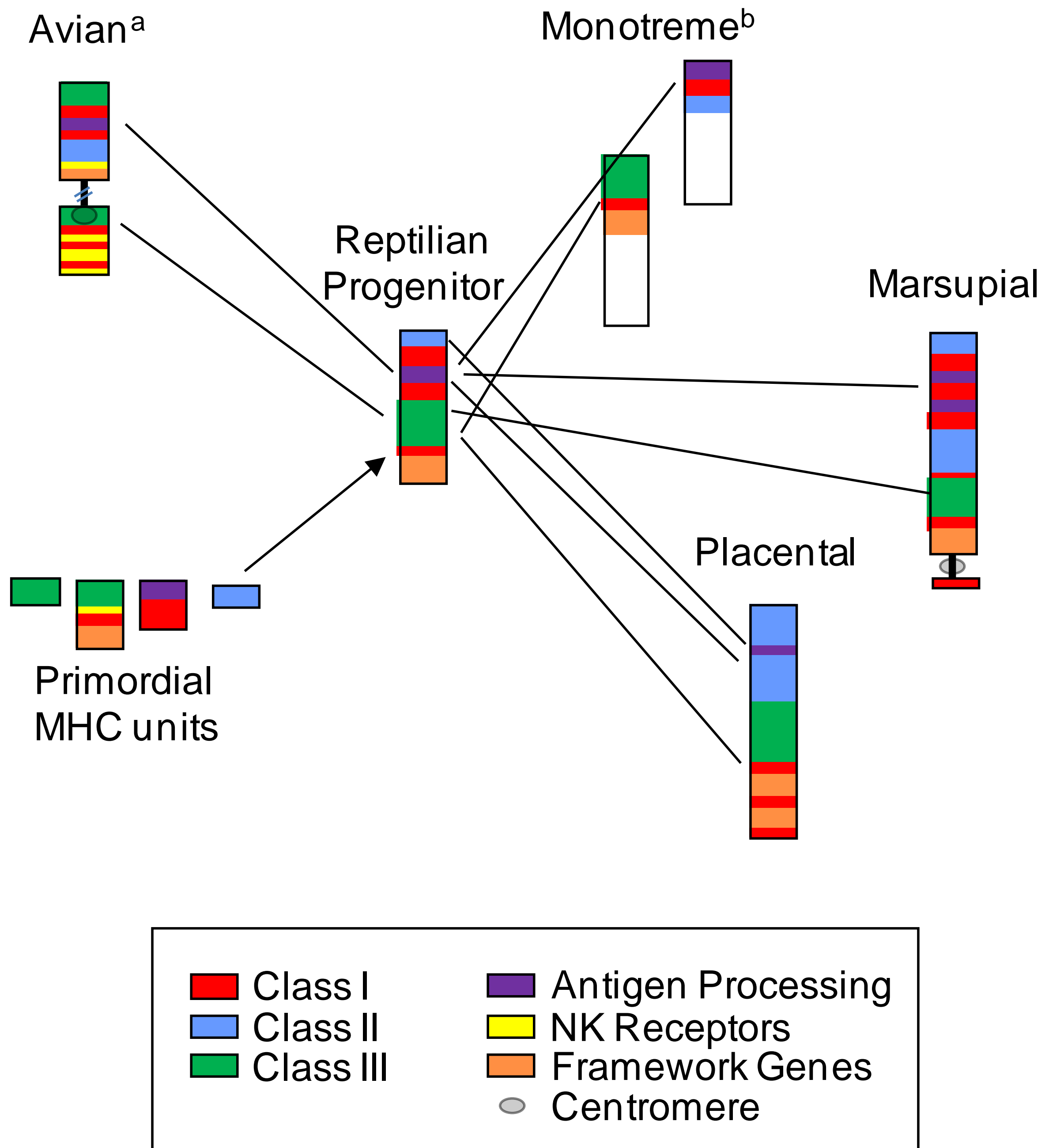
```
                      Exon 3|                                                                                              |Exon 4
RJF        AGCGCTGAAAAGGAGAGgtgaggatgggagggggacgtggggctgggctgggtgtggggcaggggctcagtgtgggctgctcagcccggtccacaacgtcattcatctgcagAGCGGCCCAA
Y-F        G...............................a...---............t........g....t..........g........at..c.......a...cc..c......T..A...TGT
YF1W*7.1   A.............................................................a.............g............c....c......cc..c......T..A....G.
YF2W*7.1   G..AG.................a.........a...........................................g............c.......a...ac..c......T..A....G.
IαYp       C.T....................cg..........t.c.a......ta............g.......g.......ga........t.ac......t....cc................TG.
IαYF1      T.T.....................c..........t...............ac...............ga....a.ga.........t.c...t...a...cc................TG.

RJF        GGTGCGAGTGTGGGGGAAGGAGGCCGACGGGATCCTGACCTTGTTCTGCCACGCTTATGGCTTCTACCCACGGCCCATCGCCATCAGCTGGATGAAGGACGGCATGGTCCGGGACCAGGAGA
Y-F        ........................G...................C........................G....................................................
YF1W*7.1   ...........................T................C..........C.C..T........G....................................................
YF2W*7.1   ...........................T................C..........C.C...........G....G.........................A.........A...........
IαYp       ..........C.............T..............T....C.....G....C.CA..........G.........A..............................A...........
IαYF1      .....A....C.............T..............T....C.....G....C.C...........GT....A..........-.....G......T....A.................

RJF        CCCACTGGGGGGGCGTCGTGCCCAACAGCGATGGCACCTACCACGCCTCGGCTAACATCGATGTGCTGCTGGAGGATGGGGACAAGTACCAGTGCTGCGTGGAGCACGCCAGCCTGCCCCAG
Y-F        .........................T..G........................GC..CT...............................G....C..........................
YF1W*7.1   ...G..........A......................................GC...T..........C..................TTG....C..........................
YF2W*7.1   ...G..........A..........T..G........T......A.T......GC...T..........C.......A..........T.G....C..........................
IαYp       ..............A................C....................C.C...........C..C..G...CT.......A.........C.T...................T....
IαYF1      ...........C..A................CA......G...T........C.C..............C..-...C.............G....C.T.......................A
                                                                                                            |Exon 5
RJF        CCTGGCCTCTTTTGTGGGgtgagcctggcagcatggggtgcgtggggttggggggtttgggggccactccttcctttcctgacaacactgctctcccccagAGCCACAGCCCAACCTGATC
Y-F        .............CA..........-.......g.....ca.........--...a..c.......g.c..................ggc................G...............
YF1W*7.1   .....T.......CA..........-.......g.....ca.........--...a..cgca...tg.c.........a........ggc.......t........G..............T
YF2W*7.1   .............CT......a...........g....a..t....a...--...a..........g.c....tg...a........gg.................G...............
IαYp       T.CA..T......CA..................g.............tc.-.....c.c....t..g.c..................g.c...............TG...............
IαYF1      ..CA.........CA...A..............G....a.c.t.............-c.......tg....................g...t....t.........G...............

RJF        CCCATTGTGGCCGGAGCGG------------------TCGTTGCCTTTGTGGCTGTCATCGCTGTTGTTGTTGGATTGGTGGTGTGGAAGAGCAAGTCAGgtaaaagcagagagctgggggg
Y-F        .......C...T..G..A...........................A.C................C..C......................................................
YF1W*7.1   ...........A..G..............................A.C................CG.C.....................................g.t..t...........
YF2W*7.1   .......A...TT.GCT........................CC....G....T.C.T....T..C.T.GA...C....AT..........TT.CT.......g..g............a...
DQ836349   .T.........T..G..A.....................A.CA..G.C.........G.G.....C.C.C..................G..............c.g......g....t...c
IαYF1      .......C......G....TCATCGCCATCATGGCTG..A.C...A.CA...............CCA................................C.....g......g..c.t...a
                                                                                                       |Exon 6
RJF        agaaagcaggaactgcaggtggggtctgaacc-ccctt-gggagtccccatgctctgacatgagcccaatgctgacacttctttctgtctgcagGGAAGG
Y-F        .................a......c.............t....a.g......g.t............c................................
YF1W*7.1   ...........g............c...g..............a.g......t.............tt...t.c............a....t........
YF2W*7.1   ......gg...g............c...g..............a.g......g.....................g.....t...................
IαYp       ....g.t-------.gg..g...aggc.g................g..tgc.........c.......................................
IαYF1      g...g.cg..gg.......g....gggct.aat.a.c.ca.....g....cat.......ca..-...c........................t......
```

**Figure 7:** Model of MHC chromosomal divergence between mammals and avians from a hypothetical reptilian progenitor. Placental map (Human) based on Horton et al. (13) Marsupial (Opossum) map adapted from Belov et al. (143), Monotreme (Platypus) based on Dohm et al. (146). (a) The two avian MHC chromosomal regions cannot be oriented with respect to each other and may be inverted. (b) Study of monotreme MHC is limited to two BAC clone sequences and may contain additional Class I and Class II loci similar to avian and marsupial.



# Chapter 5

# Haplotype Variation, Recombination, and Gene Conversion within the Turkey MHC-B

# Haplotype Variation, Recombination, and Gene Conversion within the Turkey MHC-B

**Chaves L.D., Faile G.M., S.B. Krueth, and K.M Reed**

Department of Veterinary and Biomedical Sciences,

College of Veterinary Medicine,

University of Minnesota, St Paul, MN 55108

## INTRODUCTION

The major histocompatibility complex (MHC) is a large genomic locus possessing genes involved in the adaptive immune response. It has been identified as the most polymorphic and gene dense region in the vertebrate genome (11). The MHC contains many genes including cytokines, complement proteins, and the highly polymorphic products responsible for processing and presenting endogenous and exogenous antigens to T-cells (13).

In the turkey and chicken, the MHC is currently divided into two genetically unlinked regions (MHC-Y and -B loci) located on the same microchromosome (39,47-49,96). The MHC-Y locus contains non-classical MHC genes, lectin-like loci, and other non-immune related genes (Chapter 4). The MHC-B locus is generally regarded as the homolog to the mammalian MHC, containing most of the antigen processing and presenting genes. However, the total gene number and overall locus size is significantly reduced in poultry (turkeys and chicken) as compared to mammals (Chapter 3).

The mammalian MHC is large (~ 3-7 Mb) and contains ~ 200-400 genes and pseudo-genes (11,13). In contrast, the turkey MHC-B is composed of 34 genes spanning ~200 kb (Chapter 3). This significant constriction in size and gene content was first identified in the chicken and has been described as a “minimal essential MHC” (28,147). The minimal essential MHC may not be inclusive to all bird species, as studies of the closely related quail suggest an expanded MHC-B locus (37). Single loci in the human MHC have 100’s of alleles, potentially contributing to 100,000s of haplotypes (148). In contrast, commercial chicken lines have significantly fewer alleles at loci of the MHC-B locus, thereby limiting the number of potential haplotypes within subpopulations (29).

While evidence suggests several common commercial chicken MHC-B haplotypes are recombinants, recombination is rarely observed experimentally. With a low rate of recombination, genes/alleles of the avian MHC-B can co-evolve as haplotypes. This was first hypothesized by Kaufman et al. (147), where several serological studies found no recombination between the Class Iα and Class IIβ loci in approximately 25,000 offspring. A previously published genetic map of the turkey MHC-B locus based on ~200 offspring identified a genetic distance of 0.6 cM between two markers in the region from C4 to BTN2 (96). Recombination between Class Iα and Class IIβ loci has yet to be observed in the turkey.

In chickens, haplotypes of the B locus have repeatedly been found to confer resistance or susceptibility to numerous pathogens (147). The core turkey MHC-B region has been sequenced and annotated from a single partially inbred bird (Chapter 3) and there is significant conserved synteny and homology between turkey and chicken. However, little is known of the genetic variation and diversity within the species or within the subpopulation of turkeys used by commercial breeders. This work was undertaken to analyze the extent and origin of variation within the turkey MHC-B locus. Results of this study demonstrate the potential diversity available in commercial flocks and provide genomic resources for studying the effect of turkey MHC diversity in disease susceptibility and resistance.

## MATERIALS AND METHODS

*Genomic Resequencing*

Turkey sequences[27] were used to design primers to resequence 15 interspersed amplicons across the MHC-B region (Figure 1, Table 1) using a panel of 52 DNAs consisting of the P and F1 generations of the NTE and UMN/NTBF resource populations (80,149), 6 birds from 5 Nicholas breeder lines ( [S01-S30] Chapter 7), and NT WF06, the DNA source for the original MHC-B locus sequencing (Chapter 2). Most primers were designed to anneal within exons to reduce null alleles. PCR reactions were performed as previously described (Chapter 3). Products were purified using a MinElute PCR Purification Kit (Qiagen), and sequenced with an automated ABI Sequencer. Data were manually analyzed using Sequencher software (Gene Codes, Corp.) and sequence polymorphisms (SNPs and insertion/deletion [indels]) were recorded.

*Haplotype Identification*

Polymorphisms were analysed using PHASE v2.1 and FastPHASE v2 software (150) to identify haplotype diversity and frequency within the 52 resequenced birds. Initially, polymorphisms within individual amplicons were analysed. Allele fidelity and individual haplotypes were identified using parent/offspring trios from the UMN/NTBF population. Genetically deduced haplotypes and those identified from monomorphic individuals were used to conduct a second FastPHASE analysis. Fifteen iterations with random starts were performed and majority/plurality haplotypes were assigned to

[27] Genebank DQ993255 and EU522671

individuals. Haplotype trios were re-examined to confirm correct assembly. Phylogenetic analysis of the haplotypes was performed using ClustalW software (151).

*Genetic Mapping*

The UMN/NTBF resource population was used to identify recombination frequencies within the MHC-B region. Selected SNPs were genotyped using PCR restriction fragment length polymorphism (PCR-RFLP) assays as previously described (152). Large insertion/deletions and microsatellites were genotyped by PCR and electrophoresis on agarose and/or acrylamide gels as previously described (75). Genotypes were combined with previously published data (75) and analysed using FIXED and CHROMPIC functions of CRI-Map software (153) for multi-locus linkage and gene conversion analysis.

**RESULTS**

*MHC-B Polymorphisms*

Over 14 kb of the MHC-B region was resequenced on 52 birds derived from commercially significant breeder lines. A total of 203 SNPs were identified with minor allele frequencies (MAF) ranging between 0.01-0.49 and averaging 0.15 (Figure 2). Thirty-four percent of the SNPs had an MAF $\geq 0.2$. The frequency of SNPs in this region, 1 SNP/70 bp, is higher than the 1 SNP/200 bp found in other regions of the turkey genome (154). The majority of SNPs were transition substitutions and located in introns (Table 2). Of the 34 SNPs located in coding sequence, one-third represent non-synonymous substitutions.

In addition to SNPs, several indels were present in the study samples. Nine simple repeat elements ranging from single nucleotides to hexomeric elements were identified and genotyped through resequencing. Included in these is a large polymorphic repeat element present within intron 6 of *LAAO*. Examination by electrophoresis on 3% agarose and 5% denaturing acrylamide gels found four alleles at this locus ranging in size from 180 to ~300 bp. A large pentameric repeat with ~60 repeat units occurred within intron 1 of *TRIM 27.2* in NT WF06. Another large, complex polymorphic element was found in the *TRIM 7.1* amplicon (intron 1). However, alleles at this locus could not be definitively scored. Also present within intron 6 of *TRIM 7.1* was a microsatellite repeat (TA+CTTT). A large deletion (~500 bp) was found in some individuals within the first intron of *TAP2*. Finally, the microsatellite locus MNT-482 is a pentanucleotide repeat located within intron 4 of the *TAPBP* gene.

*Haplotype Reconstruction*

Amplicons were initially analyzed separately to identify haplotypes within each locus and to identify potential null alleles. Individual loci had between three and seven haplotypes, averaging 5.3 haplotypes pre locus. Haplotype analysis was at first performed using PHASE software on a suite of 11 SNPs (*TRIM 39.1, TRIM 41*, and *BTN1_1*) with MAF $\geq$ 0.2 that were fully genotyped on all 52 individuals (Table 3). Analysis of this minimal shared dataset identified thirteen haplotypes occurring from 1 to 25 times (average occurrence of 8). Nineteen individuals were monomorphic at these loci. Four haplotypes appeared in only single individuals.

Haplotype variation differed considerably between the five commercial breeder lines. The initial analysis of 11 SNPs identified MHC-B heterozygosities between 0.5 in Line 2 and 1.0 in Line 4. Heterozygosity at the MHC-B was higher in all lines as compared to the whole genome averages (Chapter 7). The level of heterozygosity correlates to the level of MHC diversity (Figure 3). The four unique haplotypes each occurred in a separate breeder line. The number of haplotypes within lines ranged from four to six. Line 2 had the fewest haplotypes, one of which is unique (H) and another only found in that line (I). Haplotypes F (one occurrence) and G (two occurrences) were exclusively found in Line 3.

An expanded analysis with 49 SNPs (MAF $\geq 0.2$) spanning *TRIM 7.2* to *C4* was performed with FastPHASE incorporating 11 phased haplotypes identified from monomorphic samples and through analysis of UMN/NTBF trio's. F1s and their respective sires and dams were examined to identify the phase of MHC-B haplotypes present in the UMN/NTBF population. Missing genotypes were imputed by FastPHASE and haplotypes were assigned from fifteen software iterations. Forty-five haplotypes represented from 1 to 16 times were identified (Table 4). Fourteen haplotypes occurred more than once (31 unique) with an average occurrence of 2.3. Six individuals with four differing haplotypes were found to be monomorphic within the MHC-B, including NT WF06, the source for the only fully sequenced haplotype (Chapter 3).

In the expanded analysis (49 SNPs) eleven different haplotypes were identified in Line 1, four haplotypes in Line 2 (3 individuals were monomorphic), ten haplotypes in Line 3, seven haplotypes in Line 4, and eight haplotypes in Line 5 (one monomorphic

individual). Two individuals from Line 2 shared the haplotype of NT WF06 which is from a subline of line 2 (D. Harry, Pers Comm).

Phylogenetic analysis of haplotypes reconstructed from 49 SNPs identified three large clades containing haplotypes identified through trio's analysis, monomorphic individuals, and computer reconstruction (Figure 4). As few as one SNP separated some haplotypes, however variation between the major clades was much greater. Haplotypes 1-4 share approximately 40-50% of alleles. The phylogenetic relationships of the 45 haplotypes derived from 49 SNPs correlate well with the haplotypes identified from 11 SNPs (Figure 4).

*Recombination Mapping*

Markers segregating within a subset of the UMN/NTBF resource population (offspring from two females, 1044 and 1042) were selected for genotyping and linkage analysis. Nine SNPs and five indel/microsatellites were genetically mapped. Initial analysis identified a chromosome map of nearly 20 cM; however, within this region six chromosomes were identified with CRIMAP (CHROMPIC) to have undergone gene conversion in three different segments of the B locus (Figure 5). These genotypes were removed and a second linkage analysis was performed. The 14 markers within the ~ 200 kb turkey MHC-B locus generated a genetic map of 3.6 cM (Figure 5). This chromosomal region has a considerably higher recombination rate (18 cM/Mb) as compared to the whole genome average (2.59 cM per Mb) previously determined for the turkey (75).

Three MHC-B recombinants were identified. One recombinant occurred between *C4* and *TAP2*. However, additional markers would be required to identify whether this event was located between the *Class Iα2* locus and the antigen processing genes. The other two recombinants occurred between *Blec1-BTN1* and *LAAO-Bzfp2* upstream of the antigen processing and presenting loci.

**DISCUSSION**

The MHC is the most polymorphic region in vertebrate genomes, with numerous alleles at differing loci contributing to an enormous number of potential haplotypes. Contributing to this variation is random mutation, recombination, gene conversion, and in some instances, large and complex insertions, deletions, and gene duplications (36,155). An analysis of 8 human MHC haplotypes identified ~ 12.5 bp/SNP (155), considerably higher than that described here for the turkey (70 bp/SNP) or the chicken (21 bp/SNP) (36). The increased number of classical (and highly variable) MHC loci in humans as well as the reduced effective populations in commercial turkeys and chickens can explain these differences. In the human and chicken studies, whole haplotypes were resequenced including the most variable antigen presenting genes (Class I and Class II loci). The present study, however, only included selected regions not including the antigen presenting loci in order to avoid null amplifications of the exceedingly polymorphic MHC loci. Two similar studies of the turkey MHC-B examining TAP1, TAP2 and TAPBP segments identified similar levels of sequence polymorphism in a set of commercial and local domestic turkeys (63,64).

Among the MHC-B loci examined in this study, BTN1_2 showed the greatest level of polymorphism (46 bp/SNP) while the adjacent Blec1 had the fewest (101 bp/SNP). Factors contributing to the disparity of polymorphism between loci include the ratio of exon to intron sequence examined for the individual amplicons and the possibility of null amplifications due to primer-site polymorphisms thereby reducing the SNPs detected. Additionally, insertion/deletion events are problematic in direct resequencing efforts using Sanger sequencing technology, as individuals heterozygous at an indel invariably cannot be sequenced through such polymorphisms. Further, large indels (e.g. *LAAO*) may cause underrepresentation (null) alleles due to preferential amplification of the smaller product. With ~ 200 total polymorphisms detected, up to $1.6 \times 10^{60}$ combinations are possible within the B-locus assuming that all combinations are independently possible. Considering the 49 SNPs examined with FastPHASE, $5.6 \times 10^{14}$ haplotypes are possible. Linkage disequilibrium within the turkey MHC-B was high enough to reduce the number of identified haplotypes to 45 (Table 4). An analysis of 67 inbred rat lines using 67 microsatellites within the 3.5 Mb RT1 complex identified 28 haplotypes (156). The delineation of forty-five haplotypes within 52 turkeys may be an overestimate due missing/incorrect genotypes. However, only six individuals from commercial breeder lines were identified as monomorphic indicating the persistent heterozygosity within these partially inbred commercial lines.

The number of polymorphisms described in the turkey will be useful in identifying MHC-B haplotypes for genetic association studies. In the chicken, a single complex microsatellite, LEI0258, contains nearly enough alleles to independently identify haplotypes in the chicken MHC-B (157). The homologous sequence in the turkey

has thus far been found to be monomorphic (80). Alternatively, the microsatellite in *TRIM27* contains a very large penta-nucleotide repeat. Although alleles of this repeat are very diverse, genotyping with capillary sequencing technology has been less than reliable. The 5 bp differences separating alleles are difficult to consistently score in the ~400-600 bp products of PCR amplification at this locus. Additionally, the results of resequencing of the UNM/NTBF sire 7491 suggested this individual was monomorphic within the MHC-B region. However at this locus, 7491 is heterozygous suggesting the complex repeat is unstable through multiple generations, and thus an unsuitable marker for haplotype definition. At this point no single locus appears adequate to independently define turkey MHC-B haplotypes.

Turkey MHC-B haplotypes are clearly derived from a combination of novel mutation, recombination, and gene conversion, all of which were identified in the current study. Similar events have been observed in the chicken (36). Phylogenetic analysis identified loosely clustered haplotypes. Some individuals identified as sharing the same haplotype may actually possess functionally distinct haplotypes not detected by the analysis of ~50 SNPs (MAF $\geq$0.2). Rare SNP alleles may simply be novel mutations or indicative of novel functional haplotypes. The identification of turkey MHC-B haplotypes through PCR resequencing and SNP haplotyping does not necessarily identify functionally distinct alleles at the antigen processing and/or presenting alleles. Nonetheless, this report has surveyed enough individuals to identify several dominate haplotypes within commercial breeder lines. Further, several individuals have thus far been identified as monomorphic, allowing for a more thorough and in-depth resequencing effort similar to that recently performed on chicken lines (36).

The majority of haplotypes were inferred *in silico* and not experimentally identified. Alternative allele segregation and/or imputation could either increase or decrease the total number of haplotypes present in the study population. Most clusters contain many unique haplotypes differing at few SNPs. These haplotypes could be mis-identified through genotyping or *in silico* errors and may in fact be identical. The most represented—and most variable—clade contains two monomorphic individuals. Resequencing of the antigen processing and presenting genes should identify the true level of functional diversity within clades. The true number of SNPs required to identify different haplotypes may require more complete resequencing of different haplotypes as well as individuals from differing lines with the same (predicted) haplotype to identify the stability of this locus. In the chicken, resequencing of approximately 60 kb identified eight nucleotide substitutions and four indels from two chickens of different lines with the same haplotype, suggesting some haplotypes are very stable (36).

While frequent in the larger mammalian MHCs, recombination has rarely been identified between the chicken Class Iα-Class IIβ regions. It is suggested that due to their close proximity, the avian antigen presenting genes (Class Iα1-2, Class IIβ1-3) and antigen processing genes (TAP1-2, DMA, DMB1-2, TAPBPL) may have co-evolved to optimize antigen presentation given the limited number of MHC antigen genes. Results presented in this study indicate that recombination occurs within the MHC-B region at a higher frequency than the whole genome average. Similarly, recombination between the Class Iα and Class IIβ loci and BG loci located approximately 150 kb away is nearly 5 cM in chickens (147).

Along with positive selection for increased genetic diversity, gene conversion may significantly contribute to the diversity and homogenization of the turkey MHC. Six incidents among 86 offspring have been recorded within the region. It is hypothesized that the gene duplication/rearrangement of the turkey MHC-B (e.g. TAPBP and Class IIβ2) with respect to the chicken MHC-B is likely a product of non-homologous/inverted gene conversion (Chapter 3).

Despite their close physical proximity, the MHC-B and -Y loci are genetically unlinked in both turkey and chicken. The repetitive (and thus recombinogenic) nuclear organizer region was once thought to divide the two MHC loci in chickens, explaining the great genetic distance. However, research in both turkeys and chickens has disproved this early hypothesis (Chapter 2 and M Miller, Pers Comm). Instead, the two loci may indeed be genetically separated simply due to the high recombination rate along the chromosome. If one assumes an average of 20-30 cM per Mb, the distance between the MHC-B and -Y may be about 1.6-2.5 Mb, a distance not unlike primitive mammalian species (143). Alternatively, the highly repetitive, GC-rich sequence earlier considered part of the centromere may actually contribute to the extreme levels of recombination (Chapter 4). Additional MHC-B and Y sequence will soon be available for SNP identification and analysis. With additional distal loci it may be possible to identify genetic linkage between the two regions.

If the MHC-B is taken as a single locus, heterozygosity among the commercial lines is higher in this region than at other loci throughout the genome (Chapter 7), suggesting a selective pressure towards heterogeneity. The level of diversity and recombination within the commercial turkey MHC-B is consistent with the levels seen in

the closely related chicken. However, it is still not known to what extent the total genetic variation of the species is represented within commercial birds. Variation within the B-locus has yet to be studied in wild turkey populations. Additionally, no data are available for either turkey or chicken regarding the genetic diversity of the MHC-Y locus. There are likely at least 8 Class I Y loci (genes and/or pseudogenes) in the chicken of which at least one is polymorphic and transcribed (42). Furthermore, the MHC-Y contains lectin-like loci similar to natural killer receptors. The extent of polymorphism and linkage disequilibrium between these loci and Class I loci, from either MHC-B or Y, is also of interest.

**Table 1:** Primers and reaction conditions for MHC loci examined in commercial breeder birds. MNT-482 and loci denoted with “ms” or “indel” were genotyped by length polymorphism.

| Locus | Forward Primer | Reverse Primer | Product Size | TM | Ext Time |
|---|---|---|---|---|---|
| TRIM 7.2 | TTCTCCAACTCGAACCGAAG | CCAGCTGGACAAACTGCTG | 848 | 62 | 1min |
| Bzfp2 | ATTCAGCGCGATCCTAAATG | GTCTCCTTGGGTACGTGGAG | 941 | 58 | 1min |
| 20K | AGAATTGGGACCCCTAAAGC | AGGCAAGCGTGCAAAATAAC | 882 | 62 | 1min |
| LAAO | ATTTGCTGAACGAGGATTCG | CCTTCGTGCTCAAAGAGAGC | 1000 | 58 | 1min |
| LAAOms | ATTTGCTGAACGAGGATTCG | AGCAGCTCCAAGTGAGGTTC | 180 | 58 | 30s |
| TRIM 7.1 | CCTGGGTTTTCTTATCATCG | CGGTCTGCTTGGAGTTCTTC | 1682 | 62 | 2min |
| TRIM 7.1ms | AATCGGCCTTTGATTCCTG | AAGGTGTAACGGCCCCTAAG | 180 | 58 | 30s |
| TRIM 39.2 | GCAAGATCTCTGCCACCTTC | TCAGTGAGCAGATCCGTGAG | 986 | 58 | 1min |
| TRIM 27.2ms | GCATCACACCTGAATGGTGAC | TGTCCAGGAATACAAGGTAAAGG | 400 | 58 | 1min |
| TRIM 39.1 | GGCAGGGATAATGAATGAGG | TATGTGGCCCCTCTGAAAAG | 1141 | 58 | 2min |
| TRIM 41 | CGGAGTCACAAGCATCACAG | GGCCAGCAAGCTCTCATAAC | 1203 | 62 | 2min |
| BTN1_1 | AGAGCGTCACACCAAAATAGG | GGGAAGTCTTGGTTGTGTTTC | 534 | 56 | 1min |
| BTN1_2 | CCCATTGACATGAAGTGACC | AAACAGGTGGCCAAGAACAC | 924 | 60 | 1min |
| Blec1 | CACTGCTGTTGCTTTTGCAG | GTTTGGTGCAAACCCAATTC | 1419 | 62 | 2min |
| MNT-482 | GGTGGTATTCCTGTGGGATG | GGTGGACTATGGGATGATGC | 180 | 62 | 30s |
| TAP1 | GCCAGATACCACAGCAGGAG | GGCCGTGCCCTACTACAC | 458 | 60 | 30s |
| TAP2indel | CCTACATTCTGCGCCTGTC | CACCAGGCTGGAGAAGAGC | 1078 | 58 | 1min |
| C4 | GCTGTGGGTCACATGCAG | TCTCACCTTCCAGTCCTTCC | 977 | 58 | 1min |
| C4 3` | TGAGATAGGGGCTTTTGTGG | TCACCAGGTAGATGTTGTGGAG | 470 | 58 | 30s |
| CD1.1 | GCTGAGCTGATGAGCAACAC | GAGCAGAGGATAGCCGTGAG | 576 | 56 | 30s |

**Table 2:** Polymorphisms identified in commercial turkeys. Position is given with respect to DQ993255 and EU52267 (CD1.1). Under Gene, I, E/N, E/S, and N indicate intron, exon: non-synonomous, exon: synonomous, or non-codaing, respectively. Alleles use the genetic ambiguity code (R= A or G, Y= C or T, S= G or C, K= G or T, M= A or C, W= A or T) or the inserted based, where "**:**" denotes a deletion. Allele frequencies are limited to SNPs.

| | Locus | TRIM 7.2 | | | | | | | | | | | | Bzfp2 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 1462 | 1501 | 1520 | 1568 | 1620 | 1723 | 1748 | 1817 | 1957 | 2119 | 2136 | 2165 | 7579 | 7582 | 7600 | 7655 | 7724 | 7764 | 8077 | 8250 | 8292 | 8361 |
| | Gene | I | E/S | E/S | E/S | I | I | I | I | I | I | I | E/S | N | N | N | N | N | N | I | I | I | E/S |
| | Alleles | R | R | R | R | Y | R | R | R | S | Y | G | R | R | M | Y | R | Y | R | R | Y | M | Y |
| UMN/NTBF F1 | 7491 | G | G | G | G | C | C | G | A | C | C | : | G | ? | C | C | G | T | ? | ? | T | C | C |
| | | G | G | G | G | C | C | G | A | C | C | : | G | ? | C | C | G | T | ? | ? | ? | ? | C |
| | 3804 | G | G | G | G | T | C | G | A | C | C | : | G | A | C | C | G | T | A | A | T | C | ? |
| | | G | A | G | G | C | C | G | A | C | C | : | G | G | C | C | G | T | G | G | T | C | ? |
| | 1057 | G | G | G | G | C | T | G | A | C | C | : | ? | ? | C | C | G | T | A | A | ? | ? | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | ? | ? | C | C | G | T | G | G | ? | ? | ? |
| | 1049 | G | G | G | G | C | C | G | G | C | T | G | G | G | C | C | G | T | G | A | C | C | ? |
| | | G | G | G | G | C | C | G | A | C | C | : | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 1044 | G | A | G | G | T | T | G | A | C | C | : | G | G | C | C | G | T | G | A | T | C | C |
| | | G | A | G | G | C | C | A | A | C | C | : | G | ? | C | C | G | T | ? | ? | ? | ? | C |
| | 1042 | G | A | G | G | C | T | G | A | G | C | : | ? | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | C | A | A | C | C | G | ? | G | C | C | G | T | G | A | T | C | ? |
| UMN/NTBF P | 4991 | G | A | G | G | T | C | G | A | C | C | : | G | G | C | C | G | T | G | A | T | C | C |
| | | G | A | G | G | T | C | G | A | C | C | : | G | G | C | C | G | T | G | A | T | C | C |
| | 4974 | G | G | G | G | T | C | G | A | C | C | : | G | A | C | C | G | T | A | A | T | C | C |
| | | G | A | G | G | C | C | G | A | C | C | : | G | G | C | C | G | T | G | G | T | C | C |
| | 6002 | G | G | G | G | C | C | A | A | C | C | : | G | A | C | C | G | T | A | A | T | C | C |
| | | G | A | G | G | C | T | G | A | C | C | : | G | G | C | C | G | T | G | G | C | A | C |
| | 6013 | G | G | G | G | C | C | G | A | C | C | ? | G | A | C | C | G | T | A | A | T | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | C | C | G | T | G | G | C | A | C |
| | 4973 | G | A | G | G | T | C | A | A | C | C | : | G | G | C | C | G | T | G | A | T | C | C |
| | | G | A | G | G | C | T | G | A | C | C | : | G | G | C | C | G | T | G | A | C | A | C |
| | 6014 | G | G | G | G | C | C | A | A | C | C | : | G | A | C | C | G | T | A | A | T | C | ? |
| | | G | A | G | G | C | T | G | A | C | C | : | G | G | C | C | G | T | G | G | C | A | ? |
| | 5017 | G | A | G | G | T | C | G | A | G | C | G | G | G | C | C | G | T | G | A | T | C | C |
| | | G | A | G | G | C | C | G | A | C | C | : | ? | G | C | C | G | T | G | A | T | C | C |
| | 5025 | G | A | G | G | T | C | G | A | G | C | G | G | G | C | C | G | T | G | A | C | C | C |
| | | G | A | G | G | C | C | G | A | C | ? | : | G | G | C | C | G | T | G | A | T | C | C |
| | 6012 | G | G | G | G | C | C | G | A | G | T | G | G | G | C | C | G | T | G | A | C | C | ? |
| | | G | A | G | G | C | C | G | G | C | C | G | G | G | C | C | G | T | G | A | C | C | ? |
| | 4914 | G | G | G | G | C | C | G | A | G | C | G | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | G | A | G | G | C | C | G | A | C | C | : | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| NTE P and F1 | F5348 | G | A | G | G | C | T | G | A | C | C | : | G | A | C | C | G | T | A | A | C | A | ? |
| | | G | G | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | G | T | C | ? |
| | D5070 | G | A | G | G | C | T | A | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | T | A | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | D5055 | G | A | G | G | C | T | G | A | G | C | G | ? | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | ? | G | C | C | G | T | G | A | T | C | ? |
| | G1148 | G | A | G | G | C | T | G | A | G | C | G | ? | G | C | C | G | T | G | A | C | ? | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | ? | G | C | C | G | T | G | A | T | ? | ? |
| | G4600 | G | A | G | G | C | C | G | A | C | C | : | G | G | C | C | G | T | G | A | C | ? | ? |
| | | G | A | G | G | T | T | A | A | C | C | : | G | G | C | C | G | T | G | A | T | ? | ? |
| BAC | NTE | A | G | A | A | C | C | G | A | G | C | : | A | G | C | T | A | C | G | G | C | C | T |
| | WF06 | A | G | A | A | C | C | G | A | G | C | : | A | G | C | T | A | C | G | G | C | C | T |
| Commercil Breeder Line 1 | S1 | G | G | G | G | T | C | G | A | C | C | : | G | A | C | C | G | T | A | A | T | C | ? |
| | | G | A | G | G | C | C | G | A | C | C | : | G | G | C | C | G | T | G | G | T | C | ? |
| | S2 | G | A | G | G | T | C | G | A | G | C | G | ? | G | C | C | G | T | G | A | T | C | ? |
| | | G | A | G | G | C | C | G | A | C | C | : | ? | G | C | C | G | T | G | A | T | C | ? |
| | S3 | G | A | G | G | C | T | A | A | C | C | : | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | G | A | G | G | C | T | A | A | C | C | : | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | S4 | G | A | G | G | T | C | G | A | G | C | G | ? | G | C | C | G | T | G | A | T | C | ? |
| | | G | A | G | G | C | C | G | A | C | C | : | ? | G | C | C | G | T | G | A | T | C | ? |
| | S5 | G | A | G | G | T | C | G | G | C | T | G | ? | G | C | C | G | T | G | A | C | C | ? |
| | | G | G | G | G | C | C | G | A | C | C | : | ? | G | C | C | G | T | G | A | T | C | ? |
| | S6 | G | A | G | G | C | C | G | A | G | C | G | G | G | C | C | G | T | G | A | T | C | ? |
| | | G | A | G | G | C | C | G | A | G | C | G | G | G | C | C | G | T | G | A | T | C | ? |
| Commercil Breeder Line 2 | S7 | G | A | G | G | C | T | A | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | T | A | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | S8 | G | A | G | G | C | T | A | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | T | A | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | S9 | G | A | G | G | T | C | G | A | G | C | : | G | G | C | C | A | C | G | A | C | C | ? |
| | | A | G | A | A | C | C | G | A | C | C | : | A | G | A | T | G | T | G | G | T | C | ? |
| | S10 | G | A | G | G | T | C | G | A | G | C | : | G | G | C | C | A | C | G | A | C | C | ? |
| | | A | G | A | A | C | C | G | A | C | C | : | A | G | A | T | G | T | G | G | T | C | ? |
| | S11 | G | A | G | G | C | T | A | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | T | A | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | S12 | G | A | G | G | T | C | G | A | G | C | : | ? | G | A | C | A | C | G | A | C | C | ? |
| | | G | G | G | G | C | C | G | A | C | C | : | ? | G | C | T | G | T | G | G | T | C | ? |
| Commercil Breeder Line 3 | S13 | G | A | G | G | C | T | G | A | G | C | G | ? | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | ? | G | C | C | G | T | G | A | T | C | ? |
| | S14 | G | A | G | G | C | T | G | A | C | C | : | G | G | C | C | G | T | A | A | C | A | ? |
| | | G | G | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | G | T | C | ? |
| | S15 | G | A | G | G | C | T | G | A | G | C | G | ? | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | ? | A | C | C | G | T | G | A | T | C | ? |
| | S16 | G | A | G | G | C | T | G | A | C | C | : | G | G | C | C | G | T | A | A | C | ? | ? |
| | | G | G | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | G | T | ? | ? |
| | S17 | G | A | G | G | C | T | G | A | C | C | : | G | A | C | C | G | T | A | A | C | A | ? |
| | | G | G | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | G | T | C | ? |
| | S18 | G | A | G | G | C | C | G | A | G | C | G | ? | A | C | C | G | T | A | A | T | C | ? |
| | | G | G | G | G | C | C | G | A | C | C | : | ? | G | C | C | G | T | G | G | T | C | ? |
| Commercil Breeder Line 4 | S19 | G | A | G | G | T | T | G | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | A | T | C | ? |
| | S20 | G | A | G | G | T | T | G | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | A | T | C | ? |
| | S21 | G | A | G | G | T | C | G | G | C | T | G | ? | G | C | C | G | T | G | A | C | C | ? |
| | | G | G | G | G | C | C | G | A | C | C | : | ? | G | C | C | G | T | G | A | T | C | ? |
| | S22 | G | A | G | G | C | T | G | A | C | C | : | G | A | C | C | G | T | A | A | C | ? | ? |
| | | G | G | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | G | T | ? | ? |
| | S23 | G | A | G | G | C | T | G | G | C | T | G | ? | G | C | C | G | T | G | A | C | A | ? |
| | | G | G | G | G | C | C | A | A | C | C | : | ? | G | C | C | G | T | G | A | C | C | ? |
| | S24 | G | A | G | G | C | T | G | A | C | C | : | G | A | C | C | G | T | A | A | C | A | ? |
| | | G | G | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | G | T | C | ? |
| Commercil Breeder Line 5 | S25 | G | A | G | G | C | C | G | A | C | C | G | ? | A | C | C | G | T | A | A | T | C | ? |
| | | G | G | G | G | C | C | G | A | C | C | : | ? | G | C | C | G | T | G | G | T | C | ? |
| | S26 | G | A | G | G | T | T | G | A | C | C | : | G | G | C | C | G | T | G | A | C | A | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | A | T | C | ? |
| | S27 | G | A | G | G | C | C | G | A | G | C | G | G | G | C | C | G | T | G | A | C | C | ? |
| | | G | A | G | G | C | C | G | A | G | C | G | G | G | C | C | G | T | G | A | C | C | ? |
| | S28 | G | A | G | G | C | C | G | A | G | C | G | G | G | C | C | G | T | G | A | C | C | ? |
| | | G | A | G | G | C | C | G | A | G | C | G | G | G | C | C | G | T | G | A | C | C | ? |
| | S29 | G | A | G | G | T | C | G | A | C | C | : | ? | A | C | C | G | T | A | A | T | C | ? |
| | | G | G | G | G | C | C | G | A | C | C | : | ? | G | C | C | G | T | G | G | T | C | ? |
| | S30 | G | A | G | G | T | T | G | A | C | C | : | G | G | C | C | G | T | G | A | C | ? | ? |
| | | G | A | G | G | C | C | A | A | C | C | : | G | G | C | C | G | T | G | A | T | ? | ? |
| Allele Frequency | A | 0.04 | 0.71 | 0.04 | 0.04 | 0.00 | 0.00 | 0.31 | 0.95 | 0.00 | 0.00 | | 0.06 | 0.15 | 0.03 | 0.00 | 0.05 | 0.00 | 0.17 | 0.78 | 0.00 | 0.28 | 0.00 |
| | C | 0.00 | 0.00 | 0.00 | 0.00 | 0.79 | 0.69 | 0.00 | 0.00 | 0.78 | 0.95 | | 0.00 | 0.00 | 0.97 | 0.95 | 0.00 | 0.05 | 0.00 | 0.00 | 0.47 | 0.72 | 0.90 |
| | G | 0.96 | 0.29 | 0.96 | 0.96 | 0.00 | 0.00 | 0.69 | 0.05 | 0.22 | 0.00 | | 0.94 | 0.85 | 0.00 | 0.00 | 0.95 | 0.00 | 0.83 | 0.22 | 0.00 | 0.00 | 0.00 |
| | T | 0.00 | 0.00 | 0.00 | 0.00 | 0.21 | 0.31 | 0.00 | 0.00 | 0.00 | 0.05 | | 0.00 | 0.00 | 0.00 | 0.05 | 0.00 | 0.95 | 0.00 | 0.00 | 0.53 | 0.00 | 0.10 |

| | Locus | 20K | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 19507 | 19573 | 19604 | 19645 | 19712 | 19727 | 19737 | 19789 | 19844 | 19853 | 19858 | 19938 | 20020 | 20021 | 20105 | 20180 |
| | Gene | N | N | N | N | N | N | N | N | N | N | tRNA | N | N | N | N | N |
| | Alleles | R | Y | M | Y | S | R | Y | Y | S | Y | Y | Y | Y | R | W | Y |
| UMN/NTBF F1 | 7491 | G | C | C | C | G | A | T | C | G | A | C | C | C | A | T | C |
| | | G | C | C | C | G | A | T | C | G | A | C | C | C | A | T | C |
| | 3804 | ? | ? | ? | C | G | A | T | C | G | C | C | T | C | A | T | C |
| | | ? | ? | ? | T | G | G | C | C | G | A | C | T | C | A | T | C |
| | 1057 | G | C | C | C | G | A | T | C | G | C | C | C | C | A | T | C |
| | | G | C | C | C | G | A | T | T | G | A | C | T | C | A | T | C |
| | 1049 | G | C | C | C | C | A | C | C | G | C | C | C | C | A | T | C |
| | | G | T | C | C | C | G | T | C | G | A | T | T | C | A | T | C |
| | 1044 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 1042 | G | C | C | C | G | A | T | C | G | C | C | T | C | A | T | C |
| | | G | T | C | T | G | G | T | T | C | C | C | T | C | A | T | C |
| UMN/NTBF P | 4991 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 4974 | G | C | C | C | G | A | T | C | G | A | C | C | C | A | T | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6002 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6013 | G | C | C | C | G | A | T | C | G | A | C | C | C | A | T | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 4973 | G | C | C | C | G | A | T | ? | G | ? | C | ? | C | A | T | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6014 | G | ? | C | ? | G | ? | T | C | G | C | C | T | C | A | T | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 5017 | G | ? | C | ? | G | ? | T | ? | ? | C | C | T | C | A | T | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 5025 | ? | ? | ? | ? | G | ? | ? | C | G | ? | C | T | C | A | T | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6012 | G | ? | C | C | C | ? | ? | C | G | ? | ? | ? | C | A | T | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 4914 | G | ? | C | C | C | ? | ? | C | G | ? | ? | ? | ? | A | T | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| NTE P and F1 | F5348 | G | C | C | C | G | A | T | C | G | C | C | C | C | A | T | C |
| | | A | C | C | C | G | A | T | T | G | A | C | T | C | A | T | C |
| | D5070 | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | D5055 | G | C | C | C | G | A | C | C | G | C | C | T | C | A | T | C |
| | | A | T | C | T | G | G | T | T | C | C | C | T | C | A | T | C |
| | G1148 | G | C | C | C | G | A | C | C | G | C | C | T | C | A | T | C |
| | | A | T | C | T | G | G | T | T | C | C | C | T | C | A | T | C |
| | G4600 | G | C | C | C | G | A | C | C | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | G | T | T | G | C | C | T | C | A | T | C |
| BAC | NTE | G | C | A | C | C | G | T | C | G | C | C | T | T | G | A | T |
| | WF06 | G | C | A | C | C | G | T | C | G | C | C | T | T | G | A | T |
| Commercil Breeder Line 1 | S1 | G | C | C | C | G | A | C | C | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | G | T | T | G | C | C | T | C | A | T | C |
| | S2 | G | C | C | C | G | G | C | C | G | C | C | T | C | A | T | C |
| | | G | T | C | T | G | G | C | C | C | C | C | T | C | A | T | C |
| | S3 | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | | A | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | S4 | G | C | C | C | G | G | C | C | G | C | C | T | C | A | T | C |
| | | G | T | C | T | G | G | C | C | C | C | C | T | C | A | T | C |
| | S5 | G | C | C | C | G | G | C | C | G | C | C | T | C | A | T | C |
| | | G | T | C | C | C | G | C | C | G | C | T | T | C | A | T | C |
| | S6 | G | T | C | T | G | G | C | C | C | C | C | T | C | A | T | C |
| | | G | T | C | T | G | G | C | C | C | C | C | T | C | A | T | C |
| Commercil Breeder Line 2 | S7 | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | S8 | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | S9 | G | C | A | C | C | G | T | C | G | C | C | T | C | A | A | T |
| | | G | C | A | C | C | G | T | C | G | C | C | T | C | A | A | T |
| | S10 | G | C | A | C | C | G | T | C | G | C | C | T | C | A | A | T |
| | | G | C | A | C | C | G | T | C | G | C | C | T | C | A | A | T |
| | S11 | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | | A | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | S12 | G | C | C | C | C | G | T | C | G | C | C | T | C | A | A | C |
| | | G | C | A | C | G | G | C | C | G | C | C | T | C | A | T | T |
| Commercil Breeder Line 3 | S13 | ? | ? | ? | C | G | A | T | C | G | C | C | T | C | A | T | C |
| | | ? | ? | ? | T | G | G | C | C | G | C | C | T | C | A | T | C |
| | S14 | G | C | ? | C | G | A | T | T | G | C | C | C | C | A | T | C |
| | | G | C | C | C | G | A | T | C | G | A | C | T | C | A | T | C |
| | S15 | G | C | C | C | G | A | C | T | G | C | C | T | C | A | T | C |
| | | G | T | C | T | G | G | T | C | C | C | C | T | C | A | T | C |
| | S16 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | S17 | G | C | C | C | G | A | C | T | G | C | C | C | C | A | T | C |
| | | G | C | C | C | G | A | T | C | G | A | C | T | C | A | T | C |
| | S18 | G | C | C | C | G | G | C | C | G | C | C | C | C | A | T | C |
| | | G | T | C | T | G | A | T | C | C | A | C | T | C | A | T | C |
| Commercil Breeder Line 4 | S19 | ? | C | C | C | G | G | C | T | G | C | C | T | C | A | T | C |
| | | ? | C | C | T | G | A | T | C | C | C | C | T | C | A | T | C |
| | S20 | G | C | C | C | G | G | C | T | G | C | C | T | C | A | T | C |
| | | G | T | C | T | G | A | T | C | C | C | C | T | C | A | T | C |
| | S21 | G | C | C | C | G | G | C | C | G | C | C | T | C | A | T | C |
| | | G | T | C | C | C | G | C | C | G | C | T | T | C | A | T | C |
| | S22 | G | C | C | C | G | A | T | T | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | A | T | C | G | A | C | C | C | A | T | C |
| | S23 | G | C | C | C | G | G | C | T | G | C | C | T | C | A | T | C |
| | | G | T | C | C | G | A | T | C | G | C | C | T | C | A | T | C |
| | S24 | G | C | C | C | C | A | T | T | G | C | C | C | C | A | T | C |
| | | G | C | C | C | G | A | T | C | G | A | C | T | C | A | T | C |
| Commercil Breeder Line 5 | S25 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | S26 | G | C | C | C | G | A | C | T | C | C | C | T | C | A | T | C |
| | | G | C | C | T | G | G | T | C | G | C | C | T | C | A | T | C |
| | S27 | G | C | C | C | G | G | C | C | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | G | C | C | G | C | C | T | C | A | T | C |
| | S28 | G | C | C | C | G | G | C | C | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | G | C | C | G | C | C | T | C | A | T | C |
| | S29 | G | C | C | C | G | G | T | C | G | C | C | T | C | A | T | C |
| | | G | T | C | T | G | A | C | C | G | A | C | C | C | A | T | C |
| | S30 | G | C | C | C | G | G | T | T | G | C | C | T | C | A | T | C |
| | | G | C | C | C | G | A | C | C | G | C | C | T | C | A | T | C |
| Allele Frequency | A | 0.06 | 0.00 | 0.09 | 0.00 | 0.00 | 0.53 | 0.00 | 0.00 | 0.00 | 0.17 | 0.00 | 0.00 | 0.00 | 0.98 | 0.08 | 0.00 |
| | C | 0.00 | 0.81 | 0.91 | 0.82 | 0.16 | 0.00 | 0.37 | 0.68 | 0.15 | 0.83 | 0.96 | 0.16 | 0.98 | 0.00 | 0.00 | 0.92 |
| | G | 0.94 | 0.00 | 0.00 | 0.00 | 0.84 | 0.47 | 0.00 | 0.00 | 0.85 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 |
| | T | 0.00 | 0.19 | 0.00 | 0.18 | 0.00 | 0.00 | 0.63 | 0.32 | 0.00 | 0.00 | 0.04 | 0.84 | 0.02 | 0.00 | 0.92 | 0.08 |

| | Locus | LAAO | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 28905 | 28942 | 28964 | 29041 | 29218 | 29219 | 29239 | 29323 | 29347 | 29359 | 29440 | 29449 |
| | Gene | I | I | I | E/S | E/N | E/N | E/S | E/S | E/S | E/S | E/S | E/S |
| | Alleles | Y | M | W | Y | Y | R | R | Y | Y | Y | Y | Y |
| UMN/NTBF F1 | 7491 | C | A | T | C | T | G | G | C | T | C | C | T |
| | | C | A | T | C | T | G | G | C | T | C | C | T |
| | 3804 | C | A | T | C | C | G | A | C | T | C | C | T |
| | | C | A | T | C | T | G | G | C | T | C | C | T |
| | 1057 | C | A | T | C | C | G | A | C | T | C | C | T |
| | | C | A | T | T | T | G | G | C | T | T | C | T |
| | 1049 | C | A | T | T | C | G | A | C | T | T | C | T |
| | | C | A | T | C | T | G | G | C | T | C | C | T |
| | 1044 | C | A | T | C | C | G | A | C | T | C | C | T |
| | | C | A | T | C | C | G | A | C | T | C | C | T |
| | 1042 | C | A | T | C | C | G | A | C | T | C | C | T |
| | | C | A | T | T | C | G | A | C | T | T | C | T |
| UMN/NTBF P | 4991 | C | A | T | C | C | G | A | C | T | C | C | T |
| | | C | A | T | C | C | G | A | C | T | C | C | T |
| | 4974 | C | A | T | C | C | G | G | C | T | C | C | T |
| | | C | A | T | C | T | G | A | C | T | C | C | T |
| | 6002 | C | A | T | C | C | G | G | C | T | C | C | T |
| | | C | A | T | C | T | G | A | C | T | T | C | T |
| | 6013 | C | A | T | C | T | G | G | C | T | C | C | T |
| | | C | A | T | T | T | G | A | C | T | T | C | T |
| | 4973 | C | A | T | C | C | G | A | C | T | C | C | T |
| | | C | A | T | T | C | G | A | C | T | T | C | T |
| | 6014 | C | A | T | C | T | G | A | C | T | C | C | T |
| | | C | A | T | T | C | G | G | C | T | T | C | T |
| | 5017 | C | A | T | C | C | G | A | C | T | C | C | T |
| | | C | A | T | C | C | G | A | C | T | C | C | T |
| | 5025 | T | A | T | C | C | G | A | C | T | C | C | T |
| | | C | C | T | C | C | G | G | C | T | T | C | T |
| | 6012 | C | A | T | T | C | G | A | C | T | T | C | T |
| | | C | A | T | T | C | G | A | C | T | T | C | T |
| | 4914 | C | A | T | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| NTE P and F1 | F5348 | C | A | T | C | T | G | A | C | T | C | C | C |
| | | C | A | T | T | C | G | G | C | T | T | C | C |
| | D5070 | C | A | T | T | C | G | A | C | T | T | C | C |
| | | C | A | T | T | C | G | A | C | T | T | C | C |
| | D5055 | C | A | T | C | C | G | A | C | T | C | C | C |
| | | C | A | T | T | C | G | A | C | T | T | C | C |
| | G1148 | C | A | T | C | C | G | A | C | T | C | C | C |
| | | C | A | T | T | C | G | A | C | T | T | C | C |
| | G4600 | T | A | T | T | C | G | A | C | T | T | C | C |
| | | T | C | T | C | C | G | G | C | T | T | C | C |
| BAC | NTE | C | A | A | T | C | G | G | T | T | T | T | T |
| | WF06 | C | A | A | T | C | G | G | T | T | T | T | T |
| Commercil Breeder Line 1 | S1 | T | A | T | C | C | G | A | C | T | T | C | C |
| | | T | C | T | T | C | G | G | C | T | T | C | C |
| | S2 | C | A | T | C | C | G | A | C | T | C | C | C |
| | | T | C | T | C | C | G | G | C | T | T | C | C |
| | S3 | C | A | T | T | C | G | A | C | T | T | C | C |
| | | C | A | T | T | C | G | A | C | T | T | C | C |
| | S4 | T | A | T | C | C | G | A | C | T | C | C | C |
| | | T | C | T | C | C | G | G | C | T | T | C | C |
| | S5 | C | A | T | T | C | G | A | C | T | T | C | C |
| | | T | C | T | C | C | G | G | C | T | T | C | C |
| | S6 | C | A | T | C | C | G | A | C | T | C | C | C |
| | | C | A | T | C | C | G | A | C | T | C | C | C |
| Commercil Breeder Line 2 | S7 | C | A | T | T | C | G | A | C | T | T | C | C |
| | | C | A | T | T | C | G | A | C | T | T | C | C |
| | S8 | C | A | T | T | C | G | A | C | T | T | C | C |
| | | C | A | T | T | C | G | A | C | T | T | C | C |
| | S9 | T | A | A | C | C | G | G | C | T | T | C | C |
| | | C | C | T | T | C | G | G | T | T | T | T | T |
| | S10 | T | A | A | C | C | G | G | C | T | T | C | C |
| | | C | C | T | T | C | G | G | T | T | T | T | T |
| | S11 | C | A | T | T | C | G | A | C | T | T | C | C |
| | | C | A | T | T | C | G | A | C | T | T | C | C |
| | S12 | T | A | T | C | C | G | G | C | C | C | C | C |
| | | C | C | T | C | C | A | G | C | T | T | C | T |
| Commercil Breeder Line 3 | S13 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | C | G | A | C | T | T | C | C |
| | S14 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | T | G | G | C | T | T | C | C |
| | S15 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | C | G | A | C | T | T | C | C |
| | S16 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | T | G | G | C | T | T | C | C |
| | S17 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | T | G | G | C | T | T | C | C |
| | S18 | C | A | T | C | C | G | A | C | T | C | C | C |
| | | C | A | T | C | T | G | G | C | T | C | C | C |
| Commercil Breeder Line 4 | S19 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | C | G | A | C | T | T | C | C |
| | S20 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | C | G | A | C | T | T | C | C |
| | S21 | C | A | T | T | C | G | A | C | T | T | C | C |
| | | T | C | T | C | C | G | G | C | T | T | C | C |
| | S22 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | T | G | G | C | T | T | C | C |
| | S23 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | C | G | G | C | T | T | C | C |
| | S24 | C | A | T | T | C | G | A | C | T | C | C | C |
| | | C | A | T | C | T | G | G | C | T | T | C | C |
| Commercil Breeder Line 5 | S25 | C | A | T | C | C | G | A | C | T | C | C | C |
| | | C | A | T | C | T | G | G | C | T | C | C | C |
| | S26 | C | A | T | C | C | G | A | C | T | C | C | C |
| | | C | A | T | T | C | G | A | C | T | T | C | C |
| | S27 | T | C | T | C | C | G | G | C | T | T | C | C |
| | | T | C | T | C | C | G | G | C | T | T | C | C |
| | S28 | T | C | T | C | C | G | G | C | T | T | C | C |
| | | T | C | T | C | C | G | G | C | T | T | C | C |
| | S29 | C | A | T | C | T | G | A | C | T | C | C | C |
| | | C | A | T | C | C | G | G | C | T | C | C | C |
| | S30 | T | A | T | T | C | G | A | C | T | T | C | C |
| | | T | C | T | C | C | G | G | C | T | T | C | C |
| Allele Frequency | A | 0.00 | 0.85 | 0.04 | 0.00 | 0.00 | 0.01 | 0.62 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | C | 0.82 | 0.15 | 0.00 | 0.60 | 0.81 | 0.00 | 0.00 | 0.96 | 0.01 | 0.44 | 0.96 | 0.66 |
| | G | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.99 | 0.38 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | T | 0.18 | 0.00 | 0.96 | 0.40 | 0.19 | 0.00 | 0.00 | 0.04 | 0.99 | 0.56 | 0.04 | 0.34 |

| Group | Locus | TRIM 7.1 | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 42827 | 42828 | 42832 | 42866 | 42924 | 43066 | 43216 | 43301 | 43331 | 43826 | 43879 | 43893 | 43959 | 43990 | 44005 | 44014 | 44018 | 44054 | 44170 | 44198 |
| | Gene | E/N | E/N | E/S | I | I | I | I | I | I | I | I | I | I | I | I | I | I | E/S | E/N | E/S |
| | Alleles | Y | R | Y | W | M | AA | R | Y | Y | ms | Y | Y | R | R | Y | R | ms | Y | Y | Y |
| UMN/NTBF F1 | 7491 | C | G | C | T | C | :: | A | C | T | 5 | C | C | G | G | C | G | 4 | T | C | C |
| | | C | G | C | T | C | :: | A | C | T | 5 | C | C | G | G | C | G | 4 | T | C | C |
| | 3804 | C | G | C | T | C | :: | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | C | C |
| | 1057 | C | G | C | T | C | :: | ? | ? | ? | 5 | C | C | A | G | C | G | 4 | C | C | C |
| | | T | G | T | T | C | AA | ? | ? | ? | 5 | C | C | G | G | T | G | 4 | T | C | C |
| | 1049 | T | G | C | T | C | :: | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | T | T | C | AA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | C | C |
| | 1044 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 1042 | C | G | C | T | C | :: | ? | ? | ? | 4 | C | C | G | G | C | G | 3 | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | 4 | C | C | G | G | C | G | 3 | C | C | C |
| UMN/NTBF P | 4991 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 4974 | C | G | C | T | C | :: | A | C | T | 5 | C | C | G | G | C | G | 4 | T | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6002 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6013 | C | G | C | T | C | :: | A | C | T | 5 | C | C | G | G | C | G | 4 | T | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 4973 | ? | G | ? | T | C | ? | ? | ? | ? | ? | C | C | ? | G | ? | G | 4 | ? | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6014 | C | G | C | T | C | ? | ? | ? | ? | ? | C | C | G | G | C | G | ? | C | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 5017 | C | G | C | T | C | ? | ? | ? | ? | ? | C | C | G | G | C | G | ? | C | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 5025 | C | G | C | T | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6012 | ? | G | ? | T | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 4914 | ? | G | ? | T | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| NTE P and F1 | F5348 | C | G | T | T | C | :: | ? | ? | ? | 5 | C | C | A | G | C | G | 4 | C | C | C |
| | | T | G | C | T | C | AA | ? | ? | ? | 5 | C | C | G | G | T | G | 4 | T | C | C |
| | D5070 | T | G | T | T | C | :: | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | D5055 | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | G1148 | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | G4600 | T | G | T | T | C | AA | A | T | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| BAC | NTE | C | A | C | A | C | AA | G | C | T | 4 | T | T | G | A | C | G | 5 | T | T | T |
| | WF06 | C | A | C | A | C | AA | G | C | T | 4 | T | T | G | A | C | G | 5 | T | T | T |
| Commercil Breeder Line 1 | S1 | T | G | T | T | C | AA | A | T | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | S2 | C | G | C | T | C | AA | A | C | T | 4 | C | C | G | G | C | A | 3 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | G | 3 | C | C | C |
| | S3 | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | S4 | T | G | C | T | C | AA | A | C | T | 4 | C | C | G | G | C | A | 3 | C | C | C |
| | | C | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | G | 3 | C | C | C |
| | S5 | T | G | C | T | C | AA | A | C | T | 4 | C | C | G | G | C | A | 3 | C | C | C |
| | | C | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | G | 3 | C | C | C |
| | S6 | C | G | C | T | C | AA | A | C | T | 4 | C | C | G | G | C | G | 3 | C | C | C |
| | | C | G | C | T | C | AA | A | C | T | 4 | C | C | G | G | C | G | 3 | C | C | C |
| Commercil Breeder Line 2 | S7 | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | S8 | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | S9 | C | G | C | T | C | AA | A | C | T | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | T | G | T | A | A | AA | G | T | C | ? | ? | ? | ? | ? | ? | ? | ? | T | T | T |
| | S10 | C | G | C | T | C | AA | A | C | T | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | T | A | T | A | C | AA | G | T | C | ? | ? | ? | ? | ? | ? | ? | ? | T | T | T |
| | S11 | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | S12 | T | G | T | A | C | AA | A | C | T | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | G | T | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| Commercil Breeder Line 3 | S13 | T | G | T | T | C | AA | A | T | T | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | A | C | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | S14 | T | G | T | T | C | :: | ? | ? | ? | 5 | C | C | A | G | C | G | 4 | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | 5 | C | C | G | G | T | G | 4 | T | C | C |
| | S15 | T | G | T | T | C | AA | A | T | T | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | A | C | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | S16 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | S17 | T | G | T | T | C | :: | ? | ? | ? | 5 | C | C | A | G | C | G | 4 | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | 5 | C | C | G | G | T | G | 4 | T | C | C |
| | S18 | C | G | C | T | C | :: | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | C | C |
| Commercil Breeder Line 4 | S19 | T | G | T | T | C | AA | A | T | T | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | A | C | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | S20 | T | G | T | T | C | AA | A | T | T | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | A | C | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | S21 | T | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | A | 3 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | G | 3 | C | C | C |
| | S22 | T | G | T | T | C | :: | ? | ? | ? | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | 5 | C | C | G | G | C | G | 4 | T | C | C |
| | S23 | T | G | T | T | C | AA | A | T | T | 4 | C | C | A | G | T | G | 4 | C | C | C |
| | | C | G | C | T | C | AA | A | C | C | 5 | C | C | A | G | T | G | 4 | C | C | C |
| | S24 | T | G | T | T | C | :: | ? | ? | ? | 5 | C | C | A | G | C | G | 4 | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | 5 | C | C | G | G | T | G | 4 | T | C | C |
| Commercil Breeder Line 5 | S25 | C | G | C | T | C | :: | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | C | C |
| | S26 | C | G | T | T | C | AA | A | T | T | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | T | G | C | T | C | AA | A | C | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | S27 | T | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | A | 3 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | A | 3 | C | C | C |
| | S28 | T | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | A | 3 | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | 5 | C | C | G | G | C | A | 3 | C | C | C |
| | S29 | C | G | C | T | C | :: | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | C | G | C | T | C | AA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | C | C |
| | S30 | T | G | T | T | C | AA | A | T | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| | | T | G | T | T | C | AA | A | T | C | ? | ? | ? | ? | ? | ? | ? | ? | C | C | C |
| Allele Frequency | A | 0.00 | 0.03 | 0.00 | 0.06 | 0.01 | | 0.91 | 0.00 | 0.00 | | 0.00 | 0.00 | 0.42 | 0.04 | 0.00 | 0.15 | | 0.00 | 0.00 | 0.00 |
| | C | 0.47 | 0.00 | 0.47 | 0.00 | 0.99 | | 0.00 | 0.34 | 0.66 | | 0.96 | 0.96 | 0.00 | 0.00 | 0.58 | 0.00 | | 0.77 | 0.95 | 0.95 |
| | G | 0.00 | 0.97 | 0.00 | 0.00 | 0.00 | | 0.09 | 0.00 | 0.00 | | 0.00 | 0.00 | 0.58 | 0.96 | 0.00 | 0.85 | | 0.00 | 0.00 | 0.00 |
| | T | 0.53 | 0.00 | 0.53 | 0.94 | 0.00 | | 0.00 | 0.66 | 0.34 | | 0.04 | 0.04 | 0.00 | 0.00 | 0.42 | 0.00 | | 0.23 | 0.05 | 0.05 |

| | Locus | TRIM 39.2 | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 56210 | 56220 | 56238 | 56256 | 56416 | 56520 | 56559 | 56566 | 56624 | 56632 | 56663 | 56670 | 56676 | 56710 | 56713 | 56742 | 56806 | 56840 |
| | Gene | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I |
| | Alleles | M | W | R | Y | M | Y | R | R | Y | Y | R | R | R | R | R | R | W | R |
| UMN/NTBF F1 | 7491 | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | 3804 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | T | A | G | G | A | G | G | T | G |
| | 1057 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | T | G | G | G | A | G | A | A | A |
| | 1049 | C | A | A | C | A | C | A | G | T | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | T | G | G | G | A | G | A | A | A |
| | 1044 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | T | A | G | G | A | G | G | T | G |
| | 1042 | C | A | A | C | A | C | A | G | T | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | T | A | G | G | A | G | G | T | G |
| UMN/NTBF P | 4991 | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | 4974 | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | 6002 | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | 6013 | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | 4973 | C | A | A | C | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | C | A | A | C | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 6014 | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | | C | A | A | C | A | C | A | G | C | T | A | G | G | A | G | A | A | A |
| | 5017 | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | 5025 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | 6012 | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | 4914 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | A | G | C | T | A | G | G | A | G | G | T | G |
| NTE P and F1 | F5348 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | D5070 | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | D5055 | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | G1148 | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | G4600 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| BAC | NTE | C | T | G | C | A | C | G | A | C | C | A | G | A | A | G | G | A | G |
| | WF06 | C | T | G | C | A | C | G | A | C | C | A | G | A | A | G | G | A | G |
| Commercil Breeder Line 1 | S1 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | S2 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | S3 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | S4 | C | A | A | C | A | C | G | A | C | C | A | A | G | A | A | A | A | A |
| | | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | G | A | G |
| | S5 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | S6 | C | A | A | C | A | C | G | A | C | C | A | A | G | A | A | G | T | G |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| Commercil Breeder Line 2 | S7 | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | S8 | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | S9 | C | A | A | C | A | C | G | A | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | T | A | C | G | G | C | C | A | G | G | A | G | G | A | G |
| | S10 | C | A | A | C | A | C | G | A | C | C | A | G | A | A | G | A | A | A |
| | | C | T | G | C | A | C | G | G | C | C | A | G | G | A | G | G | A | G |
| | S11 | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | S12 | A | A | A | C | A | C | G | A | C | C | A | A | G | A | G | A | A | A |
| | | C | A | A | T | A | T | G | G | C | C | A | G | G | G | G | G | A | G |
| Commercil Breeder Line 3 | S13 | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | A | A | G |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | A | T | A |
| | S14 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | ? |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | A | A | ? |
| | S15 | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | G | T | G |
| | S16 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | S17 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | S18 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| Commercil Breeder Line 4 | S19 | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | G | T | G |
| | S20 | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | G | T | G |
| | S21 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | G | T | G |
| | S22 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | S23 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| | S24 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| Commercil Breeder Line 5 | S25 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | T | A | G | G | A | G | G | T | G |
| | S26 | C | A | A | C | C | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | G | T | G |
| | S27 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | S28 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | S29 | C | A | A | C | A | C | A | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | T | A | G | G | A | G | G | T | G |
| | S30 | C | A | A | C | A | C | G | G | C | C | A | G | G | A | G | A | A | A |
| | | C | A | A | C | C | C | G | G | C | C | G | G | G | A | G | A | A | A |
| Allele Frequency | A | 0.01 | 0.97 | 0.97 | 0.00 | 0.46 | 0.00 | 0.23 | 0.07 | 0.00 | 0.00 | 0.71 | 0.03 | 0.03 | 0.99 | 0.02 | 0.71 | 0.77 | 0.70 |
| | C | 0.99 | 0.00 | 0.00 | 0.98 | 0.54 | 0.99 | 0.00 | 0.00 | 0.98 | 0.83 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | G | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 0.77 | 0.93 | 0.00 | 0.00 | 0.29 | 0.97 | 0.97 | 0.01 | 0.98 | 0.29 | 0.00 | 0.30 |
| | T | 0.00 | 0.03 | 0.00 | 0.02 | 0.00 | 0.01 | 0.00 | 0.00 | 0.02 | 0.17 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.23 | 0.00 |

| | Locus | TRIM 39.1 | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 66845 | 66875 | 66904 | 66970 | 67001 | 67002 | 67004 | 67039 | 67049 | 67099 | 67142 | 67151 | 67333 | 67337 | 67364 | 67427 | 67447 | 67518 | 67532 | 67567 | 67737 | 67772 | 67801 |
| | Gene | I | I | I | I | I | I | I | I | I | I | I | I | E/S | E/N | E/N | I | I | I | I | I | I | I | I |
| | Alleles | R | W | Y | R | S | W | W | Y | S | Y | W | W | M | S | Y | Y | Y | Y | K | R | M | R | R |
| UMN/NTBF F1 | 7491 | A | T | C | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | C | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | 3804 | A | T | C | G | G | T | T | C | G | T | A | T | A | G | C | C | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | G | C | T | T | A | G | C | T | C | C | T | G | A | A | A |
| | 1057 | A | T | C | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | C | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | 1049 | ? | ? | C | G | G | T | T | C | G | T | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | ? | ? | T | G | G | T | T | C | G | C | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | 1044 | ? | ? | C | G | G | T | T | C | G | T | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | ? | ? | T | G | G | T | T | C | G | C | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | 1042 | A | T | C | G | G | T | T | C | G | T | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | G | C | T | T | A | G | C | C | C | C | T | G | A | A | A |
| UMN/NTBF P | 4991 | ? | ? | ? | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | | ? | ? | ? | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | 4974 | ? | T | C | G | G | T | T | C | G | C | A | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | | ? | T | T | G | G | T | T | C | G | T | T | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | 6002 | ? | T | C | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | | ? | T | C | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | 6013 | ? | ? | C | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | | ? | ? | C | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | 4973 | ? | ? | ? | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | | ? | ? | ? | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | 6014 | ? | ? | ? | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | | ? | ? | ? | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | 5017 | ? | ? | ? | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | | ? | ? | ? | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | 5025 | ? | ? | ? | G | G | T | T | C | G | C | A | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | | ? | ? | ? | G | G | T | T | C | C | T | T | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| | 6012 | ? | ? | ? | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | | ? | ? | ? | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | 4914 | ? | ? | ? | G | G | T | T | C | G | C | A | T | A | G | C | C | C | C | T | G | ? | ? | ? |
| | | ? | ? | ? | G | G | T | T | C | G | T | T | T | A | G | C | T | C | C | T | G | ? | ? | ? |
| NTE P and F1 | F5348 | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | D5070 | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | D5055 | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | G1148 | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | G4600 | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| BAC | NTE | G | A | C | G | C | T | T | T | G | C | A | A | C | G | T | C | C | T | G | G | C | A | A |
| | WF06 | G | A | C | G | C | T | T | T | G | C | A | A | C | G | T | C | C | T | G | G | C | A | A |
| Commercil Breeder Line 1 | S1 | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | S2 | A | T | T | G | G | T | T | C | G | T | A | T | A | G | C | C | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | C | T | T | A | G | C | T | C | C | T | G | A | A | A |
| | S3 | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | | G | T | C | A | G | T | T | C | G | C | A | T | A | C | C | T | C | C | T | A | A | A | G |
| | S4 | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | S5 | A | T | T | G | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | S6 | A | T | T | G | G | T | T | C | G | T | T | T | A | C | C | C | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| Commercil Breeder Line 2 | S7 | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | S8 | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | S9 | ? | ? | ? | G | G | A | T | C | G | C | A | A | A | G | C | T | C | C | G | G | A | A | A |
| | | ? | ? | ? | G | C | T | T | C | C | C | A | T | C | G | T | C | C | T | T | G | A | A | A |
| | S10 | ? | ? | ? | G | G | T | T | C | G | C | A | A | A | G | C | T | C | C | G | G | C | A | A |
| | | ? | ? | ? | G | C | T | T | T | C | C | A | T | C | G | T | C | C | T | T | G | A | A | A |
| | S11 | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | | G | T | C | A | G | T | T | C | G | C | A | T | A | G | C | T | C | C | T | A | A | A | G |
| | S12 | ? | ? | ? | G | G | T | T | C | G | C | A | A | A | G | C | T | C | C | G | G | A | A | A |
| | | ? | ? | ? | G | C | T | T | C | C | C | A | T | C | G | C | C | C | T | T | G | A | A | A |
| Commercil Breeder Line 3 | S13 | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | S14 | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | S15 | A | T | C | G | G | T | T | C | G | C | A | T | A | G | C | C | T | C | T | G | A | G | A |
| | | A | T | C | G | G | T | T | C | G | C | A | T | C | G | C | C | T | C | T | G | A | G | A |
| | S16 | ? | ? | ? | ? | ? | ? | ? | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | | ? | ? | ? | ? | ? | ? | ? | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | S17 | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | S18 | A | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | G | A |
| | | G | T | C | G | G | T | T | C | G | C | A | T | C | G | C | C | T | C | T | G | A | A | A |
| Commercil Breeder Line 4 | S19 | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | S20 | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | S21 | A | T | T | G | G | T | T | C | G | C | T | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | T | A | T | A | G | C | C | C | C | T | G | A | A | A |
| | S22 | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | S23 | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | S24 | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| | | G | T | C | A | G | T | A | C | G | C | A | A | A | G | C | T | C | C | T | G | A | A | A |
| Commercil Breeder Line 5 | S25 | A | T | T | G | G | T | T | C | G | T | A | T | A | G | C | C | C | C | T | G | A | A | A |
| | | A | T | C | G | G | T | T | C | G | C | T | T | A | G | C | T | C | C | T | G | A | A | A |
| | S26 | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | G | T | T | T | A | G | C | C | C | C | T | G | A | A | A |
| | S27 | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | S28 | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | S29 | A | T | C | G | G | T | T | C | G | T | A | T | A | G | C | C | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | G | C | T | T | A | G | C | T | C | C | T | G | A | A | A |
| | S30 | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| | | A | T | T | G | G | T | T | C | C | C | A | T | A | G | C | T | C | C | T | G | A | A | A |
| Allele Frequency | A | 0.60 | 0.03 | 0.00 | 0.27 | 0.00 | 0.01 | 0.13 | 0.00 | 0.00 | 0.00 | 0.72 | 0.19 | 0.93 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.13 | 0.96 | 0.96 | 0.83 |
| | C | 0.00 | 0.00 | 0.57 | 0.00 | 0.05 | 0.00 | 0.00 | 0.97 | 0.18 | 0.72 | 0.00 | 0.00 | 0.07 | 0.02 | 0.96 | 0.36 | 0.97 | 0.95 | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 |
| | G | 0.40 | 0.00 | 0.00 | 0.73 | 0.95 | 0.00 | 0.00 | 0.00 | 0.82 | 0.00 | 0.00 | 0.00 | 0.00 | 0.98 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.87 | 0.00 | 0.04 | 0.17 |
| | T | 0.00 | 0.97 | 0.43 | 0.00 | 0.00 | 0.99 | 0.87 | 0.03 | 0.00 | 0.28 | 0.28 | 0.81 | 0.00 | 0.00 | 0.04 | 0.64 | 0.03 | 0.05 | 0.95 | 0.00 | 0.00 | 0.00 | 0.00 |

| | Locus | TRIM 41 | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 77449 | 77469 | 77534 | 77545 | 77625 | 77661 | 77678 | 77726 | 77804 | 77806 | 77833 | 78036 | 78256 | 78283 | 78368 | 78434 |
| | Gene | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I |
| | Alleles | Y | W | R | G/: | Y | S | R | Y | Y | Y | Y | Y | Y | R | R | Y |
| UMN/NTBF F1 | 7491 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 3804 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 1057 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 1049 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 1044 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 1042 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| UMN/NTBF P | 4991 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | 4974 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 6002 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 6013 | ? | ? | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | ? | ? | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 4973 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 6014 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | 5017 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | 5025 | ? | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | ? | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| | 6012 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | 4914 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | ? | ? |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | ? | ? |
| NTE P and F1 | F5348 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | D5070 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | D5055 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | G1148 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | ? | ? |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | ? | ? |
| | G4600 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | T | A | G | C |
| BAC | NTE | T | T | G | : | T | C | A | C | C | C | T | C | T | G | G | T |
| | WF06 | T | T | G | : | T | C | A | C | C | C | T | C | T | G | G | T |
| Commercil Breeder Line 1 | S1 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | T | A | G | C |
| | S2 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| | S3 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | A | ? |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | ? |
| | S4 | C | A | A | G | C | G | G | T | T | C | C | C | T | A | G | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| | S5 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| | S6 | C | A | A | G | C | G | G | T | T | C | C | C | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| Commercil Breeder Line 2 | S7 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S8 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S9 | C | A | A | G | C | G | G | C | C | C | T | C | T | A | G | C |
| | | T | T | G | : | T | C | A | T | T | C | T | T | T | G | G | T |
| | S10 | T | A | A | G | T | G | A | C | C | C | T | C | T | A | G | C |
| | | T | T | G | : | C | C | G | T | T | C | T | T | T | G | G | T |
| | S11 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S12 | C | A | A | G | C | G | G | C | C | C | C | C | T | A | G | C |
| | | C | T | G | : | T | C | A | T | T | C | T | T | T | G | G | T |
| Commercil Breeder Line 3 | S13 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | G | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S14 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S15 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S16 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | ? |
| | | C | A | A | G | C | C | G | T | T | T | C | T | C | A | G | ? |
| | S17 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S18 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| Commercil Breeder Line 4 | S19 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S20 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S21 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| | S22 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | C | A | G | C |
| | S23 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | C | A | G | C |
| | S24 | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| Commercil Breeder Line 5 | S25 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S26 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | T | A | G | C |
| | S27 | C | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| | S28 | C | A | A | G | C | G | G | T | T | C | C | T | T | A | G | C |
| | | C | A | A | G | C | G | G | T | T | C | C | T | C | A | G | C |
| | S29 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | A | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | C | A | G | C |
| | S30 | C | A | A | G | C | G | G | T | T | C | C | T | C | A | G | C |
| | | C | A | A | G | C | G | G | T | T | T | C | T | T | A | G | C |
| Allele Frequency | A | 0.00 | 0.95 | 0.94 | | 0.00 | 0.00 | 0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.95 | 0.24 | 0.00 |
| | C | 0.95 | 0.00 | 0.00 | | 0.95 | 0.06 | 0.00 | 0.05 | 0.05 | 0.47 | 0.93 | 0.07 | 0.79 | 0.00 | 0.00 | 0.95 |
| | G | 0.00 | 0.00 | 0.06 | | 0.00 | 0.94 | 0.95 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.76 | 0.00 |
| | T | 0.05 | 0.05 | 0.00 | | 0.05 | 0.00 | 0.00 | 0.95 | 0.95 | 0.53 | 0.07 | 0.93 | 0.21 | 0.00 | 0.00 | 0.05 |

| | Pos | 93590 | 93619 | 93693 | 93787 | 93850 | 93863 | 93867 | 93884 | 93926 | 94022 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Locus | BTN1_1 | | | | | | | | | |
| | Gene | I | I | I | I | I | I | I | I | I | I |
| | Alleles | W | R | Y | W | W | M | Y | W | R | W |
| UMN/NTBF F1 | 7491 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | A | A | A | T | A | G | A |
| | 3804 | T | A | C | A | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | 1057 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | 1049 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | A | A | A | T | A | G | A |
| | 1044 | T | A | C | A | A | A | C | A | G | ? |
| | | T | A | T | T | T | C | T | T | G | ? |
| | 1042 | T | A | C | A | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| UMN/NTBF P | 4991 | T | A | C | T | T | C | C | T | G | T |
| | | T | A | C | T | T | C | C | T | G | T |
| | 4974 | T | A | C | A | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | 6002 | ? | A | T | A | A | A | T | A | G | A |
| | | ? | A | T | A | A | A | T | A | G | A |
| | 6013 | ? | A | T | A | A | A | T | A | G | A |
| | | ? | A | T | A | A | A | T | A | G | A |
| | 4973 | T | A | C | T | A | C | C | A | G | A |
| | | T | A | T | T | T | A | T | T | G | T |
| | 6014 | ? | A | T | A | A | A | T | A | G | A |
| | | ? | A | T | A | A | A | T | A | G | A |
| | 5017 | ? | A | C | T | T | C | C | T | G | T |
| | | ? | A | C | T | T | C | C | T | G | T |
| | 5025 | ? | A | C | T | T | C | C | T | G | T |
| | | ? | A | C | T | T | C | C | T | G | T |
| | 6012 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | A | A | A | T | A | G | A |
| | 4914 | T | A | C | T | T | C | C | A | G | A |
| | | T | A | T | A | A | A | T | T | G | T |
| NTE P and F1 | F5348 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | D5070 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | D5055 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | G1148 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | G4600 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | C | T | T | C | C | T | G | T |
| BAC | NTE | T | G | C | T | T | C | T | T | A | A |
| | WF06 | T | G | C | T | T | C | T | T | A | A |
| Commercil Breeder Line 1 | S1 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S2 | T | A | C | T | T | C | C | T | G | T |
| | | T | A | C | T | T | C | C | T | G | T |
| | S3 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S4 | T | A | C | T | T | C | C | T | G | A |
| | | T | A | C | T | T | C | C | T | G | T |
| | S5 | T | A | C | T | T | C | C | T | G | T |
| | | T | A | C | T | T | C | C | T | G | T |
| | S6 | T | A | C | T | T | C | C | T | G | A |
| | | T | A | C | T | T | C | C | T | G | T |
| Commercil Breeder Line 2 | S7 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S8 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S9 | T | A | C | T | A | A | T | A | A | A |
| | | T | G | T | T | T | C | T | T | G | A |
| | S10 | T | A | C | T | A | A | T | A | A | A |
| | | T | G | T | T | T | C | T | T | G | A |
| | S11 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S12 | A | A | C | T | A | A | T | A | G | A |
| | | T | G | T | T | A | A | T | A | G | T |
| Commercil Breeder Line 3 | S13 | T | A | C | T | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | S14 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | A | A | A | T | A | G | A |
| | S15 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S16 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S17 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S18 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | A | A | A | T | A | G | A |
| Commercil Breeder Line 4 | S19 | T | A | C | T | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | S20 | T | A | C | T | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | S21 | T | A | C | T | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | S22 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S23 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| | S24 | T | A | T | A | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| Commercil Breeder Line 5 | S25 | T | A | C | A | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | S26 | T | A | C | T | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | S27 | T | A | C | T | T | C | C | T | G | T |
| | | T | A | C | T | T | C | C | T | G | T |
| | S28 | T | A | C | T | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | S29 | T | A | C | A | A | A | C | A | G | A |
| | | T | A | T | T | T | C | T | T | G | T |
| | S30 | T | A | T | T | A | A | T | A | G | A |
| | | T | A | T | T | A | A | T | A | G | A |
| Allele Frequency | A | 0.01 | 0.95 | 0.00 | 0.29 | 0.66 | 0.66 | 0.00 | 0.66 | 0.04 | 0.72 |
| | C | 0.00 | 0.00 | 0.35 | 0.00 | 0.00 | 0.34 | 0.30 | 0.00 | 0.00 | 0.00 |
| | G | 0.00 | 0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.96 | 0.00 |
| | T | 0.99 | 0.00 | 0.65 | 0.71 | 0.34 | 0.00 | 0.70 | 0.34 | 0.00 | 0.28 |

| | Locus | BTN1_2 | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 100006 | 100042 | 100082 | 100085 | 100089 | 100188 | 100189 | 100272 | 100298 | 100343 | 100358 | 100380 | 100390 | 100481 | 100593 | 100597 | 100604 | 100653 | 100671 | 100723 |
| | Gene | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N |
| | Alleles | R | Y | M | S | C/CC/CCC | K | R | R | A/: | Y | R | Y | R | R | Y | Y | Y | Y | S | R |
| UMN/NTBF F1 | 7491 | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | 3804 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | 1057 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | 1049 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | T | G | G | A | T | G | T | G | G | C | T | T | T | G | G |
| | 1044 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | 1042 | A | T | C | C | CC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| UMN/NTBF P | 4991 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | 4974 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | 6002 | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | 6013 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 4973 | A | T | C | C | CC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CCC | G | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | 6014 | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | 5017 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | 5025 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | 6012 | A | T | C | C | CC | G | G | G | A | T | G | T | G | G | C | T | T | C | C | G |
| | | A | T | C | C | CC | G | G | G | A | T | G | T | G | G | C | T | T | C | C | G |
| | 4914 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | T | T | T | C | G |
| NTE P and F1 | F5348 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | D5070 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | D5055 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | G1148 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | G4600 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| BAC | NTE | A | C | C | C | CC | G | G | G | : | C | G | C | G | G | C | C | T | T | C | A |
| | WF06 | A | C | C | C | CC | G | G | G | : | C | G | C | G | G | C | C | T | T | C | A |
| Commercil Breeder Line 1 | S1 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | T | G | G | C | T | T | T | G | A |
| | S2 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | S3 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | S4 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | T | G | G | C | T | T | C | C | G |
| | S5 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | S6 | A | T | C | C | CCC | G | A | G | A | C | G | T | G | A | C | T | T | C | C | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | G | C | T | T | C | C | G |
| Commercil Breeder Line 2 | S7 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | S8 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | S9 | A | C | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | C | A |
| | | A | T | C | C | CC | G | G | G | : | C | G | C | G | G | C | C | T | T | G | A |
| | S10 | A | C | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | C | A |
| | | A | T | C | C | CC | G | G | G | : | C | G | C | G | G | C | C | T | T | G | A |
| | S11 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | S12 | A | C | A | C | C | G | G | G | A | C | G | C | G | A | C | C | C | T | C | A |
| | | G | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| Commercil Breeder Line 3 | S13 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | S14 | A | C | C | C | CC | T | G | G | A | C | G | C | A | A | C | C | T | T | C | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | T | C | T | T | G | G |
| | S15 | A | C | C | C | CC | G | G | G | A | C | G | C | A | A | C | C | T | T | C | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | T | C | T | T | G | G |
| | S16 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | S17 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | S18 | A | C | C | C | CC | T | G | G | A | C | G | C | A | A | C | C | T | T | C | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | T | C | T | T | G | G |
| Commercil Breeder Line 4 | S19 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | S20 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | S21 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | S22 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | S23 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | S24 | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | C | T | T | G | A |
| Commercil Breeder Line 5 | S25 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | S26 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | A |
| | | A | T | C | C | CC | G | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| | S27 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | S28 | A | T | C | C | CCC | G | A | A | A | C | A | C | A | A | C | T | T | C | C | G |
| | | A | T | C | G | CC | G | G | G | A | C | G | C | G | A | C | T | T | C | C | G |
| | S29 | A | T | C | C | CCC | G | A | G | A | C | G | C | G | A | C | C | T | C | C | G |
| | | A | T | C | C | CC | T | G | G | A | C | G | C | G | A | C | T | T | T | G | A |
| | S30 | A | T | C | C | CC | G | G | A | A | C | A | C | A | A | C | C | T | C | C | A |
| | | A | T | C | G | CC | G | G | G | A | C | G | C | G | A | C | T | T | T | G | G |
| Allele Frequency | A | 0.99 | 0.00 | 0.01 | 0.00 | | 0.00 | 0.28 | 0.02 | | 0.00 | 0.02 | 0.00 | 0.05 | 0.90 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.64 |
| | C | 0.00 | 0.08 | 0.99 | 0.98 | | 0.00 | 0.00 | 0.00 | | 0.97 | 0.00 | 0.94 | 0.00 | 0.00 | 0.97 | 0.64 | 0.01 | 0.36 | 0.44 | 0.00 |
| | G | 0.01 | 0.00 | 0.00 | 0.02 | | 0.75 | 0.72 | 0.98 | | 0.00 | 0.98 | 0.00 | 0.95 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.56 | 0.36 |
| | T | 0.00 | 0.92 | 0.00 | 0.00 | | 0.25 | 0.00 | 0.00 | | 0.03 | 0.00 | 0.06 | 0.00 | 0.00 | 0.03 | 0.36 | 0.99 | 0.64 | 0.00 | 0.00 |

| | Locus | Blec1 | | | | | | | | | | | | | | Tap 1 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 130676 | 130745 | 130750 | 130787 | 130788 | 130978 | 131029 | 131052 | 131169 | 131178 | 131280 | 131524 | 131543 | 131631 | 171845 | 171955 | 172022 | 172110 | 172117 | 172165 |
| | Gene | I | I | I | I | I | E/N | I | I | I | I | I | I | I | I | E/S | E/S | E/S | I | I | E/S |
| | Alleles | Y | R | M | K | M | Y | Y | M | Y | Y | S | R | R | R | Y | R | R | Y | Y | S |
| UMN/NTBF F1 | 7491 | T | A | A | G | A | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | | T | A | A | G | A | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | 3804 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | 1057 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | 1049 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | C | A | G | C | C | G |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | 1044 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | 1042 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| UMN/NTBF P | 4991 | C | G | C | T | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | G |
| | | C | G | C | T | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | C |
| | 4974 | C | A | A | G | A | C | C | C | C | C | G | ? | ? | ? | ? | G | G | C | C | G |
| | | T | G | C | T | C | T | T | C | C | C | G | ? | ? | ? | ? | G | G | C | C | C |
| | 6002 | T | A | A | G | A | T | T | C | C | C | G | ? | ? | ? | ? | G | G | C | C | G |
| | | T | A | A | G | A | T | T | C | C | C | G | ? | ? | ? | ? | G | G | C | C | G |
| | 6013 | T | A | A | G | A | T | T | C | C | C | G | ? | ? | ? | ? | G | G | C | C | G |
| | | T | A | A | G | A | T | T | C | C | C | G | ? | ? | ? | ? | G | G | C | C | G |
| | 4973 | C | G | C | T | C | C | C | C | C | C | G | ? | ? | ? | ? | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | C | C | G | ? | ? | ? | ? | G | G | C | C | G |
| | 6014 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | G |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | G |
| | 5017 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | C |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | C |
| | 5025 | C | G | C | T | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | C |
| | | C | G | C | T | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | C |
| | 6012 | C | G | C | T | C | C | C | C | C | C | G | ? | ? | ? | ? | A | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | ? | ? | ? | ? | A | G | C | C | G |
| | 4914 | C | A | A | T | A | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | G |
| | | C | A | C | G | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | G | G | C | C | G |
| NTE P and F1 | F5348 | T | A | A | G | A | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | D5070 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | D5055 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | G1148 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | G4600 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| BAC | NTE | T | G | C | T | C | C | C | C | T | C | C | A | A | A | C | G | G | C | C | G |
| | WF06 | T | G | C | T | C | C | C | C | T | C | C | A | A | A | C | G | G | C | C | G |
| Commercil Breeder Line 1 | S1 | C | G | C | T | C | C | C | C | C | C | C | G | G | G | C | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | T | T | G | G | G | G | T | G | G | T | C | G |
| | S2 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | S3 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S4 | C | G | C | T | C | C | C | C | C | C | C | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | T | T | G | G | G | G | C | G | G | C | C | G |
| | S5 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | S6 | C | G | C | T | C | C | C | C | C | C | C | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | T | T | G | G | G | G | C | G | G | C | C | G |
| Commercil Breeder Line 2 | S7 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | G | G | C | C | G |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | G | G | C | C | G |
| | S8 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S9 | C | G | C | T | C | C | C | C | C | C | ? | ? | ? | ? | C | G | G | C | C | G |
| | | T | G | C | T | C | C | C | C | T | C | ? | ? | ? | ? | T | G | G | C | T | G |
| | S10 | C | G | C | T | C | C | C | C | C | C | ? | ? | ? | ? | C | G | G | C | C | G |
| | | T | G | C | T | C | C | C | C | T | C | ? | ? | ? | ? | T | G | G | C | T | G |
| | S11 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S12 | C | A | C | T | C | C | C | A | C | C | ? | ? | ? | ? | C | G | G | C | C | G |
| | | T | G | C | T | C | C | C | C | T | C | ? | ? | ? | ? | T | G | G | C | C | G |
| Commercil Breeder Line 3 | S13 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S14 | T | A | A | G | A | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | | T | A | A | G | A | T | T | C | C | C | G | G | G | G | C | G | G | T | C | G |
| | S15 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | C | G | A | C | C | G |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S16 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | G | G | C | C | G |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | G | G | C | C | G |
| | S17 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S18 | T | A | A | G | A | T | T | C | C | C | G | G | G | G | C | G | A | C | C | G |
| | | T | A | A | G | A | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| Commercil Breeder Line 4 | S19 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S20 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S21 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S22 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | G | G | C | C | G |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | G | G | C | C | G |
| | S23 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S24 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | G | G | C | C | G |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | T | G | G | C | C | G |
| Commercil Breeder Line 5 | S25 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | ? | ? | ? | ? | ? | ? |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | ? | ? | ? | ? | ? | ? |
| | S26 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | T | G | G | C | C | G |
| | S27 | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | S28 | C | G | C | T | C | C | C | C | C | C | C | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | T | T | G | G | G | G | C | G | G | C | C | C |
| | S29 | C | A | A | G | A | C | C | C | C | C | G | G | G | G | C | G | G | C | C | C |
| | | T | G | C | T | C | T | T | C | C | C | G | G | G | G | C | G | G | C | C | G |
| | S30 | C | G | C | T | C | C | C | C | C | C | C | G | G | G | C | G | G | C | C | C |
| | | C | G | C | T | C | C | C | C | T | T | G | G | G | G | T | G | G | C | C | G |
| Allele Frequency | A | 0.00 | 0.26 | 0.24 | 0.00 | 0.24 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.03 | 0.03 | 0.03 | 0.00 | 0.03 | 0.02 | 0.00 | 0.00 | 0.00 |
| | C | 0.72 | 0.00 | 0.76 | 0.00 | 0.76 | 0.76 | 0.76 | 0.99 | 0.88 | 0.94 | 0.09 | 0.00 | 0.00 | 0.00 | 0.59 | 0.00 | 0.00 | 0.98 | 0.98 | 0.26 |
| | G | 0.00 | 0.74 | 0.00 | 0.24 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.91 | 0.97 | 0.97 | 0.97 | 0.00 | 0.97 | 0.98 | 0.00 | 0.00 | 0.74 |
| | T | 0.28 | 0.00 | 0.00 | 0.76 | 0.00 | 0.24 | 0.24 | 0.00 | 0.12 | 0.06 | 0.00 | 0.00 | 0.00 | 0.00 | 0.41 | 0.00 | 0.00 | 0.02 | 0.02 | 0.00 |

| | Locus | C4 | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 194767 | 194779 | 194780 | 194803 | 194807 | 194827 | 194831 | 195034 | 195092 | 195118 | 195333 | 195400 | 195435 | 195440 | 195441 | 195477 | 195592 |
| | Gene | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | E/S |
| | Alleles | Y | Y | R | GG/:: | R | R | Y | W | R | Y | R | R | S | R | R | Y | Y |
| UMN/NTBF F1 | 7491 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | 3804 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | 1057 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | 1049 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | 1044 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | 1042 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | A | G | C | G | A | T | C |
| UMN/NTBF P | 4991 | C | C | G | : | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | : | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | 4974 | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | : | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | 6002 | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | 6013 | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | 4973 | C | C | G | : | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | : | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | 6014 | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | 5017 | C | C | G | : | G | A | C | T | G | C | A | G | C | A | A | T | C |
| | | C | C | G | : | G | A | C | T | G | C | G | G | C | G | G | C | C |
| | 5025 | C | C | G | : | G | A | C | T | G | C | A | G | C | A | A | T | C |
| | | C | C | G | : | G | A | C | T | G | C | G | G | C | G | G | C | C |
| | 6012 | C | C | G | : | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | : | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | 4914 | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | : | G | A | C | A | A | C | A | G | C | A | G | C | C |
| NTE P and F1 | F5348 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | D5070 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | D5055 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | G1148 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | G4600 | C | C | G | :: | G | A | C | T | G | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | G | A | T | C |
| BAC | NTE | T | T | G | :: | G | A | C | A | A | C | G | A | C | A | G | C | C |
| | WF06 | T | T | G | :: | G | A | C | A | A | C | G | A | C | A | G | C | C |
| Commercil Breeder Line 1 | S1 | C | C | G | :: | G | A | C | A | A | C | G | A | C | A | G | C | T |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S2 | C | C | G | :: | G | A | C | T | G | C | A | G | C | G | A | T | C |
| | | C | C | G | :: | G | A | C | T | G | C | A | G | C | G | A | T | C |
| | S3 | C | C | G | :: | G | A | C | A | A | C | G | A | C | A | G | C | T |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S4 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S5 | C | C | G | :: | G | A | C | T | G | C | A | G | C | G | A | T | C |
| | | C | C | G | :: | G | A | C | T | G | C | A | G | C | G | A | T | C |
| | S6 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| Commercil Breeder Line 2 | S7 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S8 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S9 | C | C | G | :: | G | A | C | A | A | C | G | G | C | A | G | C | C |
| | | T | T | G | :: | G | A | C | T | G | C | G | A | C | A | G | C | C |
| | S10 | C | C | G | :: | G | A | C | A | A | C | G | G | C | A | G | C | C |
| | | T | T | G | :: | G | A | C | T | G | C | G | A | C | A | G | C | C |
| | S11 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S12 | C | C | G | GG | G | G | C | A | G | T | G | G | C | A | G | C | C |
| | | C | C | G | :: | A | A | C | T | G | C | G | G | G | A | G | C | C |
| Commercil Breeder Line 3 | S13 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S14 | C | C | G | :: | G | A | C | A | A | C | A | A | C | A | G | C | T |
| | | C | C | G | :: | G | A | C | A | A | C | G | G | C | A | G | C | C |
| | S15 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | A | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S16 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S17 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S18 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| Commercil Breeder Line 4 | S19 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S20 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S21 | C | C | G | :: | G | A | C | T | G | C | A | G | C | A | G | T | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | G | A | C | C |
| | S22 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S23 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S24 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| Commercil Breeder Line 5 | S25 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | S26 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | T | G | :: | G | A | T | T | G | C | G | G | C | A | G | C | C |
| | S27 | C | C | G | :: | G | A | C | T | G | C | A | G | C | G | A | T | C |
| | | C | C | G | :: | G | A | C | T | G | C | A | G | C | G | A | T | C |
| | S28 | C | C | G | :: | G | A | C | T | G | C | A | G | C | A | G | C | C |
| | | C | T | G | :: | G | A | T | T | G | C | G | G | C | G | A | T | C |
| | S29 | C | C | G | :: | G | A | C | A | A | C | A | G | C | A | G | C | C |
| | | C | T | G | :: | G | A | T | T | G | C | G | G | C | A | G | C | C |
| | S30 | C | C | G | :: | G | A | C | T | G | C | G | G | C | A | G | C | C |
| | | C | T | G | :: | G | A | T | T | G | C | G | G | C | A | G | C | C |
| Allele Frequency | A | 0.00 | 0.00 | 0.01 | | 0.01 | 0.99 | 0.00 | 0.35 | 0.34 | 0.00 | 0.39 | 0.07 | 0.00 | 0.88 | 0.12 | 0.00 | 0.00 |
| | C | 0.96 | 0.92 | 0.00 | | 0.00 | 0.00 | 0.96 | 0.00 | 0.00 | 0.99 | 0.00 | 0.00 | 0.99 | 0.00 | 0.00 | 0.88 | 0.97 |
| | G | 0.00 | 0.00 | 0.99 | | 0.99 | 0.01 | 0.00 | 0.00 | 0.66 | 0.00 | 0.61 | 0.93 | 0.01 | 0.12 | 0.88 | 0.00 | 0.00 |
| | T | 0.04 | 0.08 | 0.00 | | 0.00 | 0.00 | 0.04 | 0.65 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.03 |

| | Locus | C4 3` | | | | | | | | | CD1.1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 195613 | 195661 | 195797 | 195826 | 195848 | 195859 | 195923 | 195940 | 195957 | 85 | 127 | 223 | 265 | 299 | 350 | 398 | 457 | 590 |
| | Gene | E/S | I | I | I | I | I | I | I | I | I | I | I | I | E/S | E/S | E/S | E/S | I |
| | Alleles | R | Y | Y | Y | S | Y | Y | R | Y | Y | R | GATTGA | M | R | W | R | R | R |
| UMN/NTBF F1 | 7491 | G | C | C | T | C | T | C | G | C | C | A | : | C | G | A | G | G | A |
| | | G | C | C | T | C | T | C | G | C | C | A | : | C | G | A | G | G | A |
| | 3804 | A | C | C | T | C | T | C | G | C | C | A | GATTGA | C | G | A | G | G | : |
| | | G | T | C | T | C | T | C | G | C | C | G | : | C | G | T | G | G | : |
| | 1057 | A | C | C | T | C | T | C | G | C | C | A | GATTGA | A | G | A | G | G | : |
| | | G | T | C | T | C | T | C | G | C | T | G | : | C | G | T | G | G | : |
| | 1049 | A | C | C | T | C | T | C | G | C | C | A | GATTGA | C | G | A | G | G | : |
| | | G | C | C | T | C | T | C | G | C | C | G | : | C | G | T | G | G | : |
| | 1044 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | A |
| | | G | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | 1042 | A | C | C | C | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | A |
| | | G | C | C | T | G | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| UMN/NTBF P | 4991 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | G |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | G |
| | 4974 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | : | ? | ? | ? | ? | ? | ? |
| | 6002 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | : | ? | ? | ? | ? | ? | ? |
| | 6013 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | : | ? | ? | ? | ? | ? | ? |
| | 4973 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | A |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | A | G | T | G | G | G |
| | 6014 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | ? | ? | ? | ? | ? | ? |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | : | ? | ? | ? | ? | ? | ? |
| | 5017 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | G |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | G |
| | 5025 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | G |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | G |
| | 6012 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | A |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | GATTGA | C | G | T | G | G | A |
| | 4914 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | : | C | G | A | G | G | A |
| | | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | : | C | G | A | G | G | A |
| NTE P and F1 | F5348 | A | C | C | T | C | T | C | G | C | C | G | : | C | G | A | G | G | ? |
| | | G | C | C | T | C | T | C | G | C | C | A | GATTGA | C | G | T | G | G | ? |
| | D5070 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | A |
| | | A | C | C | T | C | T | C | G | T | T | G | GATTGA | A | G | T | G | G | G |
| | D5055 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | A |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | G1148 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | A |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | G4600 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | C | G | T | C | G | T | C | G | GATTGA | C | G | T | G | G | G |
| BAC | NTE | A | C | C | C | C | T | C | G | C | C | G | GATTGA | C | G | A | G | A | G |
| | WF06 | A | C | C | C | C | T | C | G | C | C | G | GATTGA | C | G | A | G | A | G |
| Commercil Breeder Line 1 | S1 | A | T | C | C | G | T | T | A | C | C | G | GATTGA | C | G | T | G | G | A |
| | | A | T | C | C | G | T | T | A | C | C | G | GATTGA | C | G | T | G | G | G |
| | S2 | A | T | C | C | G | T | C | G | T | T | G | GATTGA | A | G | T | G | G | A |
| | | A | T | C | C | G | T | C | G | T | T | G | GATTGA | A | G | T | G | G | A |
| | S3 | A | T | C | C | G | T | T | A | C | C | G | GATTGA | C | G | T | G | G | A |
| | | A | T | C | C | G | T | T | A | C | C | G | GATTGA | C | G | T | G | G | G |
| | S4 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | S5 | A | C | C | C | G | T | C | G | T | T | G | GATTGA | A | G | T | G | G | A |
| | | A | T | C | C | ? | T | C | G | T | T | G | GATTGA | A | G | T | G | G | A |
| | S6 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | A |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| Commercil Breeder Line 2 | S7 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | S8 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | S9 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | A | G | A | G |
| | | A | T | C | C | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | S10 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | A | G | A | G |
| | | A | T | C | C | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | S11 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | S12 | A | C | T | C | C | C | C | G | C | C | G | GATTGA | C | A | A | A | A | G |
| | | A | C | T | C | C | C | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| Commercil Breeder Line 3 | S13 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | A | G | T | G | G | A |
| | | A | T | C | T | C | T | C | G | C | T | G | GATTGA | C | G | T | G | G | G |
| | S14 | A | T | C | C | G | T | T | A | C | C | G | GATTGA | C | G | A | G | G | ? |
| | | G | T | C | C | G | T | T | A | C | C | G | : | C | G | T | G | G | ? |
| | S15 | A | C | C | T | C | T | C | G | C | T | G | GATTGA | A | G | T | G | G | A |
| | | A | T | C | T | C | T | C | G | T | T | G | GATTGA | A | G | T | G | G | A |
| | S16 | A | C | C | T | C | T | C | G | C | T | A | GATTGA | C | G | A | G | G | ? |
| | | G | T | C | T | C | T | C | G | C | T | G | : | A | G | T | G | G | ? |
| | S17 | A | C | C | T | C | T | C | G | C | T | A | GATTGA | C | G | A | G | G | ? |
| | | G | T | C | T | C | T | C | G | C | C | G | : | A | G | T | G | G | ? |
| | S18 | A | C | C | T | C | T | C | G | C | T | A | GATTGA | C | G | A | G | G | ? |
| | | G | C | C | T | C | T | C | G | T | C | G | : | A | G | T | G | G | ? |
| Commercil Breeder Line 4 | S19 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | S20 | A | C | C | T | C | T | C | G | C | C | A | GATTGA | C | G | A | G | G | ? |
| | | G | T | C | T | C | T | C | G | C | C | G | : | C | G | T | G | G | ? |
| | S21 | A | T | C | C | G | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | T | C | T | C | G | T | T | G | GATTGA | A | G | T | G | G | A |
| | S22 | A | C | C | T | C | T | C | G | C | C | A | GATTGA | C | G | A | G | G | ? |
| | | G | T | C | T | C | T | C | G | C | C | G | : | C | G | T | G | G | ? |
| | S23 | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | T | C | T | C | G | C | T | G | GATTGA | A | G | T | G | G | A |
| | S24 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | A | G | G | ? |
| | | G | T | C | T | C | T | C | G | C | C | A | : | C | G | T | G | G | ? |
| Commercil Breeder Line 5 | S25 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | A | G | G | ? |
| | | G | T | C | T | C | T | C | G | C | C | A | : | C | G | T | G | G | ? |
| | S26 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | S27 | A | T | C | C | G | T | C | G | T | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | C | G | T | C | G | T | C | G | GATTGA | A | G | T | G | G | A |
| | S28 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | G |
| | | A | T | C | C | G | T | C | G | T | C | G | GATTGA | A | G | T | G | A | A |
| | S29 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | C | G | T | G | G | ? |
| | | G | C | C | T | C | T | C | G | C | C | A | : | A | G | A | G | A | ? |
| | S30 | A | C | C | T | C | T | C | G | C | C | G | GATTGA | A | G | T | G | G | A |
| | | A | T | C | T | C | T | C | G | C | T | G | GATTGA | A | G | T | G | A | A |
| Allele Frequency | A | 0.80 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 | 0.00 | 0.00 | 0.17 | | 0.21 | 0.01 | 0.23 | 0.01 | 0.08 | 0.40 |
| | C | 0.00 | 0.40 | 0.98 | 0.26 | 0.82 | 0.02 | 0.93 | 0.00 | 0.86 | 0.81 | 0.00 | | 0.79 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | G | 0.20 | 0.00 | 0.00 | 0.00 | 0.18 | 0.00 | 0.00 | 0.93 | 0.00 | 0.00 | 0.83 | | 0.00 | 0.99 | 0.00 | 0.99 | 0.92 | 0.60 |
| | T | 0.00 | 0.60 | 0.02 | 0.74 | 0.00 | 0.98 | 0.07 | 0.00 | 0.14 | 0.19 | 0.00 | | 0.00 | 0.00 | 0.77 | 0.00 | 0.00 | 0.00 |

**Table 3:** Characterization of single nucleotide polymorphisms observed in the turkey MHC-B.

| | Observed | Frequency |
|---|---|---|
| **Non Coding** | 42 | 0.207 |
| **Intron** | 136 | 0.670 |
| **Exon** | 34 | 0.167 |
| **Syn** | 21 | 0.103 |
| **Non-Syn** | 13 | 0.064 |
| **Transition** | 154 | 0.759 |
| **Transversion** | 49 | 0.241 |
| **MAF $\geq$ 0.2** | 69 | 0.340 |

**Table 4:** MHC haplotypes derived from 52 individuals and 11 fully defined SNPs. Reference SNP positions based on DQ993255.

| Hap ID | Occurrences | 67099 | 67142 | 67427 | 77806 | 78256 | 93693 | 93787 | 93850 | 93863 | 93867 | 93884 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 25 | C | A | T | T | C | T | A | A | A | T | A |
| B | 25 | C | A | T | T | C | T | T | A | A | T | A |
| C | 1 | C | A | T | C | C | T | T | A | A | T | A |
| D | 7 | C | A | T | C | T | T | T | A | A | T | A |
| E | 9 | C | A | T | C | T | C | T | T | C | C | T |
| F | 1 | C | A | C | T | C | T | T | A | A | T | A |
| G | 2 | C | A | C | C | C | T | A | A | A | T | A |
| H | 1 | C | A | C | C | T | C | T | A | A | T | A |
| I | 4 | C | A | C | C | T | C | T | T | C | T | T |
| J | 4 | T | T | C | T | C | T | T | A | A | T | A |
| K | 3 | T | T | C | C | C | T | A | A | A | T | A |
| L | 21 | T | T | C | C | C | C | T | T | C | C | T |
| M | 1 | T | T | C | C | T | C | T | T | C | C | T |

**Table 5:** FastPHASE haplotypes reconstructed from 49 SNP loci. Reference SNP positions based on DQ993255. Individuals marked in bold were found to be monomorphic.

| ID | 1501 | 1620 | 1723 | 1748 | 1957 | 2136 | 8077 | 8250 | 8292 | 56416 | 56559 | 56663 | 56742 | 56806 | 56840 | 66845 | 66904 | 66970 | 67099 | 67142 | 67427 | 77806 | 78256 | 78368 | 93693 | 93787 | 93850 | 93863 | 93867 | 93884 | 94022 | 100188 | 100189 | 100597 | 100653 | 100671 | 100723 | 130676 | 130745 | 130750 | 130787 | 130788 | 130978 | 131029 | 171845 | 172165 | 195034 | 195092 | 195333 | Individuals |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | G | C | C | G | G | : | G | C | C | A | G | A | G | A | G | G | C | G | C | A | C | C | T | G | C | T | T | C | T | T | A | G | G | C | T | C | A | T | G | C | T | C | C | C | C | G | A | A | G | **NT WF06**, S9, S10 |
| 2 | A | C | T | A | C | : | A | C | A | C | G | G | A | A | A | G | C | A | C | A | T | T | C | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | T | G | G | F5348, Dam 5070, Dam 5055, G1148, S3, **S7, S8, S11**, S16, S17, S22, S23, S24 |
| 3 | G | C | C | G | C | : | G | T | C | A | A | A | A | A | A | A | C | G | C | A | T | T | C | G | T | A | A | A | T | A | A | T | G | C | T | G | A | T | A | A | G | A | T | T | C | G | A | A | A | **7491**, 3804, 1057, 1049, S25, S29, 4974, 6002, 6013, 6014, 4914 |
| 4 | A | T | C | G | C | : | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | G | 3804, 1044, S19, S21, S26, S29, 4991, 5017, 5025 |
| 5 | A | C | T | A | C | : | A | C | A | A | A | A | A | A | A | A | C | G | C | A | T | T | C | G | T | A | A | A | T | A | A | T | G | C | T | G | A | T | A | A | G | A | T | T | C | G | A | A | A | 1044, 1042, 6002, 6013, 6014 |
| 6 | A | T | C | G | C | : | A | T | C | A | G | A | A | A | A | A | T | G | C | A | T | C | T | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | T | G | G | G4600, S9, S10, S12, S30 |
| 7 | G | C | C | G | C | : | G | T | C | A | A | A | A | A | A | G | C | A | C | A | T | T | C | G | T | A | A | A | T | A | A | T | G | C | T | G | A | T | A | A | G | A | T | T | C | G | A | A | A | F5348, S17, S18, S22, S24 |
| 8 | A | C | C | G | G | G | A | C | C | A | G | A | A | A | A | A | T | G | C | A | T | C | T | G | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | A | **S27**, S28, 5025 |
| 9 | A | C | T | A | C | : | A | C | A | C | G | G | A | A | A | A | T | G | T | T | C | T | C | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | T | G | G | S19, S20, S26 |
| 10 | A | C | C | G | G | G | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | A | 1042, 5017 |
| 11 | A | C | C | G | G | G | A | T | C | C | G | G | A | A | A | G | C | A | C | A | T | T | C | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | T | G | G | Dam 5055, G1148 |
| 12 | A | T | C | G | C | : | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | A | S2, S5 |
| 13 | G | C | C | G | C | G | A | C | C | C | G | G | A | A | A | A | T | G | T | T | C | C | C | A | T | A | A | A | T | A | A | G | G | T | C | C | G | C | G | C | T | C | C | C | C | G | T | G | G | 1049, 6012 |
| 14 | A | C | T | A | C | : | A | C | A | C | G | G | A | A | A | A | C | G | C | A | T | T | C | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | T | G | G | 1057, 4973 |
| 15 | A | C | C | G | G | G | A | C | C | A | G | A | A | A | A | A | T | G | C | A | T | C | T | G | T | T | A | A | T | A | A | G | G | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | G | S28 |
| 16 | A | C | C | G | G | G | A | C | C | C | G | G | A | A | A | A | T | G | T | T | C | C | C | A | T | A | A | A | T | A | A | G | G | T | C | C | G | C | G | C | T | C | C | C | C | G | T | G | G | 6012 |
| 17 | A | C | C | G | G | G | A | T | C | A | G | A | A | A | A | A | T | G | C | A | T | C | T | G | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | A | S2 |
| 18 | A | C | C | G | G | G | A | T | C | A | G | A | G | A | G | A | T | G | C | A | T | C | T | G | C | T | T | C | C | T | A | G | G | T | C | C | A | C | G | C | T | C | C | C | C | G | T | G | G | S4 |
| 19 | A | C | C | G | G | G | A | T | C | C | G | A | A | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | G | S13 |
| 20 | A | C | C | G | G | G | A | T | C | C | G | A | G | T | G | A | C | G | C | A | C | C | C | G | T | A | A | A | T | A | A | T | G | C | T | C | G | T | A | A | G | A | T | T | C | G | A | A | A | S18 |
| 21 | A | C | C | G | G | G | A | T | C | C | G | A | G | T | G | A | C | G | C | A | C | C | C | G | T | T | A | A | T | A | A | G | G | C | T | C | G | C | G | C | T | C | C | C | T | G | T | G | G | S15 |
| 22 | A | C | C | G | G | G | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | A | C | T | C | C | C | C | G | A | A | A | 4914 |
| 23 | A | C | C | G | G | G | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | G | S6 |
| 24 | A | C | T | A | C | : | A | C | A | C | G | A | A | A | A | A | T | G | T | T | C | T | C | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | T | G | G | S13 |
| 25 | A | C | T | A | C | : | A | C | A | C | G | A | A | A | A | G | C | A | C | A | T | T | C | G | T | A | A | A | T | A | A | T | G | C | T | G | A | T | A | A | G | A | T | T | C | G | A | A | A | S14 |
| 26 | A | C | T | A | C | : | A | C | A | C | G | G | A | A | A | A | C | G | C | A | C | T | C | G | T | A | A | A | T | A | A | T | G | C | T | G | A | T | A | A | G | A | T | T | C | G | A | A | A | S15 |
| 28 | A | C | T | A | C | : | A | C | A | C | G | G | A | A | A | A | T | G | C | A | T | T | C | G | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | A | G4600 |
| 29 | A | C | T | A | C | : | A | C | A | C | G | G | A | A | A | A | T | G | C | A | T | T | C | G | T | T | A | A | T | A | A | G | G | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | G | S30 |
| 30 | A | C | T | A | C | : | A | C | A | C | G | G | A | A | A | G | C | A | C | A | T | T | C | A | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | A | A | G | S3 |
| 31 | A | C | T | A | C | : | A | C | A | C | G | G | A | A | A | G | C | A | C | A | T | T | C | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | A | A | A | Dam 5070 |
| 32 | A | T | C | G | C | : | A | T | C | A | G | A | A | A | A | A | T | G | C | A | T | C | T | G | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | G | S4 |
| 33 | A | T | C | G | C | : | A | T | C | A | G | A | A | A | A | A | T | G | C | A | T | C | T | G | T | T | A | A | T | A | A | G | G | T | C | C | A | C | G | C | T | C | C | C | C | G | T | G | G | S1 |
| 34 | A | T | C | G | C | : | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | G | A | A | A | S20 |
| 35 | A | T | C | G | C | : | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | G | T | G | G | 4991 |
| 36 | A | C | C | G | G | G | A | T | C | A | G | A | G | T | G | A | T | G | T | T | C | C | T | G | C | T | T | C | C | T | A | G | G | T | C | C | A | C | G | C | T | C | C | C | C | G | T | G | G | S6 |
| 37 | A | T | C | G | C | : | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | G | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | G | 4974 |
| 38 | A | T | C | G | C | : | A | T | C | C | G | A | G | T | G | A | T | A | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | T | C | G | C | G | C | T | C | C | C | C | C | T | G | G | 4973 |
| 39 | G | C | C | G | C | : | G | T | C | A | A | A | A | A | A | G | C | A | C | A | T | T | C | G | T | A | A | A | T | A | A | T | G | C | T | C | G | T | A | A | G | A | T | T | C | G | A | A | G | S14 |
| 40 | G | C | C | G | C | : | G | T | C | A | A | A | A | A | A | G | C | G | C | A | T | T | C | G | T | A | A | A | T | A | A | T | G | C | T | G | A | T | A | A | G | A | T | T | C | G | A | A | A | S16 |
| 41 | G | C | C | G | C | : | G | T | C | C | G | G | A | A | A | A | T | G | C | A | T | T | C | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | A | A | G | S1 |
| 42 | G | C | C | G | C | G | A | C | C | A | G | A | A | A | A | A | T | G | C | A | T | C | T | G | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | A | S5 |
| 43 | G | C | C | G | C | G | A | C | C | A | G | A | A | A | A | A | T | G | C | A | T | C | T | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | T | G | A | S21 |
| 44 | G | C | C | G | C | G | A | C | C | A | G | A | A | A | A | G | C | A | C | A | T | C | C | G | T | T | A | A | T | A | A | G | G | C | T | G | A | C | G | C | T | C | C | C | T | G | T | G | G | S23 |
| 45 | G | C | C | G | G | : | G | C | C | A | G | A | G | A | G | G | C | G | C | A | C | C | T | G | C | T | A | A | T | A | T | G | G | C | T | C | A | T | A | C | T | C | C | C | C | G | A | G | G | S12 |
| 46 | A | C | C | G | C | G | A | T | C | C | G | A | G | T | G | A | T | G | T | T | C | C | C | A | C | T | T | C | C | T | T | G | A | T | C | C | G | C | G | C | T | C | C | C | C | C | T | G | G | S25 |

**Figure 1:** The turkey MHC-B locus and amplicons examined in this study. Blue arrows indicate position of the resequenced amplicons, black arrows indicate amplicons resequenced and genetically mapped, and red arrows indicate microsatellites/indels genetically mapped.

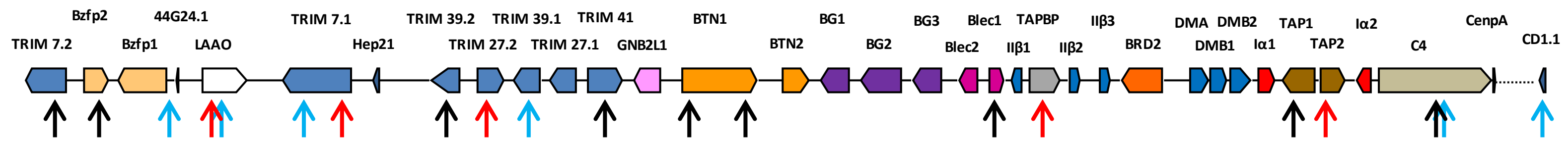


**Figure 2:** Minor allele frequencies of B-locus SNPs observed in commercial turkeys.

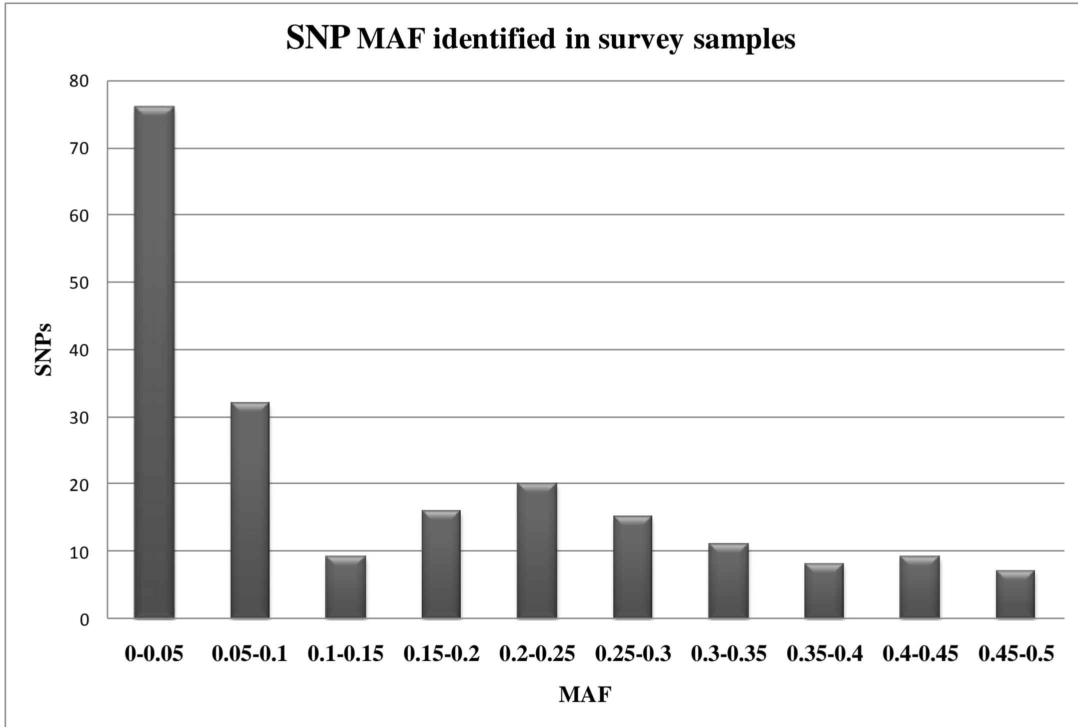


**Figure 3:** Haplotype proportions derived from analysis of 11 “minimal shared” SNPs.

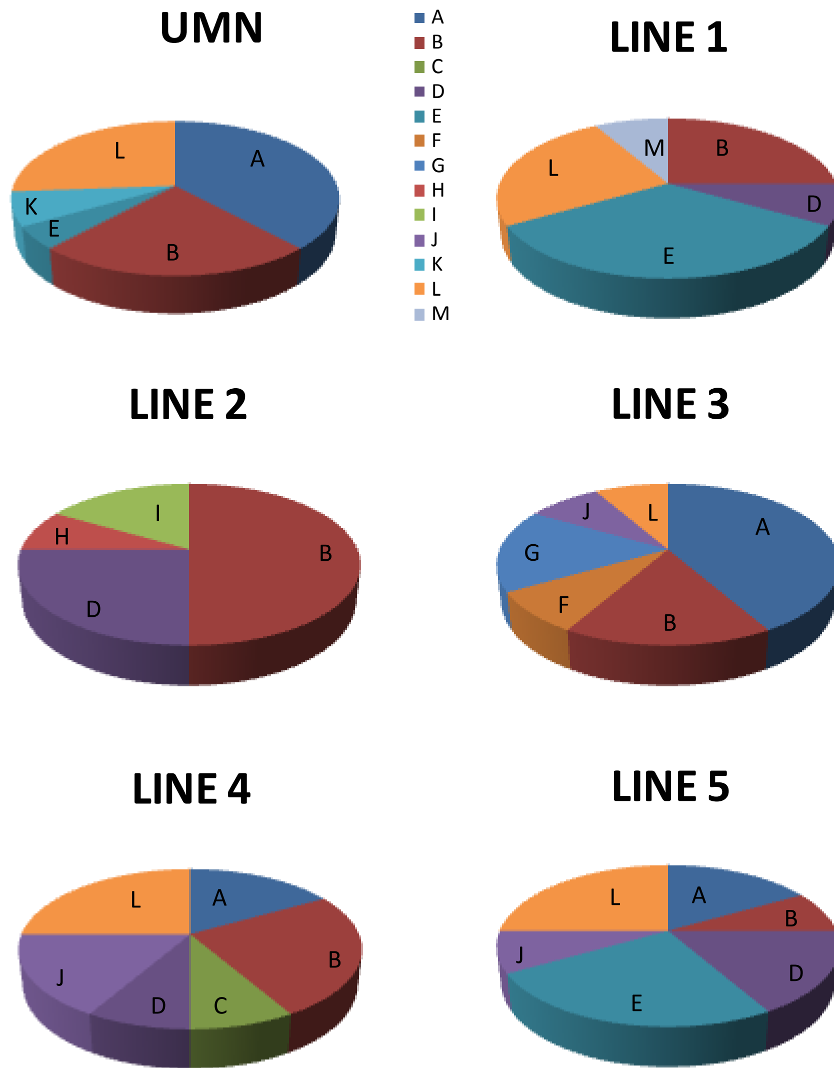


**Figure 4:** Phylogenetic relationships among the 45 FastPHASE assembled turkey MHC-B haplotypes. Shaded areas (A-M) denote haplotypes identified with the "minimal shared" SNP dataset. Haplotypes representing homozygous individuals are denoted with an asterisk.

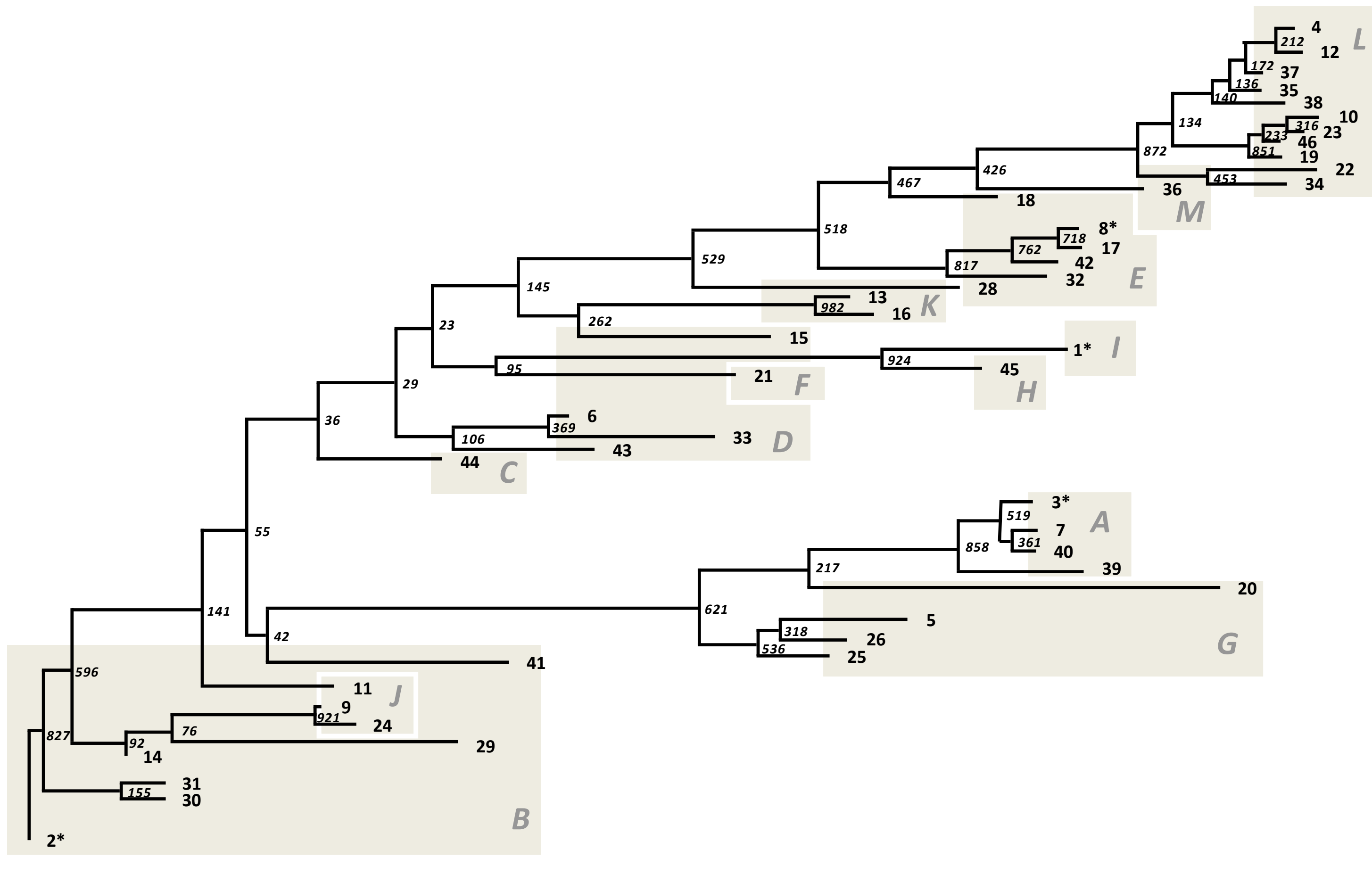
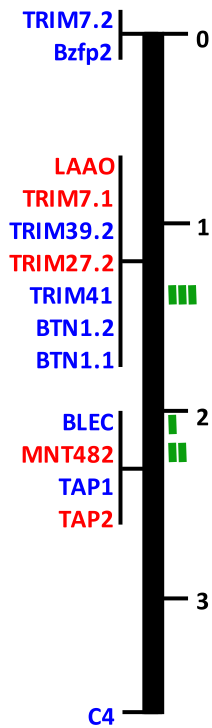


**Figure 5:** Genetic map of the turkey MHC-B locus. Single nucleotide polymorphisms are in blue and microsatellite/indels are in red. Green bars indicate position and number of gene conversion events. Scale is in cM.



# Chapter 6

## A Locus-Wide Approach to Assessing Variation in the Avian MHC: The B-Locus of the Wild Turkey

# A Locus-Wide Approach to Assessing Variation in the Avian MHC: the B-Locus of the Wild Turkey

**Chaves L. D.[1], G. M. Faile[1], K. E. Mock[2], and K. M. Reed[1]**

[1]Department of Veterinary and Biomedical Sciences

College of Veterinary Medicin

University of Minnesota, St Paul, MN 55108

[2] Wildland Resources Department

Utah State University, Logan, UT 84322-5230

## INTRODUCTION

The major histocompatibility complex (MHC) is the most polymorphic and gene dense region in the vertebrate genome (11). The locus is large, containing numerous genes involved in immune responses. Included in the MHC are genes for cytokines, complement proteins, antigen processing, and highly polymorphic glycoproteins responsible for presenting endogenous and exogenous antigens (MHC Class I and MHC Class II, respectively) to T-cells. The large mammalian MHC (~ 3-7 Mb) contains 200-400 genes (11). The best characterized MHC, that of the human (HLA), includes 3 classical (HLA-A, B, C) and 3 non-classical (HLA-E, F, G) Class I loci and Class I-like pseudo-genes as well as a highly variable number of Class II loci.

Variation in genes at the MHC is vital to generate a species-wide diverse array of antigenic peptides for T-cell presentation. The rapidly evolving pathogen repertoire requires multiple MHC loci with molecules highly polymorphic in order to present pathogen derived peptides to effector T-cells. Single loci in the human MHC have 100's of alleles and allele combinations may generate 100,000s of potential haplotypes (148). In contrast to mammals, birds show more constrained levels of diversity in their MHC (Chapter 5, (36)). For example, a recent study in domestic chickens identified nearly 100 MHC-B haplotypes (157). MHC haplotypes have repeatedly been found to confer resistance or susceptibility in chickens to numerous pathogens (147). Detailed studies of avian species are just beginning to describe the total extent of both gene content and genetic diversity.

In addition to Class I and Class II genes, clusters of other related genes can be found within the MHC such as tripartite motif loci, the butyrophilin loci, and an extensive

array of variable olfactory receptors. The presence of tightly linked olfactory receptors drew researchers to study the link between MHC genotype, odor perception, and potential mate selection (158), with numerous studies in different species sometimes drawing contradictory conclusions (159). Nonetheless, a role of the MHC in olfaction and mate selection has been repeatedly documented in mammals and some fish.

Until recently the role of the MHC in odor perception and/or mate selection has not been examined in birds, owing to a lack of genomic resources (160). It is known from the chicken genome that birds possess a large array of olfactory receptors (161) of which the vast majority cannot be positioned in the chicken genome assembly. Interestingly, the chicken MHC chromosome (GGA16) is one of the least assembled and annotated chromosomes in the genome (30).

The best studied bird species are those of agricultural significance. MHC regions have been examined in chicken, quail, duck and turkey (28,37,125). The core MHC locus of the turkey and chicken (MHC-B) are comprised of 5 and 4 MHC class I and II genes, respectively, spanning ~50 kb (Chapter 3). Studies of commercial chicken lines have clearly demonstrated that allelic diversity in commercial genetics is limited (162). Fewer alleles at genes of the MHC-B locus may limit the number of potential haplotypes within subpopulations (29). Likewise, a sampling of genetic variation and diversity within commercial turkeys suggests some haplotypes occur at high frequency within breeder lines (Chapter 5).

Population genetic studies of the MHC in birds have focused mainly on two technologies. Southern hybridization methods have been used to estimate gene copy number for the MHC Class I and II loci, as well as provide a rough estimate of genetic

variation (113,163-165). Alternatively, PCR techniques involving amplifying, cloning, and sequencing the highly variable MHC Class I and II exons (~200 bp) encoding the peptide binding domains provide a good approximation of genetic diversity within populations and gives a reasonable estimate of the number of loci present in the genome. Typically, these studies have identified 2 to 4 class IIβ loci in birds and provided evidence detailing an increased level of MHC diversity corresponding to the extent of out-breeding in the populations. However, these methods can also lead to false assumptions regarding gene number, as similar early studies of the chicken overestimated the number of loci (29,91).

The turkey is the only domesticated agricultural species native to North America. Six subspecies of the North American indigenous wild turkey have been described (166). The Eastern (*Meleagris gallopavo silvestris*) dominates in geographical distribution residing from the east coast through the Great Plains. Florida (*Meleagris gallopavo osceola*) is an exclusive inhabitant to the state of its namesake. Merriam's (*Meleagris gallopavo merriami*) is the western-most subspecies located in parts of Colorado, New Mexico, and Arizona. Rio Grande (*Meleagris gallopavo intermedia*) is present in most of Texas and parts of northern Mexico. Gould's (*Meleagris gallopavo mexicana*) and South Mexican (*Meleagris gallopavo gallopavo*) subspecies resides in western and southern Mexico, respectively. An additional, closely related species, the ocellated turkey (*Meleagris ocellata*), is found on the Yucatán Peninsula in Mexico as well as parts of Belize and the northern part of Guatemala. Commercial breeds are believed to be derived primarily from the South Mexican subspecies via reintroduction to the New World from

Europe (66). Thus, commercial turkeys represent only a small, selected sampling of the total genetic variation of the species.

This work was undertaken to analyze the extent of variation within the MHC-B locus of wild turkeys. Results of this study clearly demonstrate that MHC diversity in wild birds is significantly greater than that of commercial lines.

## MATERIALS AND METHODS

### *Genomic Resequencing*

A panel of 40 DNAs from 3 subspecies (Eastern, Merriam's, and Rio Grande) of turkey (*Meleagris gallopavo*) was examined in this study (Figure 1, Table 1). Birds were from 11 different locales representing a subset of samples examined by Mock et al. (167), Latch et al. (168) and a hunter harvested bird from Winona, MN. Individuals were amplified at 9 interspersed locations across the MHC-B region (Table 2) as previously described (Chapter 5). PCR products were purified using a MinElute PCR Purification Kit (Qiagen) and sequenced with an automated ABI Sequencer. Sequence data were manually analyzed using Sequencher software (Gene Codes, Corp.) and sequence polymorphisms (single nucleotide polymorphisms and deletion insertion polymorphisms, SNPs and DIPs, respectively) were recorded.

### *Haplotype Identification and phylogenetic analysis*

Polymorphisms were analysed using Arlequin, PHASE, and Haploview software to identify haplotype diversity and frequency within the 40 resequenced birds

(150,169,170). Analysis was performed using 42 SNPs and 3 DIPs successfully identified on all 40 individuals and that had minor allele frequencies (MAF) of 0.2 or greater. An additional haplotype inference analysis included samples from a previous study of commercial turkeys (Chapter 5) using SNPs and DIPs of high MAF common among the two groups. Phylogenetic analysis of the haplotypes was performed using ClustalW software (151).

**RESULTS**

*MHC-B Polymorphisms*

Over 9 kb of the MHC-B region was resequenced on 40 wild turkeys from across North America. A total of 238 polymorphisms were identified with minor allele frequencies (MAF) ranging between 0.01 to 0.5 and averaging 0.15 (Table 3). Thirty-seven percent of the loci had an MAF $\geq 0.2$. The frequency of SNPs in this region, 1 SNP/40 bp, is higher than the 1 SNP/200 bp found in other regions of the turkey genome (154) or the 1/70 bp identified within the MHC-B in commercial breeder lines (Chapter 5). The majority of SNPs were transition substitutions and located in introns (Table 4). Of the 45 SNPs located in coding sequence, 18 represent non-synonymous substitutions.

In addition to SNPs, 9 DIPs were present in the study samples. Included in these is a 57 bp biallelic DIP in the first intron of Bzfp2 not previously identified in commercial populations and a large polymorphic repeat element present within intron 6 of *LAAO*. Examination of this locus by electrophoresis of the PCR amplicons on 3%

agarose and 5% denaturing acrylamide gels found five alleles (one more than identified in commercial breeder lines [Chapter 5]) at this locus ranging in size from 180 to ~350bp.

Of the 238 polymorphisms, 112 (109 SNPs and 3 DIPs) were not previously identified in commercial turkeys (Chapter 5). The novel SNPs had a mean minor allele frequency of 0.07 and ranged between 0.01 to 0.49. Sixteen SNPs previously found in commercial birds were not present in the wild turkey data set. However, many of these may have been lost through technical failures (poor amplification or unidirectional resequencing).

Two tri-allelic SNPs were identified at positions 56210 (H) and 93867 (H). Allele frequencies at these loci ranged from 0.038 to 0.875 and averaged 0.25. SNPs at these loci were excluded from further analysis. The TRIM39.1 amplicon showed poor amplification and sequencing success. This is perhaps due to null alleles (nucleotide variations within the primer binding site(s)) within wild birds. This locus was not included in haplotype inference. Finally, no polymorphisms in the TAP1 amplicon had a MAF $\geq$ 0.1.

*Haplotype Reconstruction*

Haplotype inference was conducted using 45 polymorphic loci that displayed varying allele representation between the separate subspecies (Figure 3) and that had MAF $\geq$ 0.2. Initial analysis combining all individuals identified 4 monomorphic individuals (MCS6, MCS7, ECL10, and MINN). Of the 40 individuals examined, 70 potential haplotypes were inferred by PHASE software analysis. The majority of individuals were estimated to possess unique haplotypes; however, the Colorado Springs

samples had two monomorphic individuals as well as two additional individuals who shared common haplotypes with the four monomorphic individuals.

A second haplotype analysis incorporated genotypes from a panel of 52 commercial breeder turkeys (Chapter 5) and included 37 markers with MAF ≥ 0.2. This analysis identified a total of 99 haplotypes, of which the wild turkeys possessed 66 (Table 5). Due to extensive variation, many of the wild turkey MHC haplotypes could not be definitively inferred. A single haplotype was shared between the commercial and wild populations; EOM6 was identified to have a haplotype in common with NT WF06, the only fully sequenced turkey haplotype (Chapter 3). The wild turkey MHC-B haplotypes differed from each other at 1 to 26 SNPs with an average of 15 differences (standard deviation of 5.5). The two rounds of PHASE inference had similar posterior probabilities when averaged per locus. The posterior probabilities for haplotype allele assignment at a given SNP position ranged from 0.5 (no support) to 1.0 (complete linkage disequilibrium (LD)) within the sampled individuals.

*Phylogenetic Analysis*

Rare alleles within individuals residing in unique locations/subspecies were more frequent in the EWW sample and the Eastern subspecies, respectively (Table 6). As expected given the physical and genetic distance between populations, loci in Hardy-Weinberg equilibrium occurred more frequently within locations rather than subspecies (Table 6). Only a limited extent of linkage disequilibrium can be identified from the wild turkey samples (Figure 4).

Pairwise $F_{ST}$ analysis found significant variation between the three subspecies. The Eastern and Merriam's subspecies had the lowest $F_{ST}$ value (0.08) whereas the value for Eastern and Rio Grande was 0.12. The greatest pairwise $F_{ST}$ value occurred between Rio Grande versus Merriam's (0.18). Corrected average pairwise differences between subspecies were similar, with values between Eastern and Merriam's, Eastern and Rio Grande, and Merriam's and Rio Grande 1.51, 2.52, and 3.75, respectively. These values appear to contradict the distances previously identified with genomic markers for these subspecies (167), where Merriam's and Rio Grande were found to be more closely related than Eastern and Merriam's.

Phylogenetic analysis of haplotypes derived from analysis of 37 SNPs was performed in order to compare the origin and variation of wild turkey MHC haplotypes within subspecies and to those of common commercial breeder lines (Figure 5). The phylogenetic tree shows an interleaving of commercial and wild haplotypes along with a separate cluster of considerably divergent wild haplotypes. Most of the commercial haplotypes cluster near each other. Interestingly, the haplotype of NT WF06 (the only fully sequenced turkey MHC-B haplotype) is found in a distant node and is the only commercial haplotype completely shared with the wild birds surveyed.

**DISCUSSION**

Population genetic studies of avian MHCs have traditionally relied on Southern hybridizations to elucidate gene content and variation (57,59-61,171). While this approach can be an easy and powerful tool to quickly examine genomic variation within the species, the technique ultimately fails to successfully capture the true extent of

genetic diversity. Southern hybridizations rely on relatively few variable nucleotides which may be shared by multiple functional haplotypes and limited by the inability to select probes to detect all sequence variants, thereby underestimating the total level of diversity within the population. Resequencing of the exon(s) encoding Class I and/or II PBR may not accurately identify the diversity within the species either. PBRs are highly polymorphic and occur in multiple copies within the genome. PCR amplification, cloning, and sequencing is a standard procedure, however, often ignored is the high level of mosaic sequences (nearly 25%) generated in this process leading to false identification of alleles/loci (172). Additionally, neither the locus of origin nor the possibility of null amplifications can be assessed using this technique. Further, as has been previously identified in commercial turkeys, PBR alleles can be shared between loci, which would not be identified by cloning and sequencing. Additionally, detailed genomic descriptions of avian MHC is limited to a few species so that the number of MHC Class I and Class II loci is typically not known in most species where heterozygosity and rare SNPs may inadvertently be misidentified as novel haplotypes or additional genes.

MHC variation can be found at both extremes in wild bird populations, with some species having seemingly unlimited alleles and others presenting almost no polymorphism. A study of Class IIβ gene diversity in Red-winged Blackbirds (*Agelaius phoeniceus*) found extensive variation within and between individuals while a sequencing survey of the peptide binding region of a Class IIβ locus from the endangered Galápagos penguin (*Spheniscus mendiculus*) found only 3 highly similar alleles (173,174). Sequencing of cDNA from a single individual identified seven unique Class IIβ sequences in the Great Reed Warbler (*Acrocephalus arundinaceus*) suggesting at least 4

loci in the species (113). PCR-based techniques for surveying the polymorphic, multi-copy MHC genes also have the potential to identify pseudogenes and/or non-classical loci and can also miss loci and/or alleles due to null amplification (175).

The resequencing effort described in the present study examined 9 kb across 40 individuals and found nearly twice as many polymorphisms per base pair as compared those identified in commercial turkeys. Over 100 new polymorphisms were identified, however many of the novel SNPs had lower MAF compared to those shared among wild and commercial birds. The lower MAF can be attributed to relatively recent mutations or simply be a result of the individuals sampled in the survey. The significant loss of the presence of species-wide minor alleles in commercial populations has been well documented in the chicken (162).

Expansion of MHC haplotypes in wild birds relative to commercial lines has also been observed in a previous study of turkeys (60). In that study a single, predominate haplotype (identified through Southern hybridizations) was present in the commercial lines as well as wild turkeys. The wild birds also had the highest proportion of unique haplotypes. Further, in a survey of 40 game farm wild turkeys, Buchholz et al. identified 40 unique genotypes by Southern hybridization techniques (176). Interestingly, the single turkey sampled from Minnesota was found to be monomorphic. The significance of this is not fully known. Wild turkeys in Minnesota were at one point extirpated due to over hunting. In fact, it was once thought that wild turkeys would freeze in the colder winters. However, in 1973 the state of Minnesota traded 85 ruffed grouse for 29 Eastern wild turkeys from Missouri, which they released in the vicinity of Winona and Houston

Counties (Minnesota Department of Natural Resources). Whether this bottleneck lead to significantly reduced genetic variation in this region has yet to be determined.

The high level of sequence diversity is reflected in the inability to establish common haplotypes among wild turkeys, even while using relatively common alleles of extremely close proximity. These results can be explained by high levels of recombination and gene conversion within the region, as has been previously suggested (Chapter 5). Novel mutations are not likely responsible for the inability to infer and assemble common haplotypes in the subspecies as the polymorphisms used were of higher allele frequencies and thus are likely ancient polymorphisms segregating within the species.

Linkage disequilibrium in the chicken has been studied on various chromosomes in several different studies. In a large study of several commercial breeding lines, Andreescu et al. found the extent of LD in the chicken genome lower than other agricultural species (177). Strong LD ($r^2 > 0.8$) was observed in 10% of SNPs within 0.5 cM. Further, 24% of SNP pairs (within 0.5 cM) had an $r^2$ value greater than 0.5. Interestingly, in a study of wild chickens (Red Jungle Fowl) LD was found to be higher across a 200 kb region of chromosome 1 compared to two other commercial chicken lines (178). No study has yet been reported on the extent of LD within the chicken MHC.

LD was virtually non-existent within the wild turkey MHC-B, although the MHC is one of the highest LD regions of the human genome (13). The SNPs used in this study span a distance of 100 kb, and based on a previous study this would correspond to a distance of 0.25 cM (Chapter 5). Within the wild population, only four SNP pairs (0.6%) had an $r^2$ value greater than 0.8 and 14 (2%) had an $r^2 > 0.5$. These values are lower in the

wild birds. Several factors could be contributing to the low LD. As stated above, high levels of recombination and/or gene conversion could lead to the loss of LD. Based on the results of this study, the total number of haplotypes present in wild turkeys is high and the total number of samples surveyed may not have been enough to accurately and successfully identify alleles in disequilibrium. Further, resequencing samples in this manner may not sufficiently address the potential for null alleles, which can lead to the loss of LD.

As demonstrated, possession of a reference haplotype and the ability to resequence interspersed regions of the MHC presented here has significant advantages over several other methods of surveying the MHC. Resequencing several proximal non-MHC regions combined with haplotype inference identifies a greater extent of diversity throughout the region without the technical limitations incurred by Southern hybridizations or selective cloning and sequencing. Yet to be determined is whether this large extent of diversity in proximal regions translates into a large level of diversity at the most relevant regions of the MHC: the sequence encoding the peptide binding region of the MHC molecules.

**Table 1:** Distribution of subpopulations and regional representatives of wild turkeys examined for MHC variation.

| Sample IDs | Location | Subspecies | # |
|---|---|---|---|
| ECL | Camp Lejeune, NC | Eastern | 6 |
| EMC | NSWC, Indians | Eastern | 2 |
| EOM | Ozark Mtns, MO | Eastern | 3 |
| EWW | Waterhorn WMA, SC | Eastern | 4 |
| MINN | Winona, MN | Eastern | 1 |
| MCS | Colorado Springs, CO | Merriam's | 5 |
| MGD | Larimer Co., CO | Merriam's | 3 |
| MSL | Stoneman Lake, AZ | Merriam's | 5 |
| REN | Kleberg County TX | Rio Grande | 4 |
| RKR | King Ranch, TX | Rio Grande | 1 |
| RKW | Kerr WMA, TX | Rio Grande | 6 |

**Table 2:** Primers and reaction conditions for MHC loci examined in wild turkeys.

| Locus | Forward Primer | Reverse Primer | Product Size | TM | Ext Time |
|---|---|---|---|---|---|
| **TRIM 7.2** | TTCTCCAACTCGAACCGAAG | CCAGCTGGACAAACTGCTG | 848 | 62 | 1min |
| **Bzfp2** | ATTCAGCGCGATCCTAAATG | GTCTCCTTGGGTACGTGGAG | 941 | 58 | 1min |
| **LAAO** | ATTTGCTGAACGAGGATTCG | CCTTCGTGCTCAAAGAGAGC | 1000 | 58 | 1min |
| **TRIM 39.1** | GGCAGGGATAATGAATGAGG | TATGTGGCCCCTCTGAAAAG | 1141 | 58 | 2min |
| **TRIM 41** | CGGAGTCACAAGCATCACAG | GGCCAGCAAGCTCTCATAAC | 1203 | 62 | 2min |
| **BTN1_1** | AGAGCGTCACACCAAAATAGG | GGGAAGTCTTGGTTGTGTTTC | 534 | 56 | 1min |
| **BTN1_2** | CCCATTGACATGAAGTGACC | AAACAGGTGGCCAAGAACAC | 924 | 60 | 1min |
| **Blec1** | CACTGCTGTTGCTTTTGCAG | GTTTGGTGCAAACCCAATTC | 1419 | 62 | 2min |
| **TAP1** | GCCAGATACCACAGCAGGAG | GGCCGTGCCCTACTACAC | 458 | 60 | 30s |

.

**Table 3:** Polymorphisms identified in wild turkeys. Position is given with respect to DQ993255 and EU52267 (CD1.1). Under Gene, I, E/N, E/S, and N indicate intron, exon: non-synonomous, exon: synonomous, or non-codaing, respectively. Alleles use the genetic ambiguity code (R= A or G, Y= C or T, S= G or C, K= G or T, M= A or C, W= A or T) or the inserted base, where ":" denotes a deletion. Shaded loci were not previously identified in commercial turkeys.

| | Locus | TRIM 7.2 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 1462 | 1501 | 1520 | 1568 | 1638 | 1662 | 1723 | 1738 | 1748 | 1817 | 1856 | 1870 | 1881 | 1887 | 1888 | 1916 | 1923 | 1957 | 1972 | 1974 | 1992 | 2057 | 2058 | 2092 | 2119 | 2136 | 2157 | 2165 |
| | Gene | I | E/S | E/S | E/S | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | E/N | E/S |
| | Alleles | R | R | R | R | K | R | R | R | R | R | W | Y | K | Y | R | Y | S | S | R | Y | R | Y | R | W | Y | G | R | R |
| Eastern | ECL1 | G | G | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | T | G | C | G | T | T | G | ? | ? |
| | ECL1 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | ? | ? |
| | ECL10 | G | A | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| | ECL10 | G | A | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| | ECL3 | G | G | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | ECL3 | G | A | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| | ECL5 | G | A | G | G | G | G | C | G | A | A | T | T | G | T | G | T | C | G | G | C | G | C | G | T | C | : | G | ? |
| | ECL5 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | A | T | G | C | G | C | G | C | G | T | C | : | G | ? |
| | ECL6 | G | A | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | G | G | C | G | C | G | A | C | G | G | ? |
| | ECL6 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | A | ? |
| | ECL8 | G | G | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | C | G | T | G | T | C | : | A | ? |
| | ECL8 | G | G | G | G | G | G | C | G | G | G | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | ? |
| | EMC1 | G | G | G | G | G | G | C | G | G | G | T | T | G | T | G | C | G | G | G | C | G | C | G | T | C | : | G | A |
| | EMC1 | G | G | G | G | G | G | C | A | G | G | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| | EMC2 | G | G | G | G | G | G | C | G | G | G | T | T | G | T | G | C | G | G | G | C | G | C | G | T | C | : | G | A |
| | EMC2 | G | G | G | G | G | G | C | A | G | G | T | T | G | T | G | C | G | G | G | C | G | C | G | T | C | : | G | G |
| | EOM3 | G | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | EOM3 | A | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | A | A | C | : | G | G |
| | EOM4 | G | G | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | A | C | G | T | C | : | G | G |
| | EOM4 | G | G | G | G | G | G | C | G | G | G | T | T | G | T | G | C | G | G | G | C | G | C | G | T | C | : | G | A |
| | EOM6 | G | G | G | A | G | G | C | A | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | ? |
| | EOM6 | A | G | A | G | G | G | C | G | G | G | T | T | G | T | G | C | G | G | G | C | G | C | G | T | C | : | G | ? |
| | EWW11 | G | A | G | G | G | G | T | G | A | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | EWW11 | G | A | G | G | G | G | T | G | A | G | T | T | G | T | A | T | C | G | G | C | G | C | G | T | C | : | G | G |
| | EWW13 | G | G | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | EWW13 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | A |
| | EWW14 | G | A | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| | EWW14 | G | A | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | A | G |
| | EWW4 | G | G | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | C | G | T | G | T | C | G | ? | ? |
| | EWW4 | G | A | G | G | G | G | T | G | A | A | T | T | G | T | G | T | G | C | G | C | G | T | G | T | C | : | ? | ? |
| | MINN | G | A | G | G | G | G | T | G | A | A | T | T | G | T | G | T | G | C | G | C | G | T | G | T | C | : | ? | ? |
| | MINN | G | A | G | G | G | G | T | G | A | A | T | T | G | T | G | T | G | C | G | C | G | T | G | T | C | : | ? | ? |
| Merriam's | MCS1 | G | G | A | A | G | G | C | G | A | A | T | T | G | T | G | T | G | G | A | C | G | C | G | T | C | G | G | ? |
| | MCS1 | A | G | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | ? |
| | MCS3 | G | G | G | A | G | G | C | G | A | A | T | T | G | T | G | T | G | G | A | T | G | C | G | T | T | G | G | G |
| | MCS3 | G | A | G | G | G | A | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | G | G | G |
| | MCS4 | G | G | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | T | G | C | G | T | C | : | ? | ? |
| | MCS4 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | T | G | T | C | : | ? | ? |
| | MCS6 | G | A | G | G | G | A | C | G | G | A | T | T | G | T | G | T | G | G | G | T | G | C | G | T | T | G | G | G |
| | MCS6 | G | A | G | G | G | A | C | G | G | A | T | T | G | T | G | T | G | G | G | T | G | C | G | T | T | G | G | G |
| | MCS7 | G | G | G | A | G | G | C | G | A | A | T | T | G | T | G | T | G | C | A | C | G | C | G | T | C | G | G | G |
| | MCS7 | G | G | G | A | G | G | C | G | A | A | T | T | G | T | G | T | G | C | A | C | G | C | G | T | C | G | G | G |
| | MGD10 | G | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | T | G | C | G | T | C | : | G | G |
| | MGD10 | A | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | T | G | C | G | T | C | : | G | A |
| | MGD8 | G | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | T | G | C | G | T | C | : | G | G |
| | MGD8 | A | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | A |
| | MGD9 | A | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | MGD9 | A | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| | MSL11 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | T | G | C | G | T | C | : | G | G |
| | MSL11 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| | MSL12 | G | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | T | G | C | G | T | C | : | G | A |
| | MSL12 | A | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| | MSL4 | G | G | A | G | G | G | T | G | G | A | T | T | G | T | G | T | G | G | G | T | G | C | G | T | C | : | G | G |
| | MSL4 | A | A | G | G | T | G | C | G | G | G | T | T | G | T | G | T | G | G | G | T | G | C | G | T | C | : | G | A |
| | MSL6 | G | G | G | A | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | T | G | C | G | T | C | G | G | ? |
| | MSL6 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | A | C | G | C | G | T | C | : | G | ? |
| | MSL9 | G | G | G | A | G | G | C | G | A | A | T | T | G | T | G | T | G | C | A | T | G | C | G | T | C | G | G | G |
| | MSL9 | G | A | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| Rio Grande | REN19 | G | G | A | A | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| | REN19 | A | G | G | G | G | G | C | G | G | G | A | T | G | T | G | C | G | G | G | C | G | C | G | T | C | : | G | A |
| | REN24 | A | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| | REN24 | A | G | A | G | G | G | C | G | G | G | T | T | G | C | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| | REN3 | G | G | G | G | G | G | C | G | A | A | T | T | G | T | G | T | G | C | G | T | G | C | G | T | T | G | G | ? |
| | REN3 | G | A | G | G | G | G | T | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | ? |
| | REN6 | G | G | G | G | G | G | C | G | G | A | T | C | T | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | REN6 | G | G | G | G | G | G | C | G | G | G | T | T | G | T | G | C | G | G | G | C | G | C | G | T | C | : | G | G |
| | RKR19 | G | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | RKR19 | A | G | A | G | G | G | C | G | G | A | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| | RKW10 | G | G | G | G | G | G | C | G | G | A | T | T | G | C | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| | RKW10 | A | G | A | G | G | G | C | G | G | G | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| | RKW20 | G | G | G | G | G | G | C | G | G | G | T | T | G | T | G | C | G | G | G | C | A | C | G | T | C | : | G | G |
| | RKW20 | G | G | G | G | G | G | C | G | G | A | T | T | G | T | G | T | G | C | G | C | G | C | G | T | C | : | G | A |
| | RKW5 | G | G | G | G | G | G | C | G | G | G | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | RKW5 | A | G | A | G | G | G | C | G | G | G | T | T | G | T | G | C | G | G | G | C | G | C | G | T | C | : | G | A |
| | RKW6 | G | G | G | G | G | G | C | G | G | A | T | T | G | C | G | T | G | C | G | C | G | C | G | T | C | : | G | G |
| | RKW6 | A | G | A | G | G | G | C | G | G | G | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| | RKW8 | G | G | A | G | G | G | C | G | G | A | T | C | T | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | RKW8 | A | G | G | G | G | G | C | G | G | G | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | RKW9 | A | G | A | G | G | G | C | G | G | G | T | T | G | C | G | T | G | G | G | C | G | C | G | T | C | : | G | G |
| | RKW9 | A | G | A | G | G | G | C | G | G | G | T | T | G | T | G | T | G | G | G | C | G | C | G | T | C | : | G | A |
| Allele Frequency | A | 0.24 | 0.38 | 0.25 | 0.10 | 0.00 | 0.04 | 0.00 | 0.04 | 0.29 | 0.75 | 0.01 | 0.00 | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 0.08 | 0.00 | 0.03 | 0.00 | 0.01 | 0.03 | 0.00 | | 0.04 | 0.34 |
| | C | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.91 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.00 | 0.05 | 0.00 | 0.11 | 0.03 | 0.39 | 0.00 | 0.81 | 0.00 | 0.93 | 0.00 | 0.00 | 0.94 | | 0.00 | 0.00 |
| | G | 0.76 | 0.63 | 0.75 | 0.90 | 0.99 | 0.96 | 0.00 | 0.96 | 0.71 | 0.25 | 0.00 | 0.00 | 0.98 | 0.00 | 0.98 | 0.00 | 0.98 | 0.61 | 0.93 | 0.00 | 0.98 | 0.00 | 0.99 | 0.00 | 0.00 | | 0.96 | 0.66 |
| | T | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.09 | 0.00 | 0.00 | 0.00 | 0.99 | 0.98 | 0.03 | 0.95 | 0.00 | 0.89 | 0.00 | 0.00 | 0.00 | 0.19 | 0.00 | 0.08 | 0.00 | 0.98 | 0.06 | | 0.00 | 0.00 |
| | D | | | | | | | | | | | | | | | | | | | | | | | | | | 0.838 | | |
| | I | | | | | | | | | | | | | | | | | | | | | | | | | | 0.163 | | |

| | Locus | Bzfp2 | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 7576 | 7579 | 7582 | 7600 | 7655 | 7679 | 7724 | 7764 | 7805 | 7935 | 7961 | 7979 | 7987 | 7990 | 8005 | 8042 | 8077 | 8147 | 8190 | 8195 | 8245 | 8247 | 8250 | 8292 | 8361 |
| | Gene | N | N | N | N | N | N | N | N | N | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | E/S |
| | Alleles | W | R | M | Y | R | W | Y | R | R | C | 57 BP | K | Y | Y | R | Y | R | S | R | M | R | Y | Y | M | Y |
| Eastern | ECL1 | T | G | C | C | G | T | T | G | G | : | : | G | C | T | A | C | A | G | G | A | G | C | C | A | C |
| | ECL1 | T | G | C | C | G | T | T | G | G | : | : | G | C | T | A | C | A | G | G | A | G | C | C | A | C |
| | ECL10 | T | G | C | C | G | T | T | G | G | : | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | ECL10 | T | G | C | C | G | T | T | G | G | : | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | ECL3 | T | G | C | C | G | T | T | G | G | : | : | G | C | T | A | C | A | G | G | A | G | C | C | A | C |
| | ECL3 | T | G | C | C | G | T | T | G | G | : | : | G | C | T | A | C | A | G | G | A | G | C | C | A | C |
| | ECL5 | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | ECL5 | ? | ? | ? | ? | ? | ? | ? | ? | ? | : | : | G | C | T | A | C | A | G | G | A | G | C | C | A | C |
| | ECL6 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | ECL6 | T | G | C | C | G | T | T | G | G | : | : | G | C | T | A | C | A | G | G | A | G | C | T | A | C |
| | ECL8 | T | G | C | C | G | T | C | G | G | C | : | G | C | T | A | C | A | G | G | C | G | C | C | C | ? |
| | ECL8 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | ? |
| | EMC1 | T | G | C | C | G | T | C | G | G | C | : | G | C | T | A | C | G | G | G | C | G | C | T | C | ? |
| | EMC1 | T | G | A | T | A | T | T | G | G | C | : | G | C | T | A | T | G | G | G | A | G | T | C | C | ? |
| | EMC2 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | ? |
| | EMC2 | T | G | C | C | G | T | T | G | G | C | : | T | T | C | G | C | G | G | A | A | G | T | C | C | ? |
| | EOM3 | T | G | C | C | A | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | EOM3 | T | G | A | T | G | T | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | T | T | C | T |
| | EOM4 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | T |
| | EOM4 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | T |
| | EOM6 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | T |
| | EOM6 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | T |
| | EWW11 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | EWW11 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | EWW13 | ? | ? | ? | ? | ? | ? | ? | ? | ? | C | : | G | T | C | G | C | A | G | G | A | G | C | C | C | C |
| | EWW13 | ? | ? | ? | ? | ? | ? | ? | ? | ? | : | : | T | C | T | A | C | G | G | A | A | G | C | T | C | C |
| | EWW14 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | EWW14 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | A | C |
| | EWW4 | T | G | C | C | G | T | T | G | A | C | : | G | C | T | A | C | A | G | G | A | A | C | T | C | ? |
| | EWW4 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | T | C | ? |
| | MINN | T | G | C | C | G | T | T | G | A | C | : | G | C | T | A | T | A | G | G | A | G | C | T | C | C |
| | MINN | T | G | C | C | G | T | T | G | A | C | : | G | C | T | A | C | A | G | G | A | G | C | T | C | C |
| Merriam's | MCS1 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MCS1 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MCS3 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MCS3 | T | G | C | C | G | T | T | G | G | C | : | T | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MCS4 | T | A | C | C | G | T | T | A | G | C | : | G | C | T | A | C | A | G | G | A | G | C | T | C | C |
| | MCS4 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | MCS6 | T | G | C | C | G | T | T | G | G | C | : | T | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MCS6 | T | G | C | C | G | T | T | G | G | C | : | T | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MCS7 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MCS7 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MGD10 | T | A | C | C | G | T | T | A | G | C | : | G | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | MGD10 | T | A | C | C | G | T | T | A | G | C | : | G | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | MGD8 | T | A | C | C | G | T | T | A | G | C | : | G | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | MGD8 | T | A | C | C | G | T | T | A | G | C | : | G | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | MGD9 | A | G | C | C | G | A | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | C |
| | MGD9 | A | G | C | C | G | A | T | G | G | C | : | G | C | T | A | C | G | G | A | A | G | C | C | C | T |
| | MSL11 | T | A | C | C | G | T | T | A | G | C | 57 | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MSL11 | T | A | C | C | G | T | T | A | G | C | 57 | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MSL12 | T | A | C | C | G | T | T | A | G | C | 57 | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MSL12 | T | A | C | C | G | T | T | A | G | C | 57 | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MSL4 | T | A | C | C | G | T | T | A | G | C | : | G | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | MSL4 | T | A | C | C | G | T | T | A | G | C | : | G | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | MSL6 | T | A | C | C | G | T | T | A | G | C | 57 | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MSL6 | T | A | C | C | G | T | T | A | G | C | : | T | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | MSL9 | T | A | C | C | G | T | T | A | G | C | 57 | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | MSL9 | T | A | C | C | G | T | T | A | G | C | 57 | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| Rio Grande | REN19 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | A | G | A | A | G | C | C | C | T |
| | REN19 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | T |
| | REN24 | T | G | A | T | A | A | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | T |
| | REN24 | T | G | A | T | A | A | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | T |
| | REN3 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | REN3 | T | G | C | C | G | T | T | G | G | C | : | T | C | T | A | C | A | C | A | A | G | C | C | C | C |
| | REN6 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | C | G | A | G | C | C | C | C |
| | REN6 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | C | A | A | G | C | C | C | C |
| | RKR19 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | T |
| | RKR19 | T | G | A | T | A | A | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | T |
| | RKW10 | T | G | C | C | A | T | T | G | A | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | T |
| | RKW10 | T | G | A | T | G | T | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | T | C | C |
| | RKW20 | T | G | C | C | A | T | T | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | C |
| | RKW20 | T | G | A | T | G | T | C | G | G | C | : | G | C | T | A | C | G | C | G | A | G | C | C | C | T |
| | RKW5 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | A | G | G | C | G | C | T | C | C |
| | RKW5 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | A | G | G | C | G | C | T | C | C |
| | RKW6 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | C | A | A | G | C | C | C | C |
| | RKW6 | T | G | C | C | G | T | T | G | G | C | : | G | C | T | A | C | A | C | A | A | G | C | C | C | C |
| | RKW8 | T | G | C | C | A | T | T | G | G | C | : | G | C | T | A | C | A | C | G | A | G | C | C | C | C |
| | RKW8 | T | G | A | T | G | T | C | G | G | C | : | G | C | T | A | C | A | G | A | A | G | C | C | C | T |
| | RKW9 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | A | G | G | A | G | C | C | C | T |
| | RKW9 | T | G | A | T | A | T | C | G | G | C | : | G | C | T | A | C | G | G | G | A | G | C | C | C | T |
| Allele Frequency | A | 0.03 | 0.20 | 0.25 | 0.00 | 0.25 | 0.07 | 0.00 | 0.20 | 0.05 | 0.00 | | 0.00 | 0.00 | 0.00 | 0.98 | 0.00 | 0.69 | 0.00 | 0.11 | 0.95 | 0.01 | 0.00 | 0.00 | 0.09 | 0.00 |
| | C | 0.00 | 0.00 | 0.75 | 0.75 | 0.00 | 0.00 | 0.28 | 0.00 | 0.00 | 1.00 | | 0.00 | 0.98 | 0.03 | 0.00 | 0.98 | 0.00 | 0.09 | 0.00 | 0.05 | 0.00 | 0.96 | 0.75 | 0.91 | 0.76 |
| | G | 0.00 | 0.80 | 0.00 | 0.00 | 0.75 | 0.00 | 0.00 | 0.80 | 0.95 | 0.00 | | 0.91 | 0.00 | 0.00 | 0.03 | 0.00 | 0.31 | 0.91 | 0.89 | 0.00 | 0.99 | 0.00 | 0.00 | 0.00 | 0.00 |
| | T | 0.97 | 0.00 | 0.00 | 0.25 | 0.00 | 0.93 | 0.72 | 0.00 | 0.00 | 0.00 | | 0.09 | 0.03 | 0.98 | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.25 | 0.00 | 0.24 |
| | D | | | | | | | | | | | 0.913 | | | | | | | | | | | | | | |
| | I | | | | | | | | | | | 0.088 | | | | | | | | | | | | | | |

| | Locus | LAAO | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 28905 | 28942 | 28964 | 29021 | 29041 | 29059 | 29068 | 29216 | 29218 | 29219 | 29239 | 29257 | 29281 | 29323 | 29347 | 29359 | 29440 | 29449 | 29473 | 29581 |
| | Gene | I | I | I | I | E/S | E/S | E/S | E/N | E/N | E/N | E/S | E/S | E/S | E/S | E/S | E/S | E/S | E/S | E/S | E/S |
| | Alleles | Y | M | W | M | Y | Y | R | R | Y | R | R | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| **Eastern** | **ECL1** | C | A | T | C | T | C | G | G | C | A | A | C | C | C | T | C | C | T | C | C |
| | **ECL1** | C | A | T | C | C | C | G | G | C | G | G | C | C | C | T | C | C | T | T | C |
| | **ECL10** | C | A | T | C | T | C | G | G | C | G | A | C | C | C | T | T | C | T | C | C |
| | **ECL10** | C | A | T | C | T | C | G | G | C | G | A | C | C | C | T | T | C | T | C | C |
| | **ECL3** | C | A | T | C | C | C | G | G | C | A | A | C | C | C | T | C | C | T | C | C |
| | **ECL3** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | T | C |
| | **ECL5** | C | A | T | C | T | C | A | G | C | A | A | C | C | C | T | C | C | T | T | C |
| | **ECL5** | C | A | T | C | T | C | G | G | C | G | A | C | C | C | T | C | C | T | T | C |
| | **ECL6** | C | A | T | C | T | C | G | G | C | A | A | C | C | C | T | C | C | T | T | C |
| | **ECL6** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | T | C |
| | **ECL8** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | C | C |
| | **ECL8** | C | A | T | C | C | C | G | G | C | G | A | C | C | C | T | C | C | T | T | C |
| | **EMC1** | C | A | T | C | C | C | G | G | C | G | G | T | C | C | C | C | C | C | C | C |
| | **EMC1** | C | A | T | C | C | C | G | G | C | G | G | T | C | C | T | T | T | C | C | C |
| | **EMC2** | C | A | T | C | C | C | G | G | C | G | G | T | C | C | C | C | C | C | C | C |
| | **EMC2** | C | A | T | C | C | C | G | G | C | G | G | T | C | C | T | T | T | C | C | C |
| | **EOM3** | C | A | T | C | C | C | G | G | C | A | A | C | T | C | T | C | C | C | C | C |
| | **EOM3** | C | C | T | C | T | C | G | G | C | G | G | C | C | C | T | C | T | T | T | C |
| | **EOM4** | C | A | T | C | C | C | G | G | C | A | G | C | C | C | C | C | C | C | C | C |
| | **EOM4** | C | A | T | C | C | C | G | G | C | G | G | T | C | C | T | C | T | C | C | C |
| | **EOM6** | C | A | A | C | T | C | G | G | C | G | G | C | C | T | T | T | C | C | C | C |
| | **EOM6** | C | A | T | C | C | C | G | G | C | G | G | C | C | C | T | C | T | T | C | C |
| | **EWW11** | C | A | T | C | T | C | G | G | C | G | A | C | C | C | T | C | C | T | C | T |
| | **EWW11** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | T | C |
| | **EWW13** | C | A | T | C | C | C | G | G | C | G | A | C | C | C | T | C | C | T | C | C |
| | **EWW13** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | T | C |
| | **EWW14** | C | A | T | A | T | C | G | G | C | A | A | C | C | C | T | C | C | T | T | C |
| | **EWW14** | C | A | T | C | T | C | G | G | C | A | A | C | C | C | T | C | C | T | T | C |
| | **EWW4** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | C | C |
| | **EWW4** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | C | C |
| **Merriam's** | **MINN** | C | A | T | C | C | C | G | G | T | G | G | C | C | C | T | C | C | T | C | C |
| | **MINN** | C | A | T | C | C | C | G | G | T | G | G | C | C | C | T | C | C | T | C | C |
| | **MCS1** | C | A | T | C | C | C | G | G | T | A | G | C | C | C | C | C | C | C | C | C |
| | **MCS1** | C | A | T | C | C | C | G | G | C | G | G | C | C | C | T | C | C | T | C | C |
| | **MCS3** | C | A | T | C | C | C | G | G | T | G | G | C | C | C | T | C | C | T | C | C |
| | **MCS3** | C | A | T | C | C | C | G | G | T | G | G | C | C | C | T | C | C | T | C | C |
| | **MCS4** | C | A | T | C | T | C | A | G | C | G | G | C | C | C | T | C | C | T | C | C |
| | **MCS4** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | C | C |
| | **MCS6** | C | A | T | C | C | C | G | G | C | G | A | C | C | C | T | C | C | T | C | C |
| | **MCS6** | C | A | T | C | C | C | G | G | C | G | A | C | C | C | T | C | C | T | C | C |
| | **MCS7** | C | A | T | C | C | C | G | G | T | G | G | C | C | C | T | C | C | T | C | C |
| | **MCS7** | C | A | T | C | C | C | G | G | T | G | G | C | C | C | T | C | C | T | C | C |
| | **MGD10** | C | A | T | C | T | C | A | G | C | A | G | C | C | C | C | C | C | C | C | C |
| | **MGD10** | C | A | T | C | C | C | G | G | C | G | G | C | C | C | T | C | C | T | C | C |
| | **MGD8** | C | A | T | C | C | C | A | G | C | A | G | C | C | C | C | C | C | C | C | C |
| | **MGD8** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | C | T | C | C |
| | **MGD9** | C | A | T | C | C | C | G | G | C | A | G | C | C | C | C | C | C | C | C | C |
| | **MGD9** | C | A | T | C | T | C | G | G | C | A | G | T | C | C | C | C | C | C | C | C |
| | **MSL11** | C | A | T | C | C | T | A | G | C | A | G | C | C | C | T | T | C | C | C | C |
| | **MSL11** | C | C | T | C | T | C | G | G | C | G | G | T | C | C | T | C | C | T | C | C |
| | **MSL12** | C | A | A | C | T | T | A | G | C | G | G | C | C | T | T | T | C | T | C | C |
| | **MSL12** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | C | T | C | C | C |
| | **MSL4** | C | A | A | C | C | C | G | G | C | A | G | C | C | T | T | T | C | C | C | C |
| | **MSL4** | C | C | T | C | T | C | G | G | C | G | G | T | C | C | T | T | T | C | C | C |
| | **MSL6** | C | A | T | C | C | T | A | G | C | A | G | C | C | C | T | C | C | C | C | C |
| | **MSL6** | C | C | T | C | T | C | G | G | C | G | G | T | C | C | T | T | C | T | C | C |
| | **MSL9** | C | C | T | C | C | C | G | G | C | A | G | C | C | C | T | C | C | C | C | C |
| | **MSL9** | C | C | T | C | T | C | G | G | C | G | G | T | C | C | T | T | C | T | C | C |
| **Rio Grande** | **REN19** | C | A | T | C | C | C | G | G | C | A | G | C | C | C | C | C | C | C | C | C |
| | **REN19** | C | C | T | C | C | C | G | G | C | G | G | T | C | C | T | T | C | C | C | C |
| | **REN24** | C | A | A | C | T | C | G | G | C | A | G | C | C | T | C | T | C | C | C | C |
| | **REN24** | C | A | T | C | C | C | G | G | C | G | G | C | C | C | T | C | T | C | C | C |
| | **REN3** | C | A | T | C | T | C | A | G | C | G | G | C | C | C | T | C | C | T | C | C |
| | **REN3** | C | A | T | C | C | C | G | G | T | G | G | C | C | C | T | C | C | T | C | C |
| | **REN6** | C | A | T | C | C | C | G | G | C | A | G | C | C | C | T | C | C | C | C | C |
| | **REN6** | C | C | T | C | C | C | G | G | C | G | G | T | C | C | T | T | C | T | C | C |
| | **RKR19** | C | A | A | C | C | C | G | G | C | A | G | C | C | T | T | T | C | C | C | C |
| | **RKR19** | C | C | T | C | T | C | G | G | C | G | G | T | C | C | T | T | T | C | C | C |
| | **RKW10** | C | A | T | C | C | C | G | G | C | A | G | C | C | C | T | C | C | C | C | C |
| | **RKW10** | C | C | T | C | C | C | G | G | C | G | G | T | C | C | T | T | C | T | C | C |
| | **RKW20** | C | A | T | C | C | C | G | G | C | G | G | T | C | T | T | T | T | C | C | C |
| | **RKW20** | C | A | T | C | C | C | G | G | C | G | G | T | C | T | T | T | T | T | C | C |
| | **RKW5** | C | A | A | C | T | C | G | G | C | A | G | C | C | T | T | T | C | C | C | C |
| | **RKW5** | C | A | T | C | C | C | G | G | C | G | G | T | C | C | T | T | T | C | C | C |
| | **RKW6** | C | A | A | C | T | C | A | A | C | G | G | C | C | T | T | C | C | C | C | C |
| | **RKW6** | C | A | T | C | T | C | G | G | C | G | G | C | C | C | T | T | T | T | C | C |
| | **RKW8** | T | A | T | C | C | C | G | G | C | A | G | C | C | C | T | T | C | C | C | C |
| | **RKW8** | C | C | T | C | C | C | G | G | C | G | G | T | C | C | T | T | C | T | C | C |
| | **RKW9** | C | A | T | C | C | C | G | G | C | A | G | C | C | C | C | C | C | C | C | C |
| | **RKW9** | C | A | T | C | C | C | G | G | C | A | G | T | C | C | T | T | C | C | C | C |
| **Allele Frequency** | **A** | 0.00 | 0.86 | 0.09 | 0.01 | 0.00 | 0.00 | 0.11 | 0.01 | 0.00 | 0.33 | 0.19 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | **C** | 0.99 | 0.14 | 0.00 | 0.99 | 0.56 | 0.96 | 0.00 | 0.00 | 0.90 | 0.00 | 0.00 | 0.76 | 0.99 | 0.89 | 0.14 | 0.69 | 0.84 | 0.43 | 0.85 | 0.99 |
| | **G** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.89 | 0.99 | 0.00 | 0.68 | 0.81 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | **T** | 0.01 | 0.00 | 0.91 | 0.00 | 0.44 | 0.04 | 0.00 | 0.00 | 0.10 | 0.00 | 0.00 | 0.24 | 0.01 | 0.11 | 0.86 | 0.31 | 0.16 | 0.58 | 0.15 | 0.01 |
| | **D** | | | | | | | | | | | | | | | | | | | | |
| | **I** | | | | | | | | | | | | | | | | | | | | |

| | Locus | TRIM 39.2 | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 56097 | 56106 | 56210 | 56220 | 56238 | 56256 | 56259 | 56269 | 56416 | 56474 | 56481 | 56492 | 56559 | 56566 | 56622 | 56663 | 56670 | 56676 | 56694 | 56710 | 56726 | 56742 | 56788 | 56806 | 56840 |
| | Gene | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I |
| | Alleles | R | Y | H | W | R | Y | R | M | M | R | R | Y | R | R | Y | R | R | R | Y | R | Y | R | Y | W | R |
| Eastern | ECL1 | A | C | C | A | A | C | G | C | A | A | G | C | G | A | C | A | A | G | C | A | T | A | C | A | A |
| | ECL1 | A | C | C | A | A | T | G | A | C | A | G | C | G | G | C | A | G | G | C | G | T | G | C | A | G |
| | ECL10 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | ECL10 | A | C | C | A | A | T | G | C | C | A | G | C | G | G | C | G | G | G | C | A | T | A | C | A | A |
| | ECL3 | A | C | C | A | A | C | G | A | A | A | G | C | G | A | C | A | A | G | C | A | T | A | C | A | G |
| | ECL3 | A | C | C | A | A | T | G | C | A | A | G | C | G | G | C | A | G | G | C | G | T | G | C | A | G |
| | ECL5 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | G |
| | ECL5 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | A | G | G | C | A | T | G | C | T | G |
| | ECL6 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | G |
| | ECL6 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | G | G | G | C | A | T | A | C | A | G |
| | ECL8 | A | C | C | A | A | C | G | C | A | A | G | C | G | A | C | A | A | A | C | A | T | G | C | A | G |
| | ECL8 | A | C | C | A | A | T | G | C | C | A | G | C | G | G | C | A | G | G | T | A | T | G | C | T | G |
| | EMC1 | A | C | C | A | A | T | G | C | A | A | G | C | G | A | C | A | A | A | C | A | T | G | C | A | G |
| | EMC1 | A | C | C | A | A | T | G | C | A | A | G | C | G | A | C | A | G | G | T | G | T | G | C | A | G |
| | EMC2 | A | C | C | A | A | T | G | C | A | A | G | C | G | A | C | A | A | G | C | G | T | G | C | A | G |
| | EMC2 | A | C | A | A | A | T | G | C | A | A | G | C | G | A | C | A | G | G | C | G | T | G | C | A | G |
| | EOM3 | A | C | C | A | A | C | G | C | A | A | G | C | G | A | C | A | A | G | C | A | T | A | C | A | A |
| | EOM3 | A | C | A | A | A | T | G | C | A | A | G | C | G | G | C | A | G | G | C | G | T | G | T | A | G |
| | EOM4 | A | C | C | A | A | T | A | C | A | G | G | C | G | A | C | A | G | G | C | A | T | G | C | A | G |
| | EOM4 | G | T | T | T | A | T | G | C | A | A | G | C | G | G | C | A | G | G | C | G | T | G | T | A | G |
| | EOM6 | A | C | C | A | A | C | G | C | A | A | G | C | G | A | C | A | G | A | C | A | T | G | C | A | G |
| | EOM6 | A | C | C | T | G | T | G | C | A | A | G | C | G | A | C | A | G | G | C | G | T | G | C | A | G |
| | EWW11 | A | C | C | A | A | C | G | C | A | A | A | C | G | A | C | A | G | G | C | A | T | A | C | A | G |
| | EWW11 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | A | G | G | C | A | T | G | C | T | G |
| | EWW13 | A | C | C | A | A | T | G | A | A | A | A | C | G | A | C | A | A | G | C | A | C | G | C | A | G |
| | EWW13 | A | C | C | A | A | C | G | C | A | A | G | C | G | A | C | A | G | G | C | G | T | G | C | A | G |
| | EWW14 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | EWW14 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | G | G | G | C | A | T | A | C | A | A |
| | EWW4 | A | C | C | A | A | C | G | C | A | A | A | C | G | A | C | A | G | G | C | A | T | A | C | A | G |
| | EWW4 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | A | G | G | C | A | T | G | C | T | G |
| Merriam's | MINN | A | C | C | A | A | C | G | C | A | A | G | C | A | G | C | A | G | G | C | A | T | A | C | A | A |
| | MINN | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MCS1 | A | C | C | A | A | C | G | C | A | A | G | T | G | A | C | A | A | A | C | A | T | A | C | A | G |
| | MCS1 | A | C | C | A | A | T | G | C | A | A | G | C | G | G | C | A | G | G | T | A | T | G | C | A | G |
| | MCS3 | A | C | C | A | A | C | G | C | A | A | G | C | A | G | C | A | G | G | C | A | T | A | C | A | A |
| | MCS3 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | G |
| | MCS4 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MCS4 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | G | G | G | C | A | T | A | C | A | G |
| | MCS6 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | G | G | G | C | A | T | A | C | A | A |
| | MCS6 | A | C | C | A | A | C | G | C | C | A | G | C | G | G | C | G | G | G | C | A | T | A | C | A | A |
| | MCS7 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | G |
| | MCS7 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | G |
| | MGD10 | A | C | C | A | A | C | G | C | A | A | G | T | G | A | C | A | A | A | C | A | T | A | C | A | A |
| | MGD10 | A | C | C | A | A | T | G | C | A | A | G | C | G | G | C | A | G | G | T | A | T | G | C | A | G |
| | MGD8 | A | C | C | A | A | C | G | C | A | A | G | C | G | A | C | A | A | A | C | A | T | A | C | A | A |
| | MGD8 | A | C | C | A | A | T | G | C | A | A | G | T | G | G | C | A | G | G | T | A | T | G | C | A | G |
| | MGD9 | A | C | C | A | A | T | G | C | A | A | G | T | G | A | C | A | A | A | C | A | T | G | C | A | G |
| | MGD9 | A | C | A | A | A | T | G | C | A | A | G | C | G | A | C | A | A | G | T | G | T | G | C | A | G |
| | MSL11 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MSL11 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MSL12 | A | C | C | A | A | C | A | C | A | G | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MSL12 | G | T | T | T | A | T | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | G | T | A | G |
| | MSL4 | A | C | C | A | A | C | A | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MSL4 | G | T | T | T | A | T | G | C | A | G | G | C | G | G | C | A | G | G | C | A | T | G | T | A | G |
| | MSL6 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MSL6 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MSL9 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | MSL9 | A | C | C | A | A | C | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| Rio Grande | REN19 | A | C | C | A | A | T | A | A | A | G | G | C | G | A | C | A | G | G | C | A | T | G | C | A | G |
| | REN19 | G | T | T | T | A | T | G | C | A | A | G | C | G | G | C | A | G | G | C | G | T | G | T | A | G |
| | REN24 | A | C | C | A | A | T | G | C | A | A | G | C | G | A | C | A | G | G | C | G | T | G | C | A | G |
| | REN24 | A | C | C | A | A | T | G | C | A | A | G | C | G | A | C | A | G | G | C | G | T | G | C | A | G |
| | REN3 | A | C | C | A | A | C | G | C | A | A | G | C | G | A | C | A | G | G | C | A | T | A | C | A | A |
| | REN3 | A | C | C | A | A | T | G | C | A | A | G | C | G | G | C | A | G | G | C | A | T | A | C | A | A |
| | REN6 | A | C | C | A | A | C | G | C | A | A | G | C | G | A | C | A | A | A | C | A | T | A | C | A | G |
| | REN6 | A | C | C | A | A | T | G | C | A | A | G | T | G | G | C | A | G | G | T | A | T | G | C | A | G |
| | RKR19 | A | C | C | A | G | C | G | C | A | A | G | C | G | A | C | A | G | A | C | A | T | G | C | A | G |
| | RKR19 | A | C | C | T | A | T | G | C | A | A | G | C | G | A | C | A | G | G | C | G | T | G | C | A | G |
| | RKW10 | A | C | C | A | A | C | A | C | A | G | G | C | G | G | T | A | A | G | C | A | T | A | C | A | A |
| | RKW10 | G | T | T | T | A | T | G | C | C | A | G | C | G | G | C | A | G | G | C | G | T | G | C | A | G |
| | RKW20 | A | C | C | A | A | C | A | C | A | A | G | C | A | G | C | A | G | G | C | A | T | A | C | A | A |
| | RKW20 | G | T | T | T | A | T | G | C | A | G | G | C | G | G | C | A | G | G | C | A | T | G | T | A | G |
| | RKW5 | A | C | C | A | A | T | G | C | A | A | G | C | G | A | C | A | A | A | C | A | T | G | C | A | G |
| | RKW5 | A | C | C | A | A | T | G | C | A | A | G | T | G | A | C | A | G | G | T | G | T | G | C | A | G |
| | RKW6 | A | C | C | A | A | C | A | C | A | G | G | C | G | A | T | A | A | G | C | A | T | A | C | A | A |
| | RKW6 | G | T | T | T | A | T | G | C | A | A | A | C | G | G | C | A | G | G | C | G | T | G | C | A | G |
| | RKW8 | A | C | C | A | A | C | G | C | A | A | A | C | G | A | C | A | G | A | C | A | T | A | C | A | A |
| | RKW8 | A | C | C | A | A | T | G | C | A | A | G | T | G | A | C | A | A | G | T | A | T | G | C | A | G |
| | RKW9 | A | C | C | A | A | T | G | C | A | A | G | T | G | A | C | A | A | A | C | A | T | A | C | A | G |
| | RKW9 | A | C | C | A | A | T | G | C | A | A | G | C | G | A | C | A | G | G | T | G | T | G | C | A | G |
| Allele Frequency | A | 0.91 | 0.00 | 0.04 | 0.89 | 0.98 | 0.00 | 0.09 | 0.05 | 0.81 | 0.91 | 0.06 | 0.00 | 0.04 | 0.43 | 0.00 | 0.93 | 0.23 | 0.15 | 0.00 | 0.78 | 0.00 | 0.51 | 0.00 | 0.95 | 0.35 |
| | C | 0.00 | 0.91 | 0.88 | 0.00 | 0.00 | 0.56 | 0.00 | 0.95 | 0.19 | 0.00 | 0.00 | 0.90 | 0.00 | 0.00 | 0.98 | 0.00 | 0.00 | 0.00 | 0.88 | 0.00 | 0.01 | 0.00 | 0.93 | 0.00 | 0.00 |
| | G | 0.09 | 0.00 | 0.00 | 0.00 | 0.03 | 0.00 | 0.91 | 0.00 | 0.00 | 0.09 | 0.94 | 0.00 | 0.96 | 0.58 | 0.00 | 0.08 | 0.78 | 0.85 | 0.00 | 0.23 | 0.00 | 0.49 | 0.00 | 0.00 | 0.65 |
| | T | 0.00 | 0.09 | 0.09 | 0.11 | 0.00 | 0.44 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.10 | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 0.13 | 0.00 | 0.99 | 0.00 | 0.08 | 0.05 | 0.00 |
| | D | | | | | | | | | | | | | | | | | | | | | | | | | |
| | I | | | | | | | | | | | | | | | | | | | | | | | | | |

| | Locus | TRIM 39.1 | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 66829 | 66845 | 66875 | 66904 | 66927 | 67001 | 67039 | 67049 | 67099 | 67142 | 67151 | 67233 | 67255 | 67333 | 67364 | 67427 | 67448 | 67503 | 67518 | 67532 |
| | Gene | I | I | I | I | I | I | I | I | I | I | I | I | I | E/S | E/N | I | I | I | I | I |
| | Alleles | R | R | W | Y | R | S | Y | S | Y | W | W | Y | S | M | Y | Y | R | Y | Y | K |
| Eastern | ECL1 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL1 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL10 | ? | ? | ? | ? | A | G | C | C | C | A | T | C | C | A | C | T | G | C | C | T |
| | ECL10 | ? | ? | ? | ? | A | G | C | C | C | A | T | C | C | A | C | T | G | C | C | T |
| | ECL3 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL3 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL5 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL5 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL6 | ? | ? | ? | ? | A | G | C | G | C | A | A | C | C | A | C | T | G | C | C | T |
| | ECL6 | ? | ? | ? | ? | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | ECL8 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL8 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EMC1 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EMC1 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EMC2 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EMC2 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EOM3 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EOM3 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EOM4 | ? | ? | T | C | A | C | C | G | C | A | A | C | C | A | C | C | A | T | T | G |
| | EOM4 | ? | ? | T | C | G | C | T | G | C | A | A | C | G | C | C | C | G | C | T | G |
| | EOM6 | A | G | A | C | A | C | T | G | C | A | A | C | C | A | C | C | A | T | T | G |
| | EOM6 | G | G | T | C | A | C | T | G | C | A | A | C | G | C | T | C | G | C | T | G |
| | EWW11 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EWW11 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EWW13 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EWW13 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EWW14 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EWW14 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EWW4 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EWW4 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| Merriam's | MINN | A | G | A | C | A | G | C | G | C | A | T | C | C | A | C | C | G | C | C | T |
| | MINN | A | G | T | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | MCS1 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MCS1 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MCS3 | A | G | A | C | A | G | C | G | C | A | T | C | C | A | C | C | G | C | C | T |
| | MCS3 | A | G | T | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | MCS4 | ? | ? | T | C | A | G | C | G | C | A | T | C | C | A | C | C | G | C | C | T |
| | MCS4 | ? | ? | T | T | A | G | C | G | T | T | T | C | C | A | C | C | G | C | C | T |
| | MCS6 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MCS6 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MCS7 | A | G | T | C | A | G | C | G | C | A | T | C | C | A | C | C | G | C | C | T |
| | MCS7 | A | G | T | C | A | G | C | G | C | A | T | C | C | A | C | C | G | C | C | T |
| | MGD10 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MGD10 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MGD8 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MGD8 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MGD9 | A | G | A | C | A | C | C | G | C | A | A | C | C | C | C | C | G | C | T | G |
| | MGD9 | G | G | T | C | G | C | C | G | C | A | A | C | C | C | T | C | G | C | T | G |
| | MSL11 | A | A | A | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | MSL11 | A | G | T | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | MSL12 | A | A | A | C | A | C | C | G | C | A | A | C | C | A | C | C | G | C | T | T |
| | MSL12 | A | G | T | C | A | G | T | G | C | A | T | C | C | C | T | T | G | C | C | G |
| | MSL4 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MSL4 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MSL6 | A | A | A | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | MSL6 | A | G | T | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | MSL9 | A | A | T | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | MSL9 | A | A | T | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| Rio Grande | REN19 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | REN19 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | REN24 | A | G | A | C | A | C | C | G | C | A | A | C | C | C | C | C | G | C | T | G |
| | REN24 | A | G | A | C | G | C | C | G | C | A | A | C | C | C | T | C | G | C | T | G |
| | REN3 | A | A | T | C | A | G | C | G | C | A | T | C | C | A | C | C | G | C | C | T |
| | REN3 | A | G | T | C | A | G | C | G | C | A | T | C | C | A | C | T | G | C | C | T |
| | REN6 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | REN6 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKR19 | A | G | A | C | A | C | C | G | C | A | A | C | C | C | C | C | G | C | T | G |
| | RKR19 | A | G | A | C | G | C | T | G | C | A | A | T | C | C | T | C | G | C | T | G |
| | RKW10 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW10 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW20 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW20 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW5 | A | G | A | C | A | C | C | G | C | A | A | C | C | A | C | C | A | C | T | G |
| | RKW5 | A | G | A | C | A | C | T | G | C | A | A | C | G | C | T | C | G | C | T | G |
| | RKW6 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW6 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW8 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW8 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW9 | A | G | A | C | A | C | C | G | C | A | A | C | C | C | T | C | A | C | T | G |
| | RKW9 | A | G | A | C | A | C | C | G | C | A | A | C | C | C | T | C | G | C | T | G |
| Allele Frequency | A | 0.93 | 0.21 | 0.47 | 0.00 | 0.89 | 0.00 | 0.00 | 0.00 | 0.00 | 0.97 | 0.44 | 0.00 | 0.00 | 0.67 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.00 |
| | C | 0.00 | 0.00 | 0.00 | 0.97 | 0.00 | 0.42 | 0.83 | 0.06 | 0.97 | 0.00 | 0.00 | 0.97 | 0.92 | 0.33 | 0.78 | 0.61 | 0.00 | 0.94 | 0.58 | 0.00 |
| | G | 0.07 | 0.79 | 0.00 | 0.00 | 0.11 | 0.58 | 0.00 | 0.94 | 0.00 | 0.00 | 0.00 | 0.00 | 0.08 | 0.00 | 0.00 | 0.00 | 0.89 | 0.00 | 0.00 | 0.42 |
| | T | 0.00 | 0.00 | 0.53 | 0.03 | 0.00 | 0.00 | 0.17 | 0.00 | 0.03 | 0.03 | 0.56 | 0.03 | 0.00 | 0.00 | 0.22 | 0.39 | 0.00 | 0.06 | 0.42 | 0.58 |
| | D | | | | | | | | | | | | | | | | | | | | |
| | I | | | | | | | | | | | | | | | | | | | | |

| | Locus | TRIM 41 | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 77449 | 77469 | 77477 | 77491 | 77534 | 77545 | 77556 | 77625 | 77661 | 77663 | 77678 | 77726 | 77748 | 77804 | 77806 | 77833 | 77867 | 77993 | 78036 | 78068 | 78103 | 78256 | 78283 | 78368 | 78434 |
| | Gene | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | I | E/S | I | I | I | I | I | I | I | I |
| | Alleles | Y | W | R | R | R | G | R | Y | S | Y | R | Y | R | Y | Y | Y | R | R | Y | R | Y | Y | R | R | Y |
| Eastern | ECL1 | C | A | G | G | A | G | G | C | G | C | G | T | G | C | C | C | G | A | C | G | T | C | A | G | C |
| | ECL1 | C | T | G | G | G | : | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | T | A | G | C |
| | ECL10 | C | A | G | G | A | G | A | C | G | C | G | T | G | T | C | C | G | A | C | G | T | T | A | G | C |
| | ECL10 | C | A | G | G | A | G | A | C | G | C | G | T | G | T | C | C | G | A | C | G | T | T | A | G | C |
| | ECL3 | C | T | G | G | G | : | G | C | G | C | G | T | G | C | C | C | G | A | C | G | T | T | A | G | C |
| | ECL3 | C | T | G | G | G | : | G | C | G | C | G | T | G | C | C | C | G | A | T | G | T | C | A | G | C |
| | ECL5 | C | A | G | G | A | G | G | C | G | C | A | T | G | C | C | C | G | A | T | G | T | C | A | A | C |
| | ECL5 | C | T | G | G | G | : | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | ECL6 | C | A | G | G | A | G | ? | C | G | C | A | T | G | C | C | C | G | A | C | G | T | T | A | G | C |
| | ECL6 | C | T | G | G | G | : | ? | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | ECL8 | C | A | A | G | A | G | G | C | G | C | A | T | G | T | C | C | G | A | T | G | T | T | A | A | C |
| | ECL8 | C | T | G | G | G | : | G | C | G | C | G | C | G | C | C | C | G | G | T | G | T | C | A | G | C |
| | EMC1 | C | T | A | G | G | G | G | C | G | C | A | C | A | C | C | C | G | A | C | G | T | T | A | G | C |
| | EMC1 | T | A | G | G | G | : | G | C | G | C | A | C | A | C | C | C | G | G | T | G | T | T | A | G | T |
| | EMC2 | C | T | A | G | G | G | G | C | G | C | A | C | A | C | C | C | G | A | C | G | T | T | A | G | C |
| | EMC2 | T | A | G | G | G | : | G | C | G | C | A | C | G | C | C | C | G | G | T | G | T | T | A | G | T |
| | EOM3 | C | A | A | G | G | G | G | C | G | C | A | C | G | C | C | C | G | A | T | G | T | T | A | G | C |
| | EOM3 | C | T | G | G | G | : | G | C | G | C | G | T | A | T | C | C | G | G | T | G | T | C | A | G | C |
| | EOM4 | C | T | A | G | G | : | G | C | G | C | A | C | G | C | C | C | G | G | C | A | T | T | A | G | C |
| | EOM4 | C | T | G | G | G | : | G | C | G | C | G | T | G | C | C | C | G | G | T | G | C | T | A | G | C |
| | EOM6 | C | T | A | G | G | : | G | C | C | C | A | C | G | C | C | C | G | A | C | G | T | T | A | G | T |
| | EOM6 | C | T | G | G | G | : | G | T | G | C | A | C | G | C | C | T | G | G | T | G | T | T | G | G | C |
| | EWW11 | ? | A | G | G | A | G | G | C | G | C | A | T | G | T | C | C | G | A | C | G | T | T | A | G | C |
| | EWW11 | ? | A | G | G | A | G | G | C | G | C | G | T | G | T | T | C | G | A | T | G | T | C | A | G | C |
| | EWW13 | C | A | A | G | A | G | G | C | G | C | A | C | A | C | C | C | G | A | T | G | T | T | A | G | C |
| | EWW13 | C | T | G | G | G | : | G | C | G | C | G | T | G | T | C | C | G | G | T | G | T | T | A | G | T |
| | EWW14 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | A | A | T | G | T | C | A | G | C |
| | EWW14 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | A | A | T | G | T | C | A | G | C |
| | EWW4 | C | A | G | A | A | G | G | C | G | C | G | T | G | T | C | C | G | A | C | G | C | T | A | G | C |
| | EWW4 | C | A | G | G | A | G | A | C | G | C | G | T | G | T | C | C | G | A | T | G | T | T | A | G | C |
| | MINN | ? | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | C | G | T | T | A | G | C |
| | MINN | ? | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | C | G | T | T | A | G | C |
| Merriam's | MCS1 | C | A | G | G | A | G | G | C | G | C | G | T | G | C | C | C | G | A | T | G | T | T | A | G | C |
| | MCS1 | T | A | G | G | G | G | G | C | G | C | G | C | A | T | C | C | G | G | T | G | T | C | A | G | C |
| | MCS3 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MCS3 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MCS4 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | A | C |
| | MCS4 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MCS6 | ? | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MCS6 | ? | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MCS7 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MCS7 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MGD10 | C | A | G | G | A | G | G | C | G | C | A | C | A | C | C | C | G | A | T | G | T | C | A | G | C |
| | MGD10 | T | A | G | G | G | G | G | C | G | C | G | T | G | T | C | C | G | G | T | G | T | T | A | G | C |
| | MGD8 | C | A | G | G | A | G | G | C | G | C | A | C | G | C | C | C | G | A | T | G | T | C | A | G | C |
| | MGD8 | T | A | G | G | G | G | G | C | G | C | G | T | A | T | C | C | G | G | T | G | T | T | A | G | C |
| | MGD9 | C | T | A | G | G | : | G | C | G | C | A | T | G | C | C | C | G | A | C | G | T | T | A | G | T |
| | MGD9 | C | T | G | G | G | : | G | C | G | C | G | T | G | C | C | C | G | G | C | G | C | T | A | G | C |
| | MSL11 | ? | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MSL11 | ? | T | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MSL12 | ? | T | A | G | A | G | G | C | G | C | G | C | G | C | C | C | G | A | T | G | T | C | A | G | C |
| | MSL12 | ? | T | G | G | G | : | G | C | G | C | A | T | G | T | C | C | G | G | T | G | T | T | A | G | C |
| | MSL4 | ? | T | A | G | A | G | G | C | G | C | G | C | A | C | C | C | G | A | T | G | T | C | A | G | C |
| | MSL4 | ? | T | G | G | G | : | G | C | G | C | G | T | G | T | C | C | G | G | T | G | T | T | A | G | C |
| | MSL6 | C | A | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MSL6 | C | T | G | G | A | G | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MSL9 | C | T | G | G | A | G | G | C | G | C | A | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | MSL9 | C | T | G | G | A | G | G | C | G | C | A | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| Rio Grande | REN19 | C | T | A | G | G | : | G | C | C | C | A | T | G | C | C | C | G | A | C | G | C | T | G | G | T |
| | REN19 | C | T | G | G | G | : | G | C | G | C | G | T | G | C | C | C | G | G | C | G | T | T | G | G | T |
| | REN24 | C | T | A | G | G | : | G | C | G | C | A | C | G | C | C | C | G | A | C | G | T | T | A | G | C |
| | REN24 | C | T | A | G | G | : | G | C | C | C | A | T | G | C | C | C | G | G | T | G | T | T | A | G | T |
| | REN3 | C | A | G | G | A | G | G | C | G | C | A | T | G | C | C | C | G | A | T | G | T | C | A | G | C |
| | REN3 | C | T | G | G | G | : | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | C | A | G | C |
| | REN6 | C | A | G | G | A | G | G | C | G | T | A | T | G | C | C | T | G | A | C | G | T | T | A | A | C |
| | REN6 | T | T | G | G | G | : | G | T | G | C | G | C | G | C | C | C | G | A | T | G | T | T | G | G | T |
| | RKR19 | ? | ? | ? | ? | ? | : | G | C | G | C | A | C | G | C | C | T | G | A | C | G | T | T | A | G | T |
| | RKR19 | ? | ? | ? | ? | ? | : | G | T | C | C | A | T | G | C | C | C | G | A | C | G | T | T | G | G | T |
| | RKW10 | C | T | A | G | G | : | G | C | G | C | A | T | G | C | C | C | G | A | T | G | T | C | A | G | C |
| | RKW10 | C | T | G | G | G | : | G | C | G | C | A | C | G | C | C | C | G | G | T | G | T | T | A | G | C |
| | RKW20 | C | A | G | G | A | G | G | C | G | C | A | T | G | T | C | T | G | A | C | G | T | T | A | G | C |
| | RKW20 | C | A | G | G | G | : | G | C | G | C | G | T | G | T | C | C | G | A | T | G | T | T | A | G | C |
| | RKW5 | ? | T | A | G | G | : | G | C | G | C | A | C | G | C | C | C | G | G | T | G | T | T | A | G | C |
| | RKW5 | ? | T | G | G | G | : | G | C | C | C | A | C | G | C | C | C | G | G | T | G | T | T | A | G | T |
| | RKW6 | C | T | A | G | G | : | G | C | G | C | A | C | G | C | C | C | G | A | T | G | T | C | A | G | C |
| | RKW6 | C | T | G | G | G | : | G | C | G | C | A | T | G | C | C | C | G | G | T | G | T | T | A | G | C |
| | RKW8 | C | A | A | G | A | G | G | C | G | C | A | C | G | C | C | C | G | A | T | G | T | T | A | A | C |
| | RKW8 | C | T | G | G | G | : | G | C | G | C | G | T | G | T | C | C | G | G | T | G | T | T | A | G | C |
| | RKW9 | ? | T | A | G | G | : | G | C | G | C | A | T | G | C | C | C | G | A | C | G | T | T | A | G | C |
| | RKW9 | ? | T | A | G | G | : | G | C | G | C | A | C | G | C | C | C | G | G | T | G | T | T | A | G | T |
| Allele Frequency | A | 0.00 | 0.50 | 0.24 | 0.01 | 0.49 | | 0.04 | 0.00 | 0.00 | 0.00 | 0.45 | 0.00 | 0.11 | 0.00 | 0.00 | 0.00 | 0.03 | 0.73 | 0.00 | 0.01 | 0.00 | 0.00 | 0.94 | 0.06 | 0.00 |
| | C | 0.90 | 0.00 | 0.00 | 0.00 | 0.00 | | 0.00 | 0.96 | 0.06 | 0.99 | 0.00 | 0.30 | 0.00 | 0.51 | 0.99 | 0.95 | 0.00 | 0.00 | 0.29 | 0.00 | 0.05 | 0.41 | 0.00 | 0.00 | 0.84 |
| | G | 0.00 | 0.00 | 0.76 | 0.99 | 0.51 | | 0.96 | 0.00 | 0.94 | 0.00 | 0.55 | 0.00 | 0.89 | 0.00 | 0.00 | 0.00 | 0.98 | 0.28 | 0.00 | 0.99 | 0.00 | 0.00 | 0.06 | 0.94 | 0.00 |
| | T | 0.10 | 0.50 | 0.00 | 0.00 | 0.00 | | 0.00 | 0.04 | 0.00 | 0.01 | 0.00 | 0.70 | 0.00 | 0.49 | 0.01 | 0.05 | 0.00 | 0.00 | 0.71 | 0.00 | 0.95 | 0.59 | 0.00 | 0.00 | 0.16 |
| | D | | | | | | 0.55 | | | | | | | | | | | | | | | | | | | |
| | I | | | | | | 0.45 | | | | | | | | | | | | | | | | | | | |

| Group | Locus | BTN1_1 | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 93590 | 93619 | 93625 | 93693 | 93694 | 93698 | 93775 | 93787 | 93814 | 93850 | 93863 | 93867 | 93884 | 93926 | 93930 | 93953 | 94022 | 99976 |
| | Gene | I | I | I | I | I | I | I | I | E/S | I | I | I | I | I | I | I | I | N |
| | Alleles | W | R | S | Y | R | S | W | W | Y | W | M | H | W | R | K | R | W | Y |
| Eastern | ECL1 | ? | A | G | C | G | C | T | T | C | A | C | C | A | G | G | A | T | C |
| | ECL1 | ? | G | G | C | G | C | T | T | C | T | A | T | T | G | G | A | T | T |
| | ECL10 | T | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | A | C |
| | ECL10 | T | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | A | C |
| | ECL3 | ? | A | G | C | G | C | T | A | C | A | A | T | A | G | G | A | A | C |
| | ECL3 | ? | G | G | T | G | C | T | T | C | A | A | T | A | G | G | A | T | T |
| | ECL5 | T | A | G | C | G | C | T | A | C | A | C | C | A | G | T | A | A | C |
| | ECL5 | T | A | G | T | G | C | T | T | C | T | A | T | T | G | G | A | T | C |
| | ECL6 | T | A | G | C | G | C | T | A | C | A | C | C | A | G | G | A | A | C |
| | ECL6 | T | A | G | T | G | C | T | T | C | T | A | T | T | G | G | A | A | C |
| | ECL8 | ? | A | G | C | G | C | T | T | C | A | C | C | A | G | T | A | T | C |
| | ECL8 | ? | G | G | C | G | C | T | T | C | T | A | T | T | G | G | A | T | C |
| | EMC1 | A | G | G | C | G | C | T | T | C | A | A | T | A | G | G | A | T | ? |
| | EMC1 | A | G | G | C | G | C | T | T | C | A | A | T | A | G | G | A | T | ? |
| | EMC2 | A | G | G | C | G | C | T | T | C | A | A | T | A | G | G | A | T | ? |
| | EMC2 | A | G | G | C | G | C | T | T | C | A | A | T | A | G | G | G | T | ? |
| | EOM3 | ? | A | G | C | G | C | T | A | C | A | A | T | A | G | G | A | A | C |
| | EOM3 | ? | G | G | T | G | C | T | T | C | A | A | T | A | G | G | A | T | C |
| | EOM4 | ? | A | G | C | G | C | T | T | C | A | C | A | A | G | G | A | A | C |
| | EOM4 | ? | G | G | C | A | G | T | T | C | T | A | T | T | G | G | A | T | C |
| | EOM6 | A | G | G | C | G | C | T | T | C | A | C | T | A | A | G | A | A | C |
| | EOM6 | T | G | G | C | G | C | T | T | C | T | A | T | T | G | G | A | T | C |
| | EWW11 | ? | A | G | C | G | C | T | A | C | A | C | C | A | G | G | A | A | C |
| | EWW11 | ? | A | G | T | G | C | T | T | C | T | A | T | T | G | G | A | A | C |
| | EWW13 | T | A | G | C | G | C | T | T | C | A | A | T | A | G | G | A | A | C |
| | EWW13 | A | G | G | T | G | C | T | T | C | A | A | T | A | G | G | G | T | C |
| | EWW14 | ? | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | A | C |
| | EWW14 | ? | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | A | C |
| | EWW4 | T | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | A | C |
| | EWW4 | T | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | A | C |
| | MINN | ? | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | T | C |
| | MINN | ? | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | T | C |
| Merriam's | MCS1 | ? | A | G | C | A | C | T | T | T | T | C | C | T | G | G | A | T | C |
| | MCS1 | ? | A | G | C | G | G | A | T | C | T | C | A | T | G | G | G | T | C |
| | MCS3 | ? | A | G | C | G | C | T | T | C | A | C | C | A | G | G | A | T | C |
| | MCS3 | ? | A | G | T | G | C | T | T | C | T | A | T | T | G | G | A | A | C |
| | MCS4 | ? | A | G | C | G | C | T | A | C | A | C | C | A | G | G | A | A | C |
| | MCS4 | ? | A | G | T | G | C | T | T | C | T | A | T | T | G | G | A | A | C |
| | MCS6 | ? | ? | ? | T | G | C | T | T | C | A | A | T | A | G | G | A | A | C |
| | MCS6 | ? | ? | ? | T | G | C | T | T | C | A | A | T | A | G | G | A | A | C |
| | MCS7 | ? | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | T | C |
| | MCS7 | ? | A | G | C | G | C | T | T | C | T | C | C | T | G | G | A | T | C |
| | MGD10 | T | A | G | C | A | C | A | T | C | T | C | C | T | G | G | A | T | C |
| | MGD10 | T | A | G | C | G | G | T | T | T | T | C | A | T | G | G | G | T | C |
| | MGD8 | T | A | G | C | A | C | T | T | T | T | C | A | T | G | G | A | T | C |
| | MGD8 | T | A | G | C | G | G | A | T | C | T | C | C | T | G | G | G | T | C |
| | MGD9 | A | A | G | C | A | C | T | T | T | A | C | A | A | G | G | A | T | C |
| | MGD9 | T | G | G | C | G | G | A | T | C | T | A | T | T | G | G | G | T | C |
| | MSL11 | ? | A | G | C | G | C | T | A | C | A | C | C | A | G | G | A | A | C |
| | MSL11 | ? | A | G | T | G | C | T | T | C | T | A | T | T | G | G | A | A | C |
| | MSL12 | ? | A | G | C | A | C | T | A | T | A | C | A | A | G | G | A | T | C |
| | MSL12 | ? | A | G | T | G | G | A | T | C | T | A | T | T | G | G | G | A | C |
| | MSL4 | ? | A | G | C | A | C | T | A | C | A | C | A | A | G | G | A | A | C |
| | MSL4 | ? | A | G | T | G | G | T | T | C | T | A | T | T | G | G | A | A | T |
| | MSL6 | T | A | G | C | G | C | T | A | C | A | C | C | A | G | G | A | A | C |
| | MSL6 | T | A | G | T | G | C | T | T | C | T | A | T | T | G | G | A | A | C |
| | MSL9 | ? | ? | ? | T | G | C | T | A | C | A | A | T | A | G | G | A | A | C |
| | MSL9 | ? | ? | ? | T | G | C | T | A | C | A | A | T | A | G | G | A | A | C |
| Rio Grande | REN19 | ? | A | G | C | A | C | T | T | C | A | C | A | A | G | G | A | T | C |
| | REN19 | ? | G | G | C | G | G | T | T | T | T | A | T | T | G | G | G | T | C |
| | REN24 | ? | A | G | C | A | C | A | A | T | T | C | A | T | G | G | A | T | C |
| | REN24 | ? | A | G | C | G | G | T | T | C | T | C | A | T | G | G | G | T | T |
| | REN3 | ? | A | G | T | G | C | T | A | C | A | A | T | A | G | G | A | A | C |
| | REN3 | ? | A | G | T | G | C | T | T | C | A | A | T | A | G | G | A | A | C |
| | REN6 | ? | A | G | C | G | C | T | T | C | A | A | T | A | G | G | A | A | C |
| | REN6 | ? | G | G | T | G | C | T | A | C | A | A | T | A | G | G | G | T | C |
| | RKR19 | A | A | G | C | A | C | T | T | T | A | C | A | A | G | G | A | T | C |
| | RKR19 | T | G | G | C | G | G | T | T | C | T | A | T | T | G | G | G | T | C |
| | RKW10 | ? | A | G | C | G | C | T | T | C | A | A | T | A | G | G | A | A | C |
| | RKW10 | ? | G | G | T | G | C | T | A | C | A | A | T | A | G | G | G | T | C |
| | RKW20 | ? | A | G | C | G | C | T | A | C | T | C | C | T | G | G | A | T | C |
| | RKW20 | ? | A | G | C | G | C | T | T | C | T | C | A | T | G | G | A | T | T |
| | RKW5 | ? | A | G | C | A | C | T | A | T | T | C | A | T | G | G | A | T | T |
| | RKW5 | ? | A | G | C | G | G | T | T | C | T | C | A | T | G | G | G | T | T |
| | RKW6 | ? | A | G | C | A | C | T | T | C | A | C | A | A | G | G | A | A | C |
| | RKW6 | ? | A | G | T | G | G | T | T | C | T | A | T | T | G | G | A | A | T |
| | RKW8 | ? | A | G | C | G | C | T | A | C | A | A | T | A | G | G | A | A | C |
| | RKW8 | ? | G | G | T | G | C | T | T | C | A | A | T | A | G | G | G | T | T |
| | RKW9 | ? | G | C | C | G | C | T | T | C | A | C | A | A | G | G | A | T | T |
| | RKW9 | ? | G | G | C | G | C | T | T | C | T | A | T | T | G | G | G | T | T |
| Allele Frequency | A | 0.31 | 0.74 | 0.00 | 0.00 | 0.15 | 0.00 | 0.08 | 0.24 | 0.00 | 0.50 | 0.50 | 0.20 | 0.50 | 0.01 | 0.00 | 0.81 | 0.46 | 0.00 |
| | C | 0.00 | 0.00 | 0.01 | 0.73 | 0.00 | 0.85 | 0.00 | 0.00 | 0.89 | 0.00 | 0.50 | 0.29 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.86 |
| | G | 0.00 | 0.26 | 0.99 | 0.00 | 0.85 | 0.15 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.99 | 0.98 | 0.19 | 0.00 | 0.00 |
| | T | 0.69 | 0.00 | 0.00 | 0.28 | 0.00 | 0.00 | 0.93 | 0.76 | 0.11 | 0.50 | 0.00 | 0.51 | 0.50 | 0.00 | 0.03 | 0.00 | 0.54 | 0.14 |
| | D | | | | | | | | | | | | | | | | | | |
| | I | | | | | | | | | | | | | | | | | | |

| | Locus | BTN1_2 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 100006 | 100042 | 100082 | 100085 | 100086 | 100089 | 100110 | 100188 | 100209 | 100272 | 100298 | 100343 | 100350 | 100358 | 100363 | 100380 | 100389 | 100390 | 100393 | 100481 | 100543 | 100593 | 100597 | 100604 | 100613 | 100653 | 100671 | 100723 |
| | Gene | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N |
| | Alleles | R | Y | M | S | Y | C | Y | K | R | R | A | Y | T | R | R | Y | Y | R | R | R | Y | Y | Y | Y | R | Y | S | R |
| Eastern | ECL1 | A | T | C | C | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | C | A |
| | ECL1 | A | C | C | G | T | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | T | G | T | G | A |
| | ECL10 | A | T | C | G | C | C | T | G | G | A | A | C | T | A | A | C | C | A | G | A | C | C | T | T | G | C | C | G |
| | ECL10 | A | T | C | G | C | C | T | G | G | A | A | C | T | A | A | C | C | A | G | A | C | C | T | T | G | C | C | G |
| | ECL3 | A | C | C | C | C | C | T | G | G | G | A | T | T | G | A | T | C | G | G | G | C | C | C | T | G | C | C | A |
| | ECL3 | A | T | C | C | T | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | T | T | G | T | C | G |
| | ECL5 | A | T | C | C | C | C | T | G | G | G | A | C | T | G | A | T | C | G | G | A | C | C | C | T | G | C | C | A |
| | ECL5 | A | T | C | G | C | C | T | T | G | G | A | T | T | G | A | C | C | G | G | A | C | C | T | T | A | T | G | G |
| | ECL6 | A | T | C | C | C | C | T | G | G | A | A | C | T | A | A | C | C | A | G | A | C | C | T | T | G | C | C | G |
| | ECL6 | A | T | C | G | C | C | T | G | G | G | A | T | T | A | A | C | C | A | G | A | C | C | T | T | G | C | C | G |
| | ECL8 | A | T | C | G | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | C | A |
| | ECL8 | A | C | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | T | A | T | G | A |
| | EMC1 | A | C | C | G | C | C | T | G | G | G | A | C | T | G | G | C | C | G | G | G | C | C | C | T | G | T | C | G |
| | EMC1 | A | C | C | G | C | C | T | G | G | G | A | C | : | G | A | C | C | G | G | G | C | C | C | T | G | T | C | G |
| | EMC2 | A | C | C | G | C | C | T | G | G | G | A | C | T | G | G | C | C | G | G | G | C | C | C | C | G | C | C | G |
| | EMC2 | A | C | C | G | C | C | T | G | G | G | A | C | : | G | A | C | C | G | G | G | C | C | C | T | G | T | C | G |
| | EOM3 | A | C | C | C | C | C | T | G | G | G | A | T | T | G | A | T | C | G | G | G | C | C | C | T | G | T | C | A |
| | EOM3 | A | T | C | C | C | C | T | G | G | G | A | C | : | G | G | C | C | G | G | G | C | C | T | T | G | T | C | G |
| | EOM4 | A | C | C | G | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | C | G | C | C | G |
| | EOM4 | A | C | C | G | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | C | G | C | C | G |
| | EOM6 | A | C | C | G | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | C | G | C | C | A |
| | EOM6 | A | C | C | C | C | C | T | G | G | G | : | C | T | G | A | C | C | G | G | G | C | C | C | T | G | T | C | G |
| | EWW11 | A | T | C | G | C | C | T | G | G | A | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | C | C | G |
| | EWW11 | A | T | C | C | C | C | T | G | G | G | A | C | T | A | A | T | C | G | G | G | C | C | T | T | G | C | C | G |
| | EWW13 | A | T | C | G | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | C | G | C | C | A |
| | EWW13 | A | C | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | T | G | T | G | G |
| | EWW14 | A | T | C | G | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | C | C | A |
| | EWW14 | A | T | C | C | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | A | T | G | G |
| | EWW4 | A | T | C | G | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | C | G |
| | EWW4 | A | T | C | C | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | T | T | T | G | C | C | G |
| | MINN | A | T | C | C | C | C | T | G | G | G | A | T | T | G | A | T | C | G | G | G | C | C | T | T | G | C | C | G |
| | MINN | A | T | C | C | C | C | T | G | G | G | A | T | T | G | A | T | C | G | G | G | C | C | T | T | G | C | C | G |
| Merriam's | MCS1 | A | T | C | C | C | : | T | G | A | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | C | C | A |
| | MCS1 | G | C | A | C | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | T | C | G |
| | MCS3 | A | T | C | G | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | C | C | A |
| | MCS3 | A | T | C | C | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | T | C | T | G | T | C | G |
| | MCS4 | A | T | C | C | C | C | T | G | G | G | A | C | T | G | A | T | C | G | A | G | T | C | T | T | G | C | C | A |
| | MCS4 | A | T | C | C | C | C | T | G | G | G | A | T | T | G | A | T | C | G | A | G | C | C | C | T | G | T | C | G |
| | MCS6 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | T | C | T | G | T | C | A |
| | MCS6 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | T | C | T | G | T | C | A |
| | MCS7 | A | T | C | C | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | C | C | G |
| | MCS7 | A | T | C | C | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | C | C | G |
| | MGD10 | A | T | C | C | C | : | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | T | C | A |
| | MGD10 | G | C | A | C | C | C | T | G | A | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | C | C | G |
| | MGD8 | A | T | C | C | C | : | T | G | A | G | A | C | T | G | A | C | C | G | G | A | C | C | C | C | G | C | C | G |
| | MGD8 | G | C | A | C | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | T | T | G | T | C | G |
| | MGD9 | G | C | A | C | C | : | T | G | A | G | A | C | T | G | A | C | T | G | G | A | C | C | C | C | G | T | C | A |
| | MGD9 | G | C | A | C | C | : | T | G | A | G | A | C | T | G | A | C | C | G | G | A | C | C | C | C | G | T | C | A |
| | MSL11 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | MSL11 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | MSL12 | A | T | C | G | C | : | T | G | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | C | A |
| | MSL12 | G | C | A | C | C | C | T | T | A | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | MSL4 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | C | G | C | C | A |
| | MSL4 | A | C | C | C | T | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | T | G | T | G | G |
| | MSL6 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | MSL6 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | MSL9 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | MSL9 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| Rio Grande | REN19 | A | C | C | C | C | C | T | G | G | G | : | C | T | G | A | C | C | G | G | G | C | C | C | T | G | T | C | A |
| | REN19 | A | C | C | C | C | C | T | G | G | G | : | C | T | G | A | C | C | G | G | G | C | C | C | T | G | T | C | A |
| | REN24 | A | C | C | C | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | T | G | C | C | A |
| | REN24 | A | C | C | C | C | C | T | T | G | G | : | C | T | G | A | C | C | G | G | G | C | C | C | T | G | T | C | G |
| | REN3 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | C | A |
| | REN3 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | T | C | T | G | T | G | G |
| | REN6 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | REN6 | A | T | C | G | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | RKR19 | G | C | A | C | C | : | T | G | A | G | A | C | T | G | A | C | C | G | G | A | C | C | C | C | G | C | C | G |
| | RKR19 | G | C | A | C | C | : | T | G | A | G | A | C | T | G | A | C | C | G | G | A | C | C | C | C | G | C | C | G |
| | RKW10 | A | T | C | C | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | RKW10 | A | T | C | C | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | A |
| | RKW20 | A | C | C | C | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | C | G | C | C | A |
| | RKW20 | A | T | C | C | T | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | T | G | T | G | G |
| | RKW5 | A | C | C | C | T | C | C | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | C | G | C | C | G |
| | RKW5 | A | C | C | C | T | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | C | G | C | C | G |
| | RKW6 | A | C | C | C | C | C | T | G | G | G | A | C | T | G | A | C | C | G | G | G | C | C | T | C | G | C | C | G |
| | RKW6 | A | T | C | C | T | C | T | T | G | G | A | C | T | G | A | T | C | G | G | G | C | C | C | T | G | C | C | G |
| | RKW8 | A | C | C | C | C | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | C | G | C | C | A |
| | RKW8 | A | T | C | C | T | C | T | T | G | G | A | C | T | G | A | C | C | G | G | A | C | C | C | T | G | T | G | G |
| | RKW9 | A | C | C | C | T | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | C | G | C | C | G |
| | RKW9 | A | C | C | C | T | C | T | T | G | G | A | C | T | G | A | C | C | G | G | G | C | C | C | C | G | C | C | G |
| Allele Frequency | A | 0.90 | 0.00 | 0.10 | 0.00 | 0.00 | | 0.00 | 0.00 | 0.10 | 0.05 | | 0.00 | | 0.06 | 0.96 | 0.00 | 0.00 | 0.05 | 0.03 | 0.59 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.46 |
| | C | 0.00 | 0.41 | 0.90 | 0.58 | 0.88 | | 0.01 | 0.00 | 0.00 | 0.00 | | 0.91 | | 0.00 | 0.00 | 0.89 | 0.99 | 0.00 | 0.00 | 0.00 | 0.99 | 0.94 | 0.71 | 0.23 | 0.00 | 0.45 | 0.75 | 0.00 |
| | G | 0.10 | 0.00 | 0.00 | 0.43 | 0.00 | | 0.00 | 0.55 | 0.90 | 0.95 | | 0.00 | | 0.94 | 0.04 | 0.00 | 0.00 | 0.95 | 0.98 | 0.41 | 0.00 | 0.00 | 0.00 | 0.00 | 0.96 | 0.00 | 0.25 | 0.54 |
| | T | 0.00 | 0.59 | 0.00 | 0.00 | 0.13 | | 0.99 | 0.45 | 0.00 | 0.00 | | 0.09 | | 0.00 | 0.00 | 0.11 | 0.01 | 0.00 | 0.00 | 0.00 | 0.01 | 0.06 | 0.29 | 0.78 | 0.00 | 0.55 | 0.00 | 0.00 |
| | D | | | | | | 0.10 | | | | | 0.05 | | 0.04 | | | | | | | | | | | | | | | |
| | I | | | | | | 0.90 | | | | | 0.95 | | 0.96 | | | | | | | | | | | | | | | |

| | Locus | Blec | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 130446 | 130463 | 130515 | 130676 | 130709 | 130745 | 130750 | 130779 | 130787 | 130788 | 130978 | 130998 | 131029 | 131052 | 131155 | 131169 | 131178 | 131179 | 131280 | 131291 | 131293 |
| | Gene | I | I | I | I | I | I | I | I | I | I | E/N | I | I | I | I | I | I | I | I | I | I |
| | Alleles | Y | Y | Y | Y | Y | R | M | Y | K | M | Y | K | Y | M | Y | Y | Y | R | S | Y | M |
| Eastern | ECL1 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | C | A |
| | ECL1 | T | C | C | T | C | G | C | T | T | C | C | G | C | C | T | T | T | G | C | T | C |
| | ECL10 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | T | C | C | G | C | T | A |
| | ECL10 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | T | C | C | G | G | T | A |
| | ECL3 | T | C | C | C | C | A | C | C | G | A | C | G | C | C | T | C | C | G | C | C | A |
| | ECL3 | T | C | C | T | C | G | C | T | T | C | T | G | T | C | T | T | C | G | G | T | C |
| | ECL5 | T | C | C | C | C | A | C | C | G | A | C | G | C | C | T | C | C | G | G | T | A |
| | ECL5 | T | C | C | C | C | G | C | C | T | C | T | G | T | C | T | C | C | G | G | T | A |
| | ECL6 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL6 | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | ECL8 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | T | C | C | G | C | T | A |
| | ECL8 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | T | G | G | T | A |
| | EMC1 | ? | C | C | T | C | G | C | C | G | C | C | G | C | C | T | T | T | G | C | T | A |
| | EMC1 | ? | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | T | A |
| | EMC2 | ? | ? | ? | T | C | G | C | C | G | C | C | G | C | A | T | T | T | G | C | T | A |
| | EMC2 | ? | ? | ? | T | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | T | A |
| | EOM3 | T | C | C | C | C | G | C | C | G | C | C | G | C | C | T | T | T | G | C | T | A |
| | EOM3 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | EOM4 | T | C | C | T | C | A | C | C | T | C | C | G | C | C | T | T | C | G | C | T | A |
| | EOM4 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | T | A |
| | EOM6 | C | C | C | C | C | G | C | C | T | C | C | G | C | C | ? | ? | ? | ? | ? | ? | ? |
| | EOM6 | T | C | C | T | C | G | C | T | T | C | C | G | C | C | ? | ? | ? | ? | ? | ? | ? |
| | EWW11 | T | T | T | C | C | A | C | C | T | C | C | G | C | C | T | C | C | A | G | T | A |
| | EWW11 | T | C | C | T | C | G | C | C | T | C | C | T | C | C | T | T | C | G | G | T | A |
| | EWW13 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | ? | ? | ? | ? | ? | ? | ? |
| | EWW13 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | ? | ? | ? | ? | ? | ? | ? |
| | EWW14 | T | T | C | C | T | A | C | C | T | C | C | G | C | C | T | C | C | G | C | T | A |
| | EWW14 | T | C | C | T | C | A | C | C | T | C | C | G | C | C | T | C | C | G | G | T | A |
| | EWW4 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | T | C | C | G | C | T | A |
| | EWW4 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | C | C | G | G | T | A |
| Merriam's | MINN | T | C | C | T | C | A | C | C | T | A | C | G | C | C | C | C | C | G | C | T | A |
| | MINN | T | C | C | T | C | G | A | C | G | C | T | T | T | C | T | C | C | G | C | T | A |
| | MCS1 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | C | A | C | T | A |
| | MCS1 | T | C | C | T | C | G | C | C | T | C | C | T | C | C | C | C | C | G | C | T | A |
| | MCS3 | T | C | C | T | C | A | A | C | G | A | C | G | C | C | T | C | C | G | C | T | A |
| | MCS3 | T | C | C | T | C | G | C | C | T | C | T | T | T | C | C | C | C | G | G | T | A |
| | MCS4 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | T | C | C | G | G | T | A |
| | MCS4 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | C | C | C | G | G | T | A |
| | MCS6 | T | C | C | C | C | G | A | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | MCS6 | T | C | C | C | C | G | A | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | MCS7 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | C | C | C | G | C | T | A |
| | MCS7 | T | C | C | T | C | G | C | C | T | C | C | T | C | C | T | C | C | G | C | T | A |
| | MGD10 | T | C | C | C | C | A | C | C | T | A | C | G | C | C | C | C | C | G | C | T | A |
| | MGD10 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | T | A |
| | MGD8 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | C | C | A | C | T | A |
| | MGD8 | T | C | C | T | C | G | C | C | T | C | C | T | C | C | C | T | C | G | C | T | A |
| | MGD9 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | T | A |
| | MGD9 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | MSL11 | T | C | C | T | C | A | A | C | G | A | T | G | T | C | T | C | C | G | C | C | A |
| | MSL11 | T | C | C | T | C | A | A | C | G | A | T | G | T | C | T | C | C | G | C | T | A |
| | MSL12 | C | C | C | T | C | A | A | C | G | A | C | G | C | C | T | C | C | G | C | C | A |
| | MSL12 | T | C | C | T | C | G | C | C | T | C | T | G | T | C | T | T | T | G | C | T | A |
| | MSL4 | C | C | C | T | C | A | A | C | G | A | C | G | C | C | T | C | C | G | C | C | A |
| | MSL4 | T | C | C | T | C | G | C | C | T | C | T | G | T | C | T | T | T | G | C | T | A |
| | MSL6 | T | C | C | T | C | A | A | C | T | A | T | G | T | C | T | C | C | G | G | C | A |
| | MSL6 | T | C | C | T | C | A | C | C | G | C | T | G | T | C | T | C | C | G | C | T | A |
| | MSL9 | T | C | C | T | C | A | A | C | G | A | T | G | T | C | T | C | C | G | C | C | A |
| | MSL9 | T | C | C | T | C | A | A | C | G | A | T | G | T | C | T | C | C | G | C | C | A |
| Rio Grande | REN19 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | T | A |
| | REN19 | T | C | C | T | C | G | C | T | T | C | C | G | C | C | T | T | T | G | C | C | C |
| | REN24 | T | C | T | T | C | A | C | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | REN24 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | REN3 | T | C | C | C | C | A | A | C | G | A | C | G | C | C | T | C | C | G | C | T | A |
| | REN3 | T | C | C | T | C | G | C | C | T | C | T | G | T | C | T | T | T | G | G | T | A |
| | REN6 | T | C | T | C | C | G | C | C | T | C | C | G | C | C | T | C | C | G | C | T | A |
| | REN6 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | RKR19 | T | C | C | C | C | G | A | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | RKR19 | T | C | C | T | C | G | A | C | T | C | C | G | C | C | T | T | T | G | C | T | A |
| | RKW10 | T | C | C | T | C | A | A | C | T | A | C | G | C | C | T | C | C | G | C | T | A |
| | RKW10 | T | C | C | T | C | G | C | C | G | C | T | G | T | C | T | T | T | G | G | T | A |
| | RKW20 | T | C | C | T | C | A | A | C | G | A | C | G | C | C | T | C | C | A | C | T | A |
| | RKW20 | T | C | C | T | C | A | C | C | G | C | T | G | T | C | T | T | C | G | G | T | A |
| | RKW5 | C | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | C | A |
| | RKW5 | T | C | C | T | C | G | C | T | T | C | C | G | C | C | T | T | T | G | C | T | C |
| | RKW6 | T | C | C | C | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | C | A |
| | RKW6 | T | C | C | T | C | G | C | T | T | C | C | G | C | C | T | T | T | G | C | T | C |
| | RKW8 | T | C | C | T | C | A | A | C | G | A | C | G | C | C | T | C | C | G | C | C | C |
| | RKW8 | T | C | C | T | C | G | C | T | T | C | T | G | T | C | T | T | C | G | G | T | A |
| | RKW9 | T | C | C | T | C | G | C | C | T | C | C | G | C | C | T | T | C | G | C | C | C |
| | RKW9 | T | C | C | T | C | G | C | T | T | C | C | G | C | C | T | T | C | G | C | C | C |
| Allele Frequency | A | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.29 | 0.22 | 0.00 | 0.00 | 0.21 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.05 | 0.00 | 0.00 | 0.89 |
| | C | 0.05 | 0.97 | 0.96 | 0.29 | 0.99 | 0.00 | 0.78 | 0.90 | 0.00 | 0.79 | 0.79 | 0.00 | 0.79 | 0.99 | 0.09 | 0.47 | 0.72 | 0.00 | 0.77 | 0.19 | 0.11 |
| | G | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.71 | 0.00 | 0.00 | 0.24 | 0.00 | 0.00 | 0.92 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.95 | 0.23 | 0.00 | 0.00 |
| | T | 0.95 | 0.03 | 0.04 | 0.71 | 0.01 | 0.00 | 0.00 | 0.10 | 0.76 | 0.00 | 0.21 | 0.08 | 0.21 | 0.00 | 0.91 | 0.53 | 0.28 | 0.00 | 0.00 | 0.81 | 0.00 |
| | D | | | | | | | | | | | | | | | | | | | | | |
| | I | | | | | | | | | | | | | | | | | | | | | |

| | Locus | Tap 1 | | | | | | | | | | | | CD1.1 | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pos | 171878 | 171879 | 171942 | 171955 | 171957 | 172009 | 172010 | 172022 | 172110 | 172116 | 172117 | 172165 | 223 | 265 | 298 | 299 | 302 | 324 | 330 | 350 | 351 | 398 | 416 | 438 | 457 | 590 | 597 |
| | Gene | I | I | E/N | E/S | E/N | E/S | E/N | E/N | I | I | I | E/N | I | I | E/N | E/N | E/S | E/N | E/N | E/S | E/N | E/S | E/S | E/N | E/N | I | I |
| | Alleles | Y | R | Y | R | Y | Y | R | R | Y | R | Y | S | GATTGA | M | Y | R | Y | R | Y | W | Y | R | Y | R | R | R | Y |
| Eastern | ECL1 | C | G | C | A | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | ECL1 | C | A | C | G | C | C | G | G | C | G | C | G | GATTGA | A | C | G | C | G | C | T | C | G | C | G | A | G | C |
| | ECL10 | T | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | A | C |
| | ECL10 | T | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | A | C |
| | ECL3 | C | A | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | A | C |
| | ECL3 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | ECL5 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | A | C |
| | ECL5 | C | G | C | G | C | C | G | G | C | G | C | C | GATTGA | C | C | G | C | G | C | T | C | G | C | G | G | A | C |
| | ECL6 | C | G | C | G | C | C | G | G | T | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | A | C |
| | ECL6 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | ECL8 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | A | C |
| | ECL8 | C | G | C | G | C | C | G | G | C | G | C | C | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | EMC1 | C | G | C | G | C | C | G | G | C | G | C | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EMC1 | C | G | C | G | C | C | G | G | C | G | C | C | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EMC2 | C | G | C | G | C | C | G | G | C | G | C | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EMC2 | C | G | C | G | C | C | G | G | C | G | C | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | EOM3 | C | G | C | G | C | C | G | G | T | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | EOM3 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | EOM4 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | A | C | G | C | A | C | A | C | G | A | G | C |
| | EOM4 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | EOM6 | C | G | C | G | C | C | A | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | EOM6 | C | A | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | EWW11 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | T | T | C | G | C | G | G | A | C |
| | EWW11 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | T | T | C | G | C | G | G | A | C |
| | EWW13 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | EWW13 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | A | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | EWW14 | C | G | C | G | C | C | G | G | C | G | C | C | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | A | C |
| | EWW14 | C | G | C | G | C | C | G | G | C | G | C | C | GATTGA | C | C | G | C | G | C | T | C | G | C | G | G | G | T |
| | EWW4 | C | G | C | A | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | A | C |
| | EWW4 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | A | C |
| Merriam's | MINN | C | G | C | G | C | C | G | G | T | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | G | T |
| | MINN | C | G | C | G | C | C | G | G | T | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | G | T |
| | MCS1 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | MCS1 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | A | C | G | T | A | C | T | T | G | C | G | G | G | C |
| | MCS3 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MCS3 | C | G | C | G | C | C | G | G | C | G | C | G | : | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MCS4 | C | G | C | A | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | T | C | G | C | G | G | A | C |
| | MCS4 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | MCS6 | C | G | C | G | C | C | G | G | C | G | C | G | : | A | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | MCS6 | C | G | C | G | C | C | G | G | C | G | C | G | : | A | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | MCS7 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | A | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | MCS7 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | A | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | MGD10 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | MGD10 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | A | C | G | T | A | C | T | T | G | C | G | G | G | C |
| | MGD8 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | MGD8 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | A | C | G | T | A | C | T | T | G | C | G | G | G | C |
| | MGD9 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | A | C | G | C | A | C | A | C | G | A | G | C |
| | MGD9 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | MSL11 | C | G | C | A | C | C | G | G | C | G | C | G | GATTGA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MSL11 | C | G | C | G | C | C | G | G | C | G | C | G | : | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MSL12 | C | G | C | A | C | C | G | G | C | G | C | G | GATTGA | C | C | A | C | G | C | A | C | G | T | G | A | G | C |
| | MSL12 | C | G | C | A | C | C | G | G | C | G | C | G | GATTGA | A | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | MSL4 | C | G | C | G | C | C | G | G | C | A | C | G | GATTGA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MSL4 | C | G | C | G | C | C | G | G | C | G | C | G | : | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MSL6 | C | G | C | G | C | C | G | A | C | G | C | G | GATTGA | A | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | MSL6 | C | G | C | A | C | C | G | G | C | G | C | G | GATTGA | A | C | G | C | G | C | T | C | G | C | G | G | G | C |
| | MSL9 | C | G | C | A | C | C | G | G | C | G | C | G | GATTGA | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | MSL9 | C | G | C | G | C | C | G | G | C | G | C | G | : | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| Rio Grande | REN19 | C | G | C | G | C | C | G | G | C | A | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | REN19 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | A | C | G | C | A | C | A | C | G | A | G | C |
| | REN24 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | A | C | G | C | A | C | A | C | G | A | G | C |
| | REN24 | C | G | C | G | C | C | G | G | C | G | T | C | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | REN3 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | ? | ? |
| | REN3 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | ? | ? |
| | REN6 | C | G | C | G | C | C | G | G | C | A | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | A | C |
| | REN6 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | G | C |
| | RKR19 | C | G | C | G | C | C | G | G | C | A | C | G | GATTGA | C | T | A | C | G | C | A | C | A | C | A | A | G | C |
| | RKR19 | C | G | C | G | C | C | G | G | C | G | C | C | GATTGA | C | C | G | C | G | C | A | C | G | C | G | A | G | C |
| | RKW10 | C | G | C | G | T | T | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | A | C | G | A | G | C |
| | RKW10 | C | G | T | G | C | C | G | G | C | G | C | G | GATTGA | C | C | A | C | G | C | A | C | G | C | G | G | G | C |
| | RKW20 | C | G | C | G | T | C | G | G | C | G | C | G | GATTGA | C | C | A | C | G | C | A | C | A | C | G | A | A | C |
| | RKW20 | C | G | T | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | G | C |
| | RKW5 | C | G | C | G | T | T | G | G | C | G | C | G | GATTGA | C | C | A | C | G | C | A | C | A | C | G | A | G | C |
| | RKW5 | C | G | T | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | A | A | G | C |
| | RKW6 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | C | G | C | A | C | G | C | A | A | A | C |
| | RKW6 | C | G | C | G | C | C | G | G | C | G | C | C | GATTGA | C | C | G | C | G | C | A | C | G | C | G | G | G | C |
| | RKW8 | C | G | C | G | C | C | G | G | C | G | C | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW8 | C | G | C | G | C | C | G | G | C | G | C | G | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| | RKW9 | C | G | C | G | C | C | G | G | C | G | T | G | GATTGA | C | C | G | T | A | C | A | T | G | C | G | A | G | C |
| | RKW9 | C | G | C | G | C | C | G | G | C | G | C | G | GATTGA | C | C | G | T | A | C | A | T | G | C | G | A | G | C |
| Allele Frequency | A | 0.00 | 0.04 | 0.00 | 0.10 | 0.00 | 0.00 | 0.01 | 0.01 | 0.00 | 0.05 | 0.00 | 0.00 | | 0.18 | 0.00 | 0.14 | 0.00 | 0.08 | 0.00 | 0.70 | 0.00 | 0.12 | 0.00 | 0.05 | 0.55 | 0.25 | 0.00 |
| | C | 0.98 | 0.00 | 0.96 | 0.00 | 0.96 | 0.98 | 0.00 | 0.00 | 0.95 | 0.00 | 0.98 | 0.10 | | 0.82 | 0.98 | 0.00 | 0.92 | 0.00 | 0.97 | 0.00 | 0.92 | 0.00 | 0.98 | 0.00 | 0.00 | 0.00 | 0.95 |
| | G | 0.00 | 0.96 | 0.00 | 0.90 | 0.00 | 0.00 | 0.99 | 0.99 | 0.00 | 0.95 | 0.00 | 0.90 | | 0.00 | 0.00 | 0.86 | 0.00 | 0.92 | 0.00 | 0.00 | 0.00 | 0.88 | 0.00 | 0.95 | 0.45 | 0.75 | 0.00 |
| | T | 0.03 | 0.00 | 0.04 | 0.00 | 0.04 | 0.03 | 0.00 | 0.00 | 0.05 | 0.00 | 0.03 | 0.00 | | 0.00 | 0.02 | 0.00 | 0.08 | 0.00 | 0.03 | 0.30 | 0.08 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.05 |
| | D | | | | | | | | | | | | | 0.08 | | | | | | | | | | | | | | |
| | I | | | | | | | | | | | | | 0.92 | | | | | | | | | | | | | | |

**Table 4:** Characterization of single nucleotide polymorphisms identified within wild turkey subpopulations.

| | Observed | Frequency | Newly Identified | Frequency |
|---|---|---|---|---|
| **Non Coding** | 38 | 0.165 | 12 | 0.112 |
| **Intron** | 154 | 0.667 | 76 | 0.710 |
| **Exon** | 45 | 0.195 | 22 | 0.206 |
| **Syn** | 27 | 0.117 | 11 | 0.103 |
| **Non-Syn** | 18 | 0.078 | 11 | 0.103 |
| **Transition** | 173 | 0.749 | 88 | 0.822 |
| **Transversion** | 54 | 0.234 | 19 | 0.178 |
| **Triallelic** | 2 | 0.009 | 2 | 0.019 |
| **Indel** | 8 | NA | 3 | NA |
| **MAF $\geq$ 0.2** | 84 | 0.367 | 5 | 0.047 |
| **Total SNP** | 229 | | 107 | |

**Table 5:** MHC haplotypes in wild turkey inferred from 37 SNPs and DIPs. Deletions are indicated by "-". Lower case letters and "~" indicate SNP or DIP alleles with haplotype assignment posterior probabilities ≤ 0.9.

| Individual ID | Haplotype | Other ID |
|---|---|---|
| ECL1b | ACGGGACTCACTCCGAAAGGTCCCTTCTTTGTACTGA | |
| ECL1a | GCGG-ACCCGCTTAAGTG-GCCTCTAAATCCTGCTCA | ECL3a |
| ECL10a | ACAC-ACTCATTCCGAAAGGTCTCTTCTATGGATCCG | |
| ECL10b | ACAC-ACTCATTTCGAAAGGTCTCTTCTATGGATCCG | |
| ECL3b | ACAC-ACTCACTCAGGTG-GCCCTAAAAATCGGTCCG | |
| ECL3a | GCGG-ACCCGCTTAAGTG-GCCTCTAAATCCTGCTCA | ECL1a |
| ECL5b | ACac-ACTCACTCAGGTG-ACCCTAAAAATCGATCCG | ECL6b |
| ECL5a | ACgg-ACTCACTCCGGAAGGTCCCTTCTTTGTACTGA | |
| ECL6b | ACac-ACTCACTCAGGTG-ACCCTAAAAATCGATCCG | ECL5b |
| ECL6a | ACggGATTCGCTCCGGAAGGTCTCTTCTATGGATCCG | |
| ECL8b | GCac-ACTCACTCCGGAAGGTCCCTTCTTTGTACTGA | |
| ECL8a | GCgg-ACCCGCTTAAGTG-ACCTCTAAATCGGGCTCA | |
| EMC1b | GCGG-GcCCGTCTAAGAGGACCTCTAAATCGGGCTCG | EMC2b |
| EMC1a | GCGG-GtCCGCCTAAGTG-ACCTCTAAATCGGGCTCG | |
| EMC2a | GCGG-GCCCGCCTAAGTG-ACCTCTAAATCGGGCCCG | EOM4b |
| EMC2b | GCGG-GCCCGTCTAAGAGGACCTCTAAATCGGGCTCG | EMC1b |
| EOM3b | aCGG-gtTCACTCAGAAGGGTCCTAAAAATCGGTTCG | |
| EOM3a | gCGG-acCCGCCTAAGTG-ACCTCTAAATCCGGCTCA | REN19a |
| EOM4b | GCGG-GCCCGCCTAAGTG-ACCTCTAAATCGGGCCCG | EMC2a |
| EOM4a | GCGG-GCCCGCCTAGGTG-GCCTCTTCTACGGGCCCG | |
| EOM6a | GCGG-aCCCGCCTAAGTG-ACCTCTAAATCGGGCCCG | |
| EOM6b | GCGG-gCTCGTTCAAGTG-ACCTCTTCTACCGGCTCA | |
| EWW11b | ATAG-ACTCaCTCaaGAAGaTTCTAAAAATCGGTCCG | |
| EWW11a | ATAG-ACTCgCTCcgGAAGgTCTCTTCTATGGATCCG | |
| EWW13b | aCag-accCaCTTAAGTG-ACCTCTAAATCGGGCCCG | |
| EWW13a | gCgc-gttCgCTCAAGAAGGTCTTTAAAATGTACTGA | |
| EWW14a | ACAc-ACTCACTCCGAAAGGTCCCTTCTATcGATCCG | |
| EWW14b | ACAg-ACTCACTCCGAAAGGTCCCTTCTATgTACTGA | |
| EWW4b | atAC~ATTCGCTCAAGAAGGTCTCTTCTATcTACTCG | |
| EWW4a | gcACgATTCGCTCCGGAAGGTCTCTTCTATgGATCCG | |
| MINN-M | ATAC-ATCTGCTCAGAAAGGTCTCTTCTTTCGGTCCG | |
| MCS1b | GCACGACCTGCTCAGGAAGGTCCCTTCTTTCGATCCG | MCS3a, MCS7-M |
| MCS1a | GCGG-ACCCGCCTAAGAGGGCCTCTTCTTCCGATTCA | |
| MCS3b | ACGGGACCTGCTCAGAAAGGTCCTTAAAATGTACTCA | |
| MCS3a | GCACGACCTGCTCAGGAAGGTCCCTTCTTTCGATCCG | MCS1b, MCS7-M |
| MCS4b | aCaC-aTTCGCTCcGaAAGGTCCTAAAAATCGGTCCG | |
| MCS4a | gCgC-gTTCGCTCaGgAAGGTCCCTTCTATCGGCTCA | |

| Individual ID | Haplotype | Other ID |
|---|---|---|
| MCS6-M | ACGGGACCCACTCCGAAAGGTCCTTAAAATGTACTCA | |
| MCS7-M | GCACGACCTGCTCAGGAAGGTCCCTTCTTTCGATCCG | MCS1b, MCS7-M |
| MGD10b | aCGG-GTTCGCTCAGAAAGATCCCTTCTTTCGATCCG | MGD8b |
| MGD10a | gCGC-GTCCGCCTAAGAGGGCCTCTTCTTCCGATTCA | |
| MGD8b | aCGG-GTTCGCTCAGAAAGATCCCTTCTTTCGATCCG | MGD10b |
| MGD8a | gCGC-GTCCGCCTAAGAGGGCCTCTTCTTCCGACTCG | |
| MGD9a | GCGG-GCcCGCCTAAGTG-ACCTCTAAATCCGACTCA | |
| MGD9b | GCGG-GCtCGCCTAAGTG-GCCTCTTCTTCCGACTCA | |
| MSL11a | ACGC-ACCCGCTCAGAAAGGTCCCTTCTATGTACTGA | MSL6b |
| MSL11b | ACGC-ACTCGTCCAGATAGGTCCTAAAAATGTACTGA | |
| MSL12b | ACGc-ACTCGTCCAGATAGaTCCTAAAAATGTACTGA | MSL9b |
| MSL12a | GCGg-ACTCGCTTAGGTG-gCCTCTTCTTCCGACTCA | |
| MSL4a | acGG-GTCCGTCTAGGTG-GCCTCTTCTACCTGCCCG | |
| MSL4b | gtGG-GTTCGTCCAGATAGGTCCTAAAAATGTACTGA | |
| MSL6b | aCgC~ACCCGCTCAGAAAGGTCCCTTCTATGTACTGA | MSL11a |
| MSL6a | gCaCgGTTCGTCCAGATAGGTCCTAAAAATGTACTGA | |
| MSL9b | aCgC~ACTCGTCCAGATAGATCCTAAAAATGTACTGA | MSL12b |
| MSL9a | gCaCgACCCGCTCAGATAGATCCTAAAAATGTACTGA | |
| REN19a | GCGG-ACCCGcCTAaGTG-aCCTCTAAATCCGGCTCA | EOM3a |
| REN19b | GCGG-ACCCGtCTAgGTG-gCCTCTTCTTCCGGCTCA | |
| REN24a | GCGG-GCcCGcCTAAGTG-ACCTCtTCTTCCGGCTCA | |
| REN24b | GCGG-GCtCGtCTAAGTG-ACCTCaTCTTCCTGCCCG | |
| REN3b | atagGACCTGCTCAGAAAGGTCCTTAAAATGTACTCg | |
| REN3a | gcgc-ACTCGCTTAAATG-ACCCTAAAAATGTACTGa | |
| REN6b | GCGG-ACCCGcctAaGtg~aCCTctAAAtTGTACTGA | |
| REN6a | GCGG-ACCCGttcAgGaaggCCTtaAAAaTGTACTGA | |
| RKR19a | GCGG-aCcCGTCtAAGTG-ACCTCTAAATCCGACCCG | |
| RKR19b | GCGG-gCtCGTCcAAGTG-ACCTCTTCTTCCGACCCG | |
| RKW10a | GCGc-gtCCGctccGaTG-ACCctaAAAaTCTACTGA | |
| RKW10b | GCGg-acCCGtctaGgTG-ACCtctAAAtTCTACTGA | |
| RKW20a | GCGc-aCCCGTtCAGAAAGGTCTCTTCTTTCTACTGA | |
| RKW20b | GCGg-gCCCGTcTAGGAG-ATCTCATCTTCCTGCCCG | |
| RKW5a | GCGG-ATcCGTCTAAGTG-ACCTCTTCTTCCTGCCCG | |
| RKW5b | GCGG-ATtCGTCTAAGTG-ACCTCATCTTCCTGCCCG | |
| RKW6b | GCGc-ACTCGcttAaaTG-ACCCTTAAAATCGGTCCG | |
| RKW6a | GCGg-ACTCGtccAggTG-ACCTCTTCTACCTGCCCG | |
| RKW8b | GCGG-ACCCGTCTAAGTG-ACCTCTAAATCCTGCCCG | RKW9a |
| RKW8a | GCGG-ACCCGTTCAAAAAGGTCTTAAAAATCTACTGA | |
| RKW9a | GCGG-ACCCGtCTAAGTG-ACCTCTAAATCCTGCCCG | RKW8b |
| RKW9b | GCGG-GCCCGcCTAAGTG-ACCTCTTCTTCCTGCCCG | |

**Table 6:** Population genetic parameters for sample locations with 4 or more individuals. For all polymorphisms identified, alleles exclusive to populations are noted. For each location and subspecies, the loci present in Hardy-Weinberg (H-W) equilibrium ($p \geq 0.05$), loci found to be monomorphic, and the number of haplotypes calculated for the 45 SNPs and DIPs used for initial haplotype inference are indicated. Distribution of monomorphic individuals by population and subspecies are shown.

| | Individuals | Exclusive Alleles | H-W Loci | Monomorphic Loci | Haplotypes Identified | Monomorphic Individuals |
|---|---|---|---|---|---|---|
| **ECL** | 6 | 2 | 36 | 8 | 11 | 1 |
| **EWW** | 4 | 8 | 38 | 6 | 8 | 0 |
| **MCS** | 5 | 2 | 31 | 11 | 6 | 2 |
| **MSL** | 5 | 4 | 39 | 6 | 9 | 0 |
| **REN** | 4 | 1 | 35 | 4 | 8 | 0 |
| **RKW** | 6 | 5 | 37 | 6 | 12 | 0 |
| **Eastern** | 16 | 32 | 25 | 0 | 29 | 2 |
| **Merriam's** | 13 | 9 | 19 | 1 | 19 | 2 |
| **Rio Grande** | 11 | 16 | 27 | 1 | 22 | 0 |
| **Total** | 40 | 57 | | 0 | 70 | 4 |

**Figure 1:** Historical geographical distribution of wild turkey subpopulations. The Eastern, Rio Grande, Merriam's subpopulation is represented in green, red and blue, respectively. Florida (orange) and Goulds (yellow) subpopulations were not included in the current study. The Florida subspecies is not genetically distinct from the Eastern subspecies (167). (*Adapted from Mock et al. 2002 (167)*)

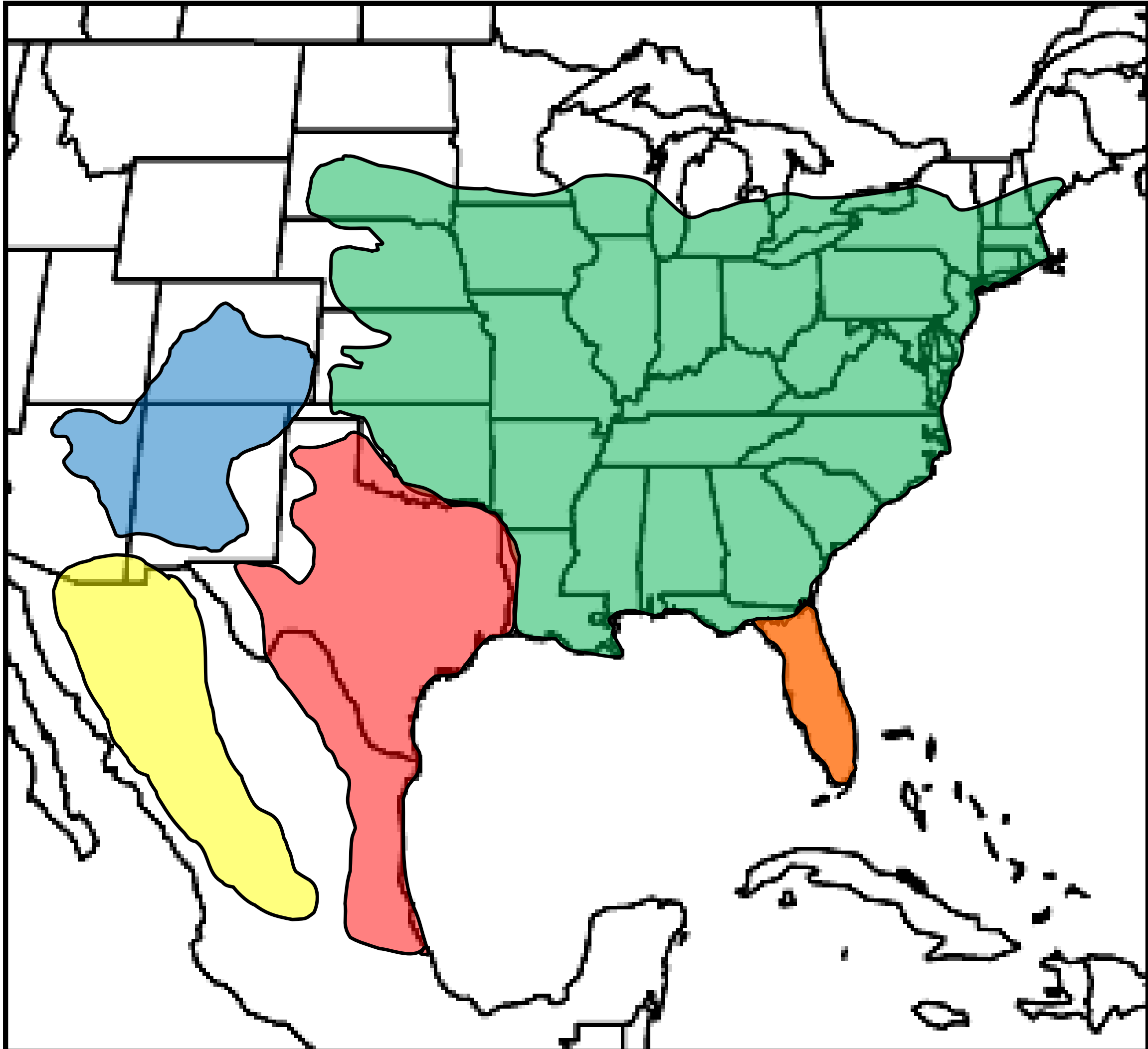


**Figure 2:** Minor allele frequencies of SNPs and DIPs identified by resequencing ~9kb of the MHC-B locus in 40 wild turkeys.

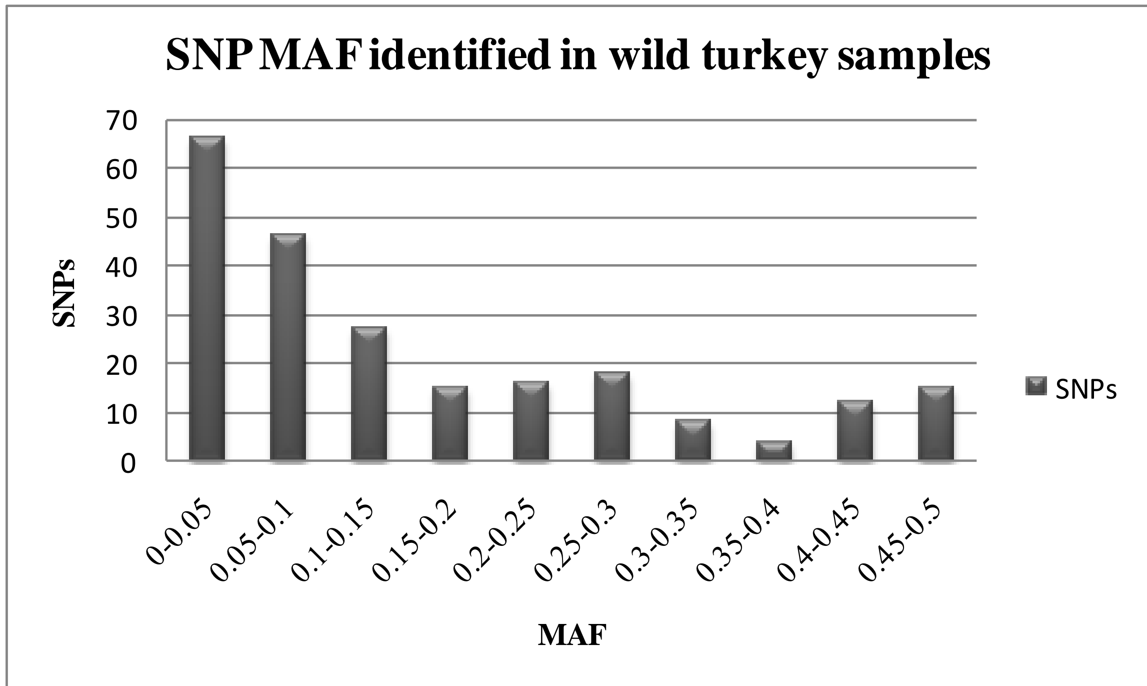


**Figure 3:** Allele frequencies from 45 loci used in haplotype inference within 3 subspecies of wild turkey. Blue, red, and green indicate proportion with genotype homozygous major allele, heterozygous, and homozygous minor allele, respectively. E-astern, M-merriam's, and R-io Grande. Approximate position within the MHC-B is indicated.

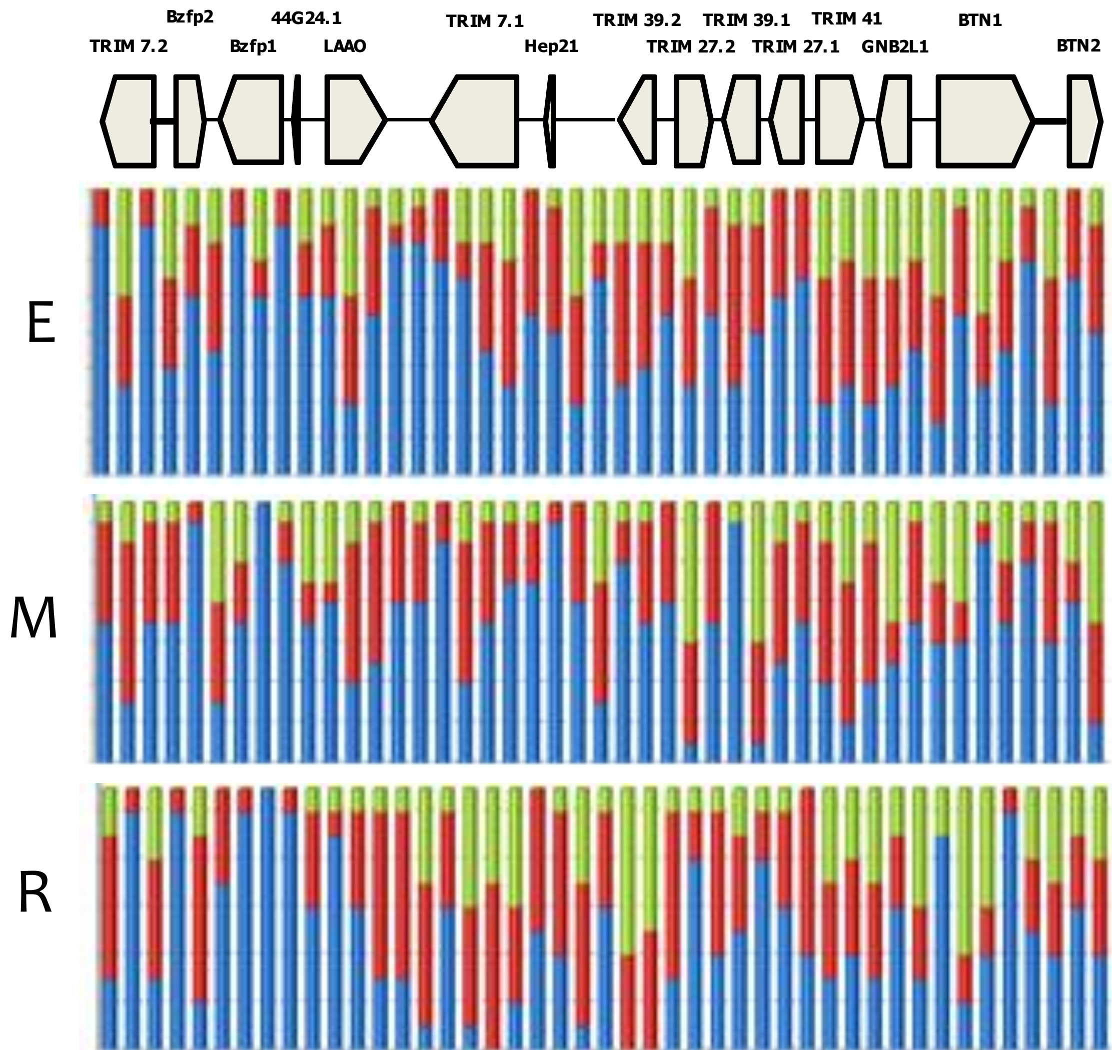


**Figure 4:** The limited extent of linkage disequilibrium across the MHC-B region identified from wild turkeys. D' values (X 100) given for all pairwise comparisons < 1.0. Values not given for blocks in complete LD (D' = 1.0). Block colors correspond to D'/LOD, where D' < 1 and LOD < 2 is white, D' < 1 and LOD ≥ 2 is shades of pink/red, D' = 1 and LOD < 2 are grey, and D' = 1 and LOD ≥ 2 are bright red.

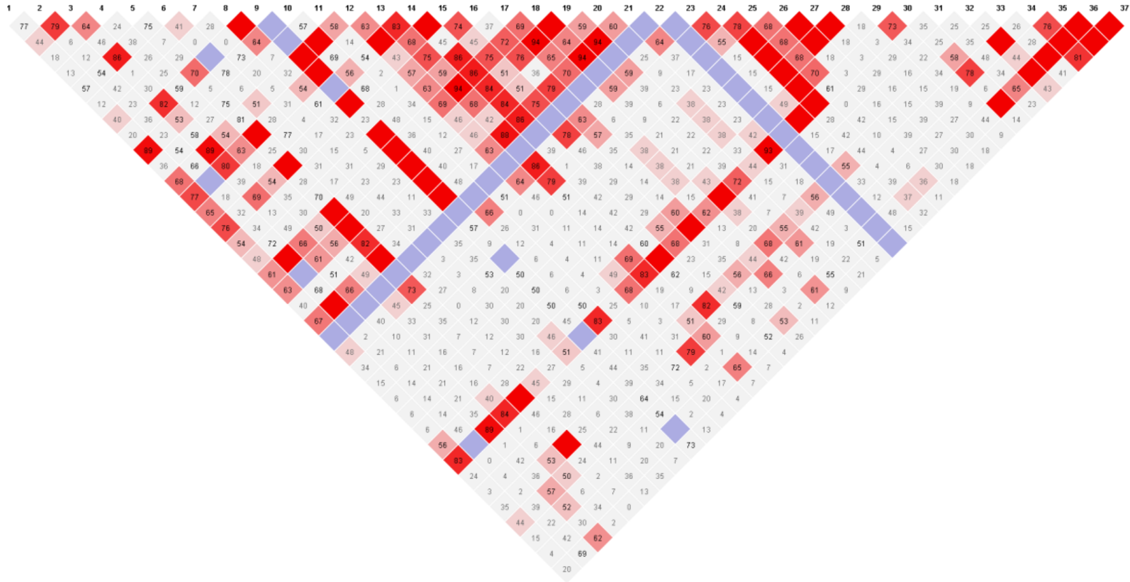
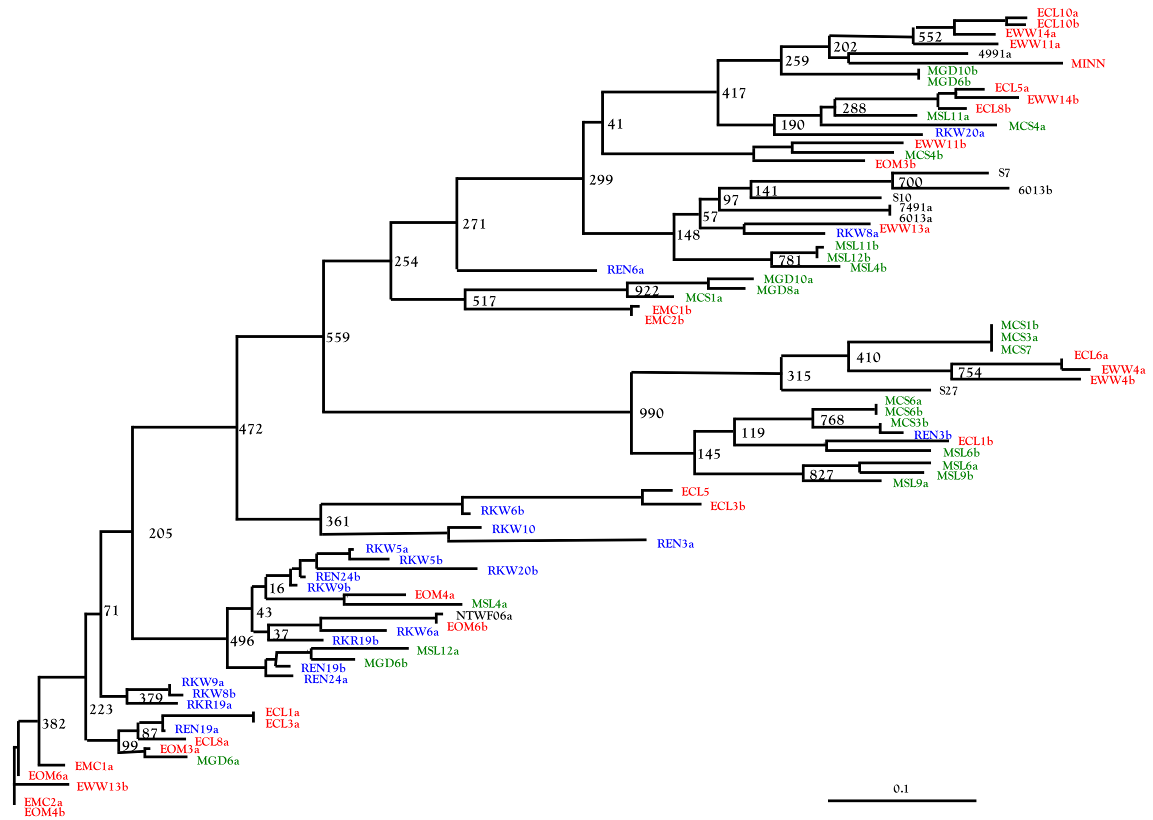


**Figure 5:** Unrooted tree depicting phylogenetic relationships of inferred haplotypes from wild and selected commercial turkeys. Haplotypes originating in Eastern, Merriam's, and Rio Grande subspecies are indicated in red, green, and blue, respectively. Commercial haplotypes are indicated in black.



# Chapter 7

# Genome-Wide Genetic Diversity of "Nici", the DNA Source for the CHORI-260 Turkey BAC Library and Candidate for Whole Genome Sequencing

# Genome-Wide Genetic Diversity of "Nici", the DNA Source for the CHORI-260 Turkey BAC Library and Candidate for Whole Genome Sequencing[28]

L. D. Chaves[1], D. E. Harry[2] and K. M. Reed[1].

[1]Department of Veterinary and Biomedical Sciences, University of Minnesota, St. Paul, MN 55108, USA

[2] Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA

Published as: L. D. Chaves, D. E. Harry and K. M. Reed. Genome-Wide Genetic Diversity of "Nici", the DNA Source for the CHORI-260 Turkey BAC Library and Candidate for Whole Genome Sequencing. *Animal Genetics* (2009).

[28] In Press, Animal Genetics, 2009

**SUMMARY**

Vertebrate whole genome sequence assembly can benefit from *a priori* knowledge of variability in the target genome, with researchers often selecting highly inbred individuals for sequencing. However, for most species highly inbred research lines are lacking requiring the use of an outbred individual(s). Here we examined the source DNA (Nici) of the CHORI-260 turkey BAC library through analysis of microsatellites and BAC sequences. Heterozygosity of Nici was compared with that of individuals from several breeder lines. Seventy-eight microsatellites were screened for polymorphism on a total of 43 birds identifying an average individual heterozygosity of 0.39 with Nici at 0.35. Additional loci (total of 147) were examined on a subset of individuals to obtain better genome coverage. The mean heterozygosity for this subset was 0.33 with Nici at 0.31. Examination of approximately 200 kb of genome sequence identified single nucleotide polymorphisms in the order of 1 per 200 bp in Nici. These data suggest the heterozygosity of Nici is comparable to other birds of selected breeder lines and whole genome sequencing would result in an abundant resource of genome wide polymorphisms.

**Keywords** microsatellites, heterozygosity, whole genome sequencing, single nucleotide polymorphism, *Meleagris gallopavo*

Whole genome characterization in many species has comprised a course of genetic linkage mapping, followed by physical mapping through bacterial artificial chromosome (BAC) library and contig development, prior to whole genome sequencing and assembly. Linkage mapping requires analysis of polymorphic DNA markers (RFLP, microsatellite, SNPs) segregating in genetically diverse, pedigreed resource populations. In contrast, BAC libraries are generated from the DNA of a single individual. Fingerprint analysis and subsequent contig assembly of the libraries is simplified when polymorphism in that DNA source is limited. Whole genome sequencing requires the generation of millions of individual sequence reads ordered and assembled into contigs assigned to chromosomes, and likewise, benefits from a lack of sequence polymorphism. Linkage and BAC contig maps are indispensable in the final finishing and assembly of the whole genome sequencing. Ideally the DNA used for BAC library generation and whole genome sequencing should be from the same individual. Assembly gaps and errors can be rapidly resolved through the selected analysis of specific BAC clones within such regions. In some cases, highly inbred individuals were identified to facilitate whole genome sequencing efforts (179-181).

The turkey is well positioned for whole genome sequencing. Extensive comparative analysis has shown the turkey and chicken genomes to be highly homologous in sequence and syntenic in locus order (72,75) with a limited number of chromosomal rearrangements as identified by linkage and physical mapping (74,75). Recent mapping efforts have placed 95% of turkey sequences to homologous chicken genome regions (75), suggesting that the chicken genome can be used as scaffolding to assemble a polymorphic whole genome turkey sequence. A dense, integrative and

comparative linkage map is available for the turkey (75) and a BAC library has been developed. Generation of a whole genome BAC contig map is underway (76).

The CHORI-260 turkey BAC library was constructed by Michael Nefedov in Pieter de Jong's laboratory at BACPAC Resources, Children's Hospital Oakland Research Institute. The library was generated from DNA isolated from a female turkey (donated by Nicholas Turkey Breeding Farms) identified as NT-WF06-2002-E0010, herewith referred to as "Nici" (Nicholas inbred). Nici is from an inbred subline (i.e. sib-mating for nine generations) originally derived from a commercially significant breeding line. From her pedigree, Nici has an increased inbreeding coefficient of 0.624 relative to the founder breeding line. Nici is also being considered as the individual source DNA for whole genome sequencing of turkey. Anticipating these future efforts, here we evaluate the genome-wide genetic diversity of Nici in microsatellite and BAC sequences.

Microsatellite markers identified as polymorphic in previous studies (87) were genotyped for Nici and six individuals from each of five Nicholas sire and dam lines. Genotypes were determined using $^{33}P$ labeled PCR products electrophoresed on acrylamide gels as previously described (182). Additional individuals from the UMN/NTBF mapping population (see Figure 1 in Reed *et al.* 2003a) were included for comparison, namely: the founder parents (P generation, including six females and four males from two separate lines) and two F1 males (D3804 and D7491).

Heterozygosity (H) refers to the proportion of loci within an individual that are heterozygous. Individual genome-wide heterozygosity can be estimated with molecular markers by determining the observed frequency of heterozygous loci (represented as *h*). The correlation between H and *h* is directly effected by the number of loci examined and

only approaches 1.0 when the number of loci examined is relatively large as in dense genome-wide SNP surveys (183). In population genetics, H typically refers to the fraction of individuals within the population heterozygous at a particular locus. Herein we denote the estimated mean population heterozygosity as $\hat{H}$ to distinguish this from estimated individual heterozygosity (*h*).

## MICROSATELLITE VARIATION

Initially, heterozygosity was examined by genotyping 78 microsatellites on a total of 43 individual turkeys including the grandparent lines and two F1 sires by counting number of heterozygous loci per individual (Table 1). On average, 39% of the 43 individuals were heterozygous at any individual marker (range 5-81%). This value however, does not reflect overall heterozygosity in turkeys because we studied only markers previously identified as polymorphic.

Among all loci, individual birds had a mean *h* of 0.39, normally distributed (R= 0.98) from 0.19 to 0.60. Z-score analysis ($\alpha = 0.05$) identified two individuals (C5025 and C6002, both P generation parents in the mapping population) with significantly less heterozygosity, and three individuals (two from Line 4, and one from line 5), which were significantly more heterozygous than the mean. Interestingly, neither F1 male was significantly different from the mean. Among the lines examined, $\hat{H}$ ranged from 0.37 to 0.47 with an overall mean value of 0.42 (Table 2). Specific allele frequencies and distributions were not calculated due to the small sample size from each line. Analysis of variance and subsequent Scheffe testing between the breeder lines and P lines identified a significant difference in heterozygosity between Line 4 (highest) and the male P line

(lowest, p = 0.005). These results are consistent with the known breeding history of these lines (internal data, Nicholas Turkey Breeding Farms). No significant difference in heterozygosity (chromosomal or overall) was observed between Nici ($h$ = 0.35) and the breeder lines, the two P generation lines, or the F1s.

In order to provide a more thorough analysis of genome variation, an additional suite of sixty-nine markers was genotyped on Nici, the P generation and two F1 sires. The larger data set (147 markers) allowed for further partitioning the variation by linkage group. Within the macrochromosomes (MGA1-9), $\hat{H}$ ranged from 0.22 to 0.54 (Table 3) with Nici often falling within one standard deviation of the P line means (Figure 1). Overall, $h$ ranged from 0.22 (C6002-P male line) to 0.50 (D3804-F1 hybrid) and had a mean of 0.33, with Nici at 0.31. Higher heterozygosities would be expected in the F1 individuals. Compared to the overall mean, heterozygosities for Nici were higher for some macrochromosomes (1, 2, 5, 7, and 8) and lower for others (3, 4, 6, and 9). Nici was monomorphic for multiple markers on chromosome 6 as well as M11 (equivalent to p-arm of GGA4). We acknowledge however, that $\hat{H}$ in the smaller linkage groups is based on a small number of markers. A combined analysis of all loci on linkage groups represented by less than four markers had a mean heterozygosity of 0.40 ($\sigma$ = 0.11) and varied between 0.24 and 0.62 among individuals (Nici = 0.29). Taken as a whole, these data suggest the heterozygosity of Nici is on average lower than birds of selected breeder lines and the F1 hybrids.

**SEQUENCE VARIATION**

In addition to microsatellite polymorphism, DNA sequence polymorphisms were identified through analysis of overlapping assembled BAC clones from the NIH NISC Comparative Vertebrate Sequencing effort[29]. Assembled BAC sequences retrieved from NCBI were aligned into contigs and analyzed using Sequencher software (Gene Codes, Corp.). Polymorphisms identified in the aligned contigs were verified through manual reanalysis of the original electropherograms. Of the 20 overlapping BAC clones included in the NISC dataset (encompassing 1.09 Mb) only one contig (4 clones) within MGA6 was polymorphic (BAC clones derived from alternate chromosomes). Interestingly, MGA6 was identified as monomorphic in Nici at all 4 microsatellite loci tested. It is unlikely that the 16 invariant BAC clones represent monomorphic regions within Nici's genome.

Examination of nearly 200 kb of overlapping sequence within the bichromosomal BAC contig[30] (contig comprised of BACs derived from the two alternate chromosomes) identified 1043 variable nucleotides, or ~1 SNP/200 bp. This number is similar to the level found in other resequencing efforts of turkeys (154). Assembly of the equine genome identified one polymorphism approximately every 1.5 kb, whereas in the chicken whole genome sequencing only one polymorphism was initially identified per 5000 bps (180,181). Polymorphism in Nici may cause some difficulty in a whole genome shotgun sequencing assembly. For example, assemblers may be unable to distinguish between allelic sequences and repeated blocks within the turkey genome, especially those not present in the chicken. To overcome this, deeper sequencing coverage with newer

[29] Genbank accession #: AC15289 ,AC155203, AC153078, AC148423, AC157842, AC148371, AC154064, AC159922, AC154927, AC148424, AC148419, AC154059, AC154060, AC154925, AC148421, AC148422, AC148370, AC159921, AC154063, AC154926
[30] Genbank accession #: AC148423, AC157842, AC152897, AC148421

sequencing technologies, close examination of paired-end reads, as well as additional genetic and/or physical mapping may be necessary.  Although some birds examined in this study were found to be less heterozygous than Nici these differences were not significant.  Moreover, resources already established using DNA from Nici outweigh the concern of polymorphism.  Use of Nici for whole genome sequencing may have an additional benefit, the identification of nearly 5 million SNP loci for fine mapping and association studies of turkey traits.

**Acknowledgements**

This research was supported by grants from the Cooperative State Research, Education and Extension Service, U.S. Department of Agriculture (2001- 35205-9494 and 2004-35205-14217).

**Table 1:** Number of alleles identified at microsatllite loci in the turkey. For each marker the Genbank accession number is given. Asterisks denote markers where locations were identified through chicken BLAST homologies; ie. not genetically mapped.

| Marker | Accession | Turkey LG | 7491 | 3804 | Nici | Line 1 | | | | | | Line 2 | | | | | | Line 3 | | | | | | Line 4 | | | | | | Line 5 | | | | | | UMN/NTBF P DAM/SIRE | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | S11 | S12 | S13 | S14 | S15 | S16 | S17 | S18 | S19 | S20 | S21 | S22 | S23 | S24 | S25 | S26 | S27 | S28 | S29 | S30 | C4974 | C5025 | C4914 | C4991 | C4973 | C5017 | C6013 | C6014 | C6002 | C6012 |
| **MNT092** | AY235102 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT060** | AY235073 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 |
| **MNT126** | AF540429 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 0 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT036** | AY235049 | 1 | 2 | 2 | 2 | 1 | 2 | 0 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 0 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 |
| **MNT144** | AF540436 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 |
| **MNT109** | AF540412 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT105** | AF540408 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT128** | AF540430 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| **MNT054** | AY235067 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT048** | AY235061 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 |
| **MNT084** | AY235095 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| **MNT160** | AY235129 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 0 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 |
| **MNT037** | AY235050 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 |
| **MNT049** | AY235062 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 |
| **MNT106** | AF540409 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 |
| **MNT032** | AY235045 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 |
| **MNT026** | AY235039 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 |
| **MNT089** | AY235100 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 |
| **MNT129** | AF540431 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT153** | AY235122 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 |
| **MNT132** | AY235108 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 0 | 1 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT070** | AY235083 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 |
| **MNT045** | AY235058 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 |
| **MNT083** | AY235094 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT086** | AY235097 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT155** | AY235124 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT080** | AY235091 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 |
| **MNT050** | AY235063 | 2 | 2 | 2 | 1 | 1 | 0 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| **MNT047** | AY235060 | 3 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 |
| **MNT041** | AY235054 | 3 | 2 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 |
| **MNT157** | AY235126 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 1 |
| **MNT146** | AF540438 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 |
| **MNT150** | AY235119 | 4 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT079** | AY235090 | 4 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 |
| **MNT142** | AF540434 | 4 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 |
| **MNT087** | AY235098 | 4 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT065** | AY235078 | 4 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT039** | AY235052 | 5 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 |
| **MNT108** | AF540411 | 5 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT156** | AY235125 | 5 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 |
| **MNT052** | AY235065 | 5 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 0 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 |
| **MNT042** | AY235055 | 5 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 0 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT163** | AY235132 | 6 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 |
| **MNT040** | AY235053 | 7 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 0 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 0 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 |
| **MNT085** | AY235096 | 7 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 1 |
| **MNT115** | AF540418 | 8* | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 2 |
| **MNT095** | AF540401 | 8 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 |
| **MNT103** | AF540406 | 8 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 0 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 |
| **MNT024** | AY235037 | 8 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT071** | AY235084 | 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 |
| **MNT078** | AY235089 | 8 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 |
| **MNT111** | AF540414 | 9 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 1 |
| **MNT062** | AY235075 | 9 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 0 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT066** | AY235079 | 9 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT159** | AY235128 | 9 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT031** | AY235044 | M10 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 |
| **MNT101** | AF540404 | M11 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 |
| **MNT117** | AF540420 | M11 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 |
| **MNT058** | AY235071 | M12 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 |
| **MNT114** | AF540417 | M12 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 |
| **MNT061** | AY235074 | M14 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 |
| **MNT056** | AY235069 | M14 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 |
| **MNT122** | AF540425 | M17 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| **MNT063** | AY235076 | M17 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 |
| **MNT051** | AY235064 | M17 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT068** | AY235081 | M21 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 0 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 |
| **MNT088** | AY235099 | M21 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 |
| **MNT120** | AF540423 | M24 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 |
| **MNT043** | AY235056 | M24 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 |
| **MNT124** | AF540427 | M24 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT112** | AF540415 | M24 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT073** | AF540399 | M28 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 |
| **MNT075** | AY235086 | M28 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 2 |
| **MNT107** | AF540410 | M29 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 1 |
| **MNT082** | AY235093 | M35 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 |
| **MNT119** | AF540422 | M40 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT044** | AY235057 | M09 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| **MNT029** | AY235042 | M09 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 0 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| **MNT057** | AY235070 | Z | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 |
| **MNT033** | AY235046 | Z | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 |
| **MNT143** | AF540435 | Z | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 |

| | | | | | | Line 1 | | | | | | Line 2 | | | | | | Line 3 | | | | | | Line 4 | | | | | | Line 5 | | | | | | UMN/NTBF P DAM/SIRE | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Marker | Accession | Turkey LG | 7491 | 3804 | Nici | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | S11 | S12 | S13 | S14 | S15 | S16 | S17 | S18 | S19 | S20 | S21 | S22 | S23 | S24 | S25 | S26 | S27 | S28 | S29 | S30 | C4974 | C5025 | C4914 | C4991 | C4973 | C5017 | C6013 | C6014 | C6002 | C6012 |
| **MNT196** | AY235164 | 1 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT174** | AY235142 | 1 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT219** | AY235184 | 1 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 |
| **MNT232** | AY235197 | 1 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 |
| **MNT181** | AY235149 | 1 | 1 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT192** | AY235160 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 |
| **MNT200** | AY235167 | 2 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 |
| **MNT184** | AY235152 | 2 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT230** | AY235195 | 2 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 1 |
| **MNT217** | AY235182 | 2 | 2 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 |
| **MNT170** | AY235139 | 2 | 1 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 2 |
| **MNT224** | AY235189 | 2 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT205** | AY235172 | 2 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 |
| **MNT185** | AY235153 | 3 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT187** | AY235155 | 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 |
| **MNT199** | AY235166 | 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 2 |
| **MNT137** | AY235113 | 3 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT190** | AY235158 | 3 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT169** | AY235138 | 3 | 1 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 2 |
| **MNT167** | AY235136 | 4 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 2 |
| **MNT216** | AY235181 | 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT189** | AY235157 | 4 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 |
| **MNT168** | AY235137 | 4* | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 |
| **MNT429** | AL592644 | 4 | 2 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT193** | AY235161 | 5 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 |
| **MNT183** | AY235151 | 5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT207** | AY235174 | 5 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 1 |
| **MNT229** | AY235194 | 5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 |
| **MNT206** | AY235173 | 6 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| **MNT231** | AY235196 | 6 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 |
| **MNT164** | AY235133 | 6 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 |
| **MNT202** | AY235169 | 7 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 |
| **MNT188** | AY235156 | 7 | 1 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT225** | AY235190 | 7 | 1 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT180** | AY235148 | 7 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| **MNT176** | AY235144 | 7 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 |
| **MNT140** | AY235116 | 8 | 1 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 |
| **MNT215** | AY235180 | 8 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 |
| **MNT197** | AY235165 | 8 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT233** | AY235198 | 8 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT194** | AY235162 | 9 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT179** | AY235147 | 9 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT204** | AY235171 | 9 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 |
| **MNT116** | AF540419 | 9 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 |
| **MNT228** | AY235193 | 9 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT201** | AY235168 | M09 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT403** | AL592737 | M10 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 |
| **MNT221** | AY235186 | M10 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| **MNT173** | AY235141 | M10 | 2 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 |
| **MNT172** | AY235140 | M11 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 1 |
| **MNT149** | AY235118 | M11 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 |
| **MNT211** | AY235176 | M11 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 1 |
| **MNT226** | AY235191 | M12 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 |
| **MNT223** | AY235188 | M12 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 |
| **MNT178** | AY235146 | M12 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 2 |
| **MNT360** | AY552929 | M13 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 |
| **MNT294** | AY552866 | M13 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 |
| **MNT479** | AL593793 | M16 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 |
| **MNT450** | AL593308 | M16 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 |
| **MNT208** | AY235200 | M17 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 |
| **MNT210** | AY235175 | M17 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 |
| **MNT409** | AY926543 | M19 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 |
| **MNT404** | AL593651 | M19 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 |
| **MNT282** | AY552854 | M20 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT430** | AL592650 | M23 | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT322** | AY552892 | M23 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT222** | AY235187 | M24 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| **MNT186** | AY235154 | M24 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 1 |
| **MNT136** | AY235112 | M31 | 1 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MNT212** | AY235177 | Z | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 |

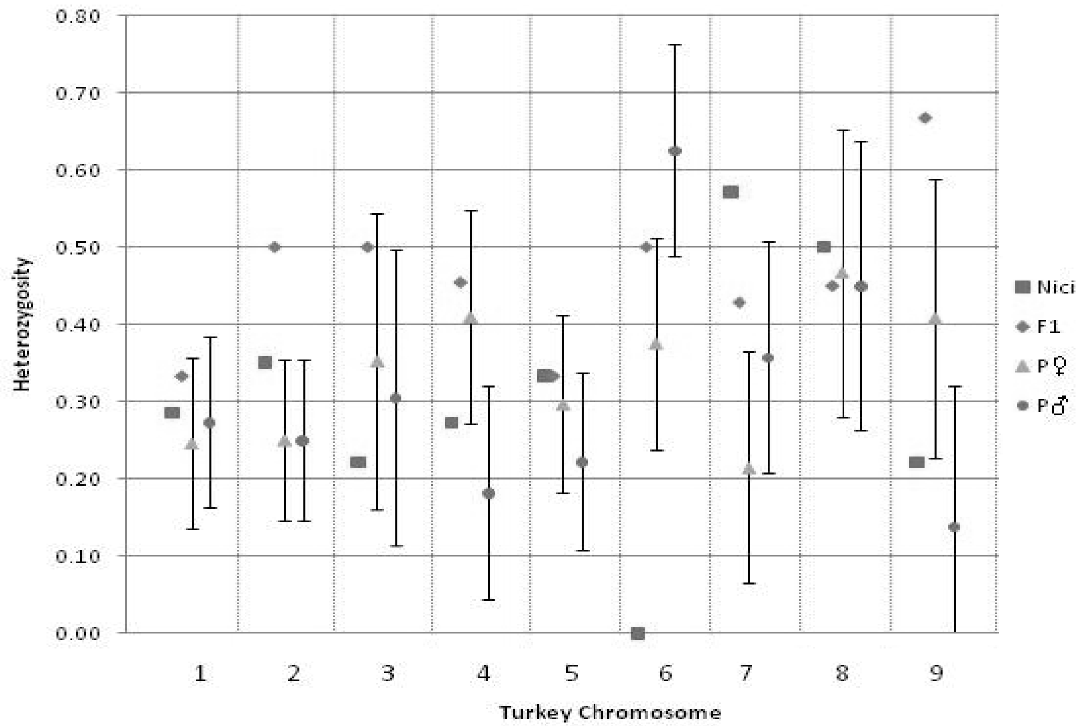
**Table 2:** Summary of heterozygosity scores determined by genotyping 78 microsatellite markers in 43 individuals. Tabulated results (below) represent the sampled groups (left, sample size in parenthesis) or individuals (F1 sires and Nici).

| LGs | Markers | Line 1 (6) | Line 2 (6) | Line 3 (6) | Line 4 (6) | Line 5 (6) | $P_{FEM}$ (6) | $P_{MALE}$ (4) | | D7491 | D3804 | Nici | | $\hat{H}_{LG}$ | Std $Dev_{LG}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Macro | 55 | 0.36 | 0.42 | 0.42 | 0.49 | 0.44 | 0.33 | 0.29 | | 0.38 | 0.55 | 0.33 | | 0.40 | 0.10 |
| Micro | 23 | 0.41 | 0.44 | 0.43 | 0.41 | 0.31 | 0.33 | 0.30 | | 0.26 | 0.43 | 0.39 | | 0.38 | 0.10 |
| | $\hat{H}_{LINE}$ | 0.37 | 0.43 | 0.42 | 0.47 | 0.40 | 0.33 | 0.29 | *h* | 0.35 | 0.51 | 0.35 | $\hat{H}$ | 0.39 | 0.08 |
| | Std Dev | 0.04 | 0.04 | 0.04 | 0.05 | 0.11 | 0.08 | 0.13 | | | | | | | |

**Table 3:** Summary of individual heterozygosities determined by genotyping 147 microsatellite markers across the P generation and F1 males of the UMN/NTBF population and Nici.

| | | $P_{FEMALE}$ | | | | | | $P_{MALE}$ | | | | F1 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LG | markers | C4974 | C5025 | C4914 | C4991 | C4973 | C5017 | C6013 | C6014 | C6002 | C6012 | D7491 | D3804 | Nici | $\hat{H}_{LG}$ | Std $Dev_{LG}$ |
| 1 | 21 | 0.29 | 0.19 | 0.24 | 0.24 | 0.43 | 0.1 | 0.24 | 0.33 | 0.33 | 0.19 | 0.33 | 0.33 | 0.29 | 0.27 | 0.09 |
| 2 | 20 | 0.3 | 0.05 | 0.3 | 0.25 | 0.25 | 0.35 | 0.3 | 0.4 | 0.15 | 0.15 | 0.4 | 0.6 | 0.35 | 0.3 | 0.14 |
| 3 | 9 | 0.11 | 0.22 | 0.56 | 0.22 | 0.56 | 0.44 | 0.33 | 0.33 | 0 | 0.56 | 0.44 | 0.56 | 0.22 | 0.35 | 0.19 |
| 4 | 11 | 0.55 | 0.36 | 0.18 | 0.55 | 0.45 | 0.36 | 0.18 | 0.27 | 0 | 0.27 | 0.45 | 0.45 | 0.27 | 0.34 | 0.16 |
| 5 | 9 | 0.33 | 0.33 | 0.44 | 0.22 | 0.33 | 0.11 | 0.11 | 0.33 | 0.22 | 0.22 | 0.22 | 0.44 | 0.33 | 0.28 | 0.11 |
| 6 | 4 | 0.25 | 0.5 | 0.5 | 0.25 | 0.25 | 0.5 | 0.75 | 0.5 | 0.5 | 0.75 | 0.5 | 0.5 | 0 | 0.44 | 0.21 |
| 7 | 7 | 0.14 | 0.43 | 0.29 | 0.14 | 0.29 | 0 | 0.71 | 0.57 | 0.14 | 0 | 0.14 | 0.71 | 0.57 | 0.32 | 0.26 |
| 8 | 10 | 0.8 | 0.3 | 0.3 | 0.5 | 0.5 | 0.4 | 0.6 | 0.5 | 0 | 0.7 | 0.2 | 0.7 | 0.5 | 0.46 | 0.22 |
| 9 | 9 | 0.33 | 0.44 | 0.67 | 0.44 | 0.44 | 0.11 | 0.22 | 0.22 | 0 | 0.11 | 0.56 | 0.78 | 0.22 | 0.35 | 0.23 |
| M09 | 3 | 0 | 0.67 | 0.33 | 0 | 0.33 | 0 | 0 | 0 | 0 | 0.67 | 0 | 0 | 0.33 | 0.18 | 0.26 |
| M10 | 4 | 0.25 | 0.25 | 0.25 | 0.25 | 0.5 | 0.25 | 0.25 | 0 | 0 | 0.25 | 0.75 | 0.5 | 0.5 | 0.31 | 0.21 |
| M11 | 5 | 0.6 | 0.2 | 0.4 | 0.6 | 0.8 | 0.4 | 0.2 | 0.4 | 0.6 | 0.2 | 0.6 | 0.2 | 0 | 0.4 | 0.23 |
| M12 | 5 | 0.4 | 0.2 | 0.6 | 0.2 | 0.8 | 0.8 | 0.2 | 0.4 | 0.8 | 0.2 | 0.6 | 0.4 | 0.2 | 0.45 | 0.25 |
| M13 | 2 | 0.5 | 0 | 0 | 0.5 | 0 | 0.5 | 1 | 0.5 | 0.5 | 0.5 | 1 | 0.5 | 0 | 0.42 | 0.34 |
| M14 | 2 | 0.5 | 0.5 | 1 | 0.5 | 1 | 1 | 0 | 0.5 | 0 | 0 | 0.5 | 1 | 0.5 | 0.54 | 0.38 |
| M16 | 2 | 0.5 | 0 | 0 | 0.5 | 0.5 | 1 | 0 | 0.5 | 0.5 | 0.5 | 0 | 0.5 | 0 | 0.35 | 0.32 |
| M17 | 5 | 0.2 | 0.4 | 0.2 | 0.4 | 0 | 0 | 0 | 0.8 | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.25 | 0.22 |
| M19 | 2 | 0 | 1 | 1 | 0.5 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0.65 | 0.47 |
| M20 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0.38 | 0.51 |
| M21 | 2 | 0 | 0 | 0.5 | 1 | 0 | 0.5 | 0 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.38 | 0.3 |
| M23 | 2 | 0 | 0.5 | 0.5 | 0.5 | 0.5 | 0 | 1 | 1 | 0.5 | 0.5 | 0 | 0.5 | 0 | 0.42 | 0.34 |
| M24 | 6 | 0.33 | 0 | 0 | 0.17 | 0.17 | 0.17 | 0 | 0.5 | 0.17 | 0.33 | 0 | 0.17 | 0.5 | 0.19 | 0.18 |
| M28 | 2 | 0.5 | 0.5 | 0.5 | 0 | 0.5 | 0.5 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0.5 | 0.41 |
| M29 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0.62 | 0.51 |
| M31 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0.08 | 0.28 |
| M35 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0.31 | 0.48 |
| M40 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.08 | 0.28 |
| | *h* | 0.33 | 0.26 | 0.34 | 0.31 | 0.4 | 0.28 | 0.31 | 0.41 | 0.22 | 0.31 | 0.37 | 0.5 | 0.31 | 0.33 | 0.07 |

**Figure 1:** Heterozygosity comparisons of the turkey macrochromosomes of P generation, F1 hybrids, and Nici. Means and standard deviations are given for the two lines representing the P generation.



# Chapter 8

# Conclusions and Future Studies

At the outset of this project decades of research had generated interesting conclusions regarding the chicken major histocompatibility complex (MHC). The chicken MHC appeared to be split between two regions (29). The MHC-B contained most of the loci traditionally associated with graft rejection and immune function (147). An additional, genetically unlinked region located distally on the same microchromosome (MHC-Y) was also found to contain MHC-like loci with non-classical immune functions (42). A 92 kb sequence assembly suggested that the chicken MHC-B was small, compact, and gene dense with a unique organization in comparison to mammalian MHCs (147). This locus was further unique compared to mammals in that both Class I and Class IIβ loci were found close together with many of the loci encoding antigen processing machinery. A small portion of the chicken MCH-Y (~20 kb) was sequenced and assembled identifying a Class I-like locus and two lectin-like loci (43).

Chicken MHC-B haplotypes have repeatedly been associated with pathogen resistance and/or susceptibility (50,51,54,55). The MHC-B locus can influence the outcome to viral, bacterial, and parasitic infections. Further, in some instances MHC-Y haplotypes were identified to effect infectious outcomes as well, however to a lesser extent (44,45). One of the most interesting and well studied interactions of MHC haplotype and infectious outcome is that of the relative resistance of the B21 haplotype to Marek's disease virus (147).

In late 2004 the first assembly of the chicken (an inbred laboratory strain [UCD-001], B21 MHC haplotype of Red Jungle Fowl; RJF-256) whole genome sequence (WGS) was published in Nature (180). Unfortunately, the chicken MHC chromosome (GGA16) was the least well assembled in the genome initially spanning ~200 kb and containing many gaps and ambiguities. The ensuing build 2.1 (May 2006) assembled a region spanning ~430 kb. However, again this assembly was hindered by numerous gaps, ambiguities and is likely misassembled, despite a further BAC shotgun sequencing effort (CHORI 261 79N09). Close examination of the raw data files (chromatograms/trace files) indicate why assembly of the region was incomplete. First, in both the whole genome and BAC subclone sequences, high-quality trace files were lacking within and flanking some of the Class I and Class IIβ loci. Additionally, the sequenced individual appears to be

heterozygous 5` upstream from the "core" MHC-B locus, near Bzfp1, as indicated by SNPs identified in the WGS where only a single allele was present in the BAC subclones.

Using data available in Genebank as well as small portions of the original 92 kb B-locus sequence (B12 haplotype), a monomorphic contig of 156 kb could be assembled. Further polymorphic sequence upstream was also assembled into additional contigs, however large DIPs in the genome limited proper assembly. In 2007 Shiina et al. published a 242 kb assembly of the RJF MHC utilizing overlapping BAC clones and long-range PCR products (30). This superior assembly, containing ~ 43 genes, benefitted from BAC clones of a single chromosome allowing easier assembly. Interestingly, a comparison of the BAC sequence from Shiina et al., CHORI 261 79N09, and WGS trace files proved the RJF was indeed heterozygous at the 5` distal end of the MHC-B region. Further addition of WGS traces to this assembly only provides an additional 1.2 and 2.3 kb to the 5` and 3` ends, respectively. Using the improved assembly, Hosomichi et al. (2008) presented an impressive study in which approximately 60 kb of the core chicken MHC-B was resequenced from 14 different haplotypes, identifying levels of mutation, recombination, and gene conversion among the haplotypes (36).

In this same period of time the genome of the turkey was just beginning to be examined with rapid progress. From initial linkage maps published in 2003 elucidated with restriction fragment length polymorphisms and microsatellite markers (80,149,184) to a large, integrated and comparative map of over 700 markers published in 2007 (75), studying the turkey genome has drawn international attention. The considerable homology between the turkey and chicken genomes has been well described (74,75). Several studies have begun to examine gene expression and the transcriptome of the turkey (152,185,186). Two BAC libraries from the same source DNA have been generated and characterized by hybridization (76). Fingerprint analysis and end sequencing is near completion (J Dogeson, pers comm.). With the support of these previous studies, the turkey whole genome sequence is now being assembled with the latest sequencing technology by an international consortium of scientists and ~4X coverage will be available to the research community by the summer of 2009. Plans for further genome sequencing as well as high density SNP arrays are now being proposed for the turkey. With all of the genomic tools available currently (or soon to be), it is an

exciting time to study the immunogenetics of the turkey. In particular, results presented in this thesis provide the opportunity to pursue many additional studies, some of which will now be discussed.

In Chapter 2 the turkey MHC-B and the genetically unlinked MHC-Y were found to reside on the same microchromosome. The NOR was located not between the two loci as previously suggested in the chicken (29), but rather at a distal location. Subsequent work in the chicken also positioned the NOR distal to the two loci (M. Miller, pers. comm.). In Chapter 2 many BAC clones were found to be positive for the 18S rDNA and only one was chosen for FISH mapping. While it is possible that some of the positive clones could represent other chromosomes with just a single or few copies of the 18S rDNA, it would be of interest to fingerprint and map the 18s containing clones to identify the NOR/MHC-Y locus boundary. The Y locus may indeed be quite large, and the single clone mapped in the turkey appears to be located at a boundary between the homologous mammalian class I/class III region. It would be prudent to map distal Y locus clones to estimate the size of the Y locus in the chromosome.

Chapter 3 presented approximately 200 kb of sequence from the core turkey MHC-B locus. Much of the sequence was highly similar to the chicken, with the exception of two additional BG-like loci and an additional Class IIβ locus. In this study, additional BAC clones extending 5` of the B-locus were identified and will be sequenced. End sequencing of the larger CHORI 260 BAC clone (56J17) indicates the 5` region extends into BG-like loci. This is especially important in that assembly of this region in turkey may actually assist the assembly of the heterozygous chicken WGS.

The CHORI 260 turkey library was exhaustively screened for overlapping 3` clones with C4 and CD1.1 probes without identifying a single clone. The additional turkey BAC library (TKN-MI) should be screened for clones containing the extended MHC regions from both the B and the Y. As new regions are examined, new sequence will be available for probe generation. Other large-insert cloning methods are available (Cosmid, Fosmid, Lambda) and have been used to study the avian MHC and could be employed in the turkey. The high degree of homology between the turkey and chicken also allow simple PCR methods to be used to generate flanking sequence. The ~30 kb chicken region from CYP21 to CD1A1 might easily be amplified (using chicken-based

primers) and sequenced in the turkey as was similarly done in Chapter 3. Indeed, a portion of the turkey CD1.1 locus has already been amplified and SNP mapped in the turkey. Additionally, initial screening of the 2x turkey whole genome sequence already provides some scaffolding in this region. Preliminary PCR efforts using chicken-based primers were unsuccessful in amplifying a turkey Class IIα homolog. Additional turkey whole genome sequencing will hopefully provide data to develop turkey-specific primers if not assemble the locus. Perhaps a BAC clone might be found for the locus as well.

Another interesting outcome of Chapter 3 was the identification of disparate expression between turkey MHC loci (Class I and Class IIβ). The sample used for the expression studies was an MHC heterozygote; however, the results were consistent at each locus between the two haplotypes. Class I expression appeared to be greater at the Iα2 locus. This is consistent with results in the chicken where the homologous BF2 locus is better expressed (92). In fact, in many chicken haplotypes the BF1 locus (Iα1 homolog) is not expressed at all. Class IIβ loci are even more interesting. The Class IIβ1 locus (BLB2 homolog) is the highest expressed in the two haplotypes. The Class IIβ3 locus was expressed at half that level. Class IIβ2 was not identified in any of the cDNA clones sequenced.

Additional expression studies using monomorphic samples with commercially prevalent haplotypes would be worthwhile to confirm loci presence and expression differences. Monomorphic samples would allow application of PCR techniques for single pass locus-specific resequencing of the exons encoding Class I and Class IIβ peptide binding regions without requiring multiple reads from cloned products. This would allow the assignment of alleles to SNP haplotypes. While the diversity of the regions proximal to the Class I and IIβ loci have been characterized, the levels of variation of the MHC antigens have not. Further, more precise, real time quantitative PCR approaches could then be performed by (potentially) eliminating allele sharing and decreased primer design restrictions. Such efforts could be extended into a similar resequencing study as that of Hosomichi et al. across the core B-locus (36). The genetic diversity of other loci like the highly polymorphic (and positively selected) NK cell receptor locus (Blec2) or the TAP proteins which are thought to co-evolve with Class I genes would also be of interest.

One of the most interesting characteristics of the avian MHC is the unique presence of the BG loci. BG loci are expressed on erythrocytes (as well as other cell types) but very little is known of their function (30). BG proteins are membrane spanning and highly polymorphic (32). They are variable in gene copy number in birds. The core MHC of chicken has one BG locus, the turkey has three, and the quail has at least eight (28,37). Upstream from the core MHC is a region with many BG loci yet to be fully described. Currently two additional loci have been identified in the chicken (30), with work currently underway to describe this region in the turkey. It might be of interest to develop genetic knockdowns/knockouts or chimeric reporter assays to try and assess a function. However, the high number of loci may make this difficult.

The MHC-Y BAC clone mapped and sequenced represents only a portion of the total Y locus. Two additional MHC-Y BAC clones have been identified and are currently being sequenced. As stated above, additional BAC library screening (and sequencing) should be completed to identify the additional Class I and Class II genes as well as screening with probe(s) for regions currently contained in the sequenced Y-clone (e.g. BAT1 or 3). Three to six Class I-like and two to four Class II-like loci can be identified in the RJF whole genome sequence (29), however the region is not a contiguous, well assembled sequence (likely due to heterozygosity) so the precise numbers cannot be determined. At least one more additional Class I-like locus (not present in the sequenced BAC clone) has been partially amplified and sequenced in Nici. Preliminary screening of the turkey whole genome sequence reads has identified at least one additional Class II-like sequence present in the turkey genome that is highly homologous to the Class-II loci of the chicken Y-locus. As more WGS for the turkey becomes available there will likely be more loci at least partially identified. For example, additional lectin-like loci have been identified in the chicken MHC-Y with an unclear immune function (43).

The Y-locus clone sequenced and mapped in Chapter 3 represents genomic sequence virtually unrepresented in the chicken WGS. Very few trace files are available for homologous sequence from the chicken. The turkey sequence identified most likely does exist in the chicken as (chicken) ESTs similar to the loci contained in the BAC clone are in the Genbank EST database and chicken DNA could be amplified using some of the primer sets used to amplify and sequence portions of the turkey locus. It may be valuable

to screen the chicken large insert libraries to identify, map, and sequence the homologous region.

The genetic diversity of the core MHC-B region has been studied in both commercial and wild turkeys with a higher level of polymorphism identified in the MHC as compared to other chromosomes. Yet, little is known of the genetic diversity of the MHC-Y. The region is thought to contain non-classical MHC loci, and therefore is expected to not be as polymorphic as the B. At least one study has identified a polymorphic and functional Class I locus with a divergent peptide binding region. Ensuing identification, sequencing, and assembly of additional MHC-Y sequences should be conducted to assess the level of polymorphism in this region. Simple PCR resequencing techniques as used in this study could be employed as well as some locus-specific amplifications of the Y-locus MHC antigens to identify the diversity of the region. Further characterization of specific genes (at both the sequence and expression levels) might suggest possible functions. As part of the work described in this thesis, portions of the MHC-Y have been amplified in Nici without identifying a polymorphism. The extent of Nici's homozygosity across the whole chromosome is not known; however, the limited heterozygosity has benefitted sequencing and assembly thus far, and would aid in completing the sequencing across the Y.

Popular culture is currently trending toward free-range, antibiotic free agriculture. Both practices lead to higher infection rates in flocks, leading producers to desire knowledge of genetic loci affecting immune status. Recent work by Abdulrahman and Hafez identified a significant difference in susceptibility among turkey lines to the parasitic protozoan *Histomonas meleagridis* (187). These results should be augmented by surveying the MHC-B haplotypes to identify potential haplotype associations. Further industry-supported research regarding loci influencing infectious outcomes and vaccine efficacy has been discussed. Other roles of the MHC to infectious outcomes of turkey pathogens can readily be studied such as *Salmonella Spp, Eimeria Spp, Pasteurella multocida,* and the avian influenza virus (H5N1) currently of great interest. It will be interesting to see what (if any) role MHC haplotypes have in immune responses to such pathogens.

Reference List


1. Punnett, R. C. 1923. *Heredity in Poultry*. Macmillan and Co,Ltd, London.

2. Piertney, S. B., and M. K. Oliver. 2005. The evolutionary ecology of the major histocompatibility complex. *Heredity* 96: 7-21.

3. Klein, J. 1986. Seeds of time: fifty years ago Peter A. Gorer discovered the H-2 complex. *Immunogenetics* 24: 331-338.

4. Snell, G. D. 1992. The Nobel Lectures in Immunology. Lecture for the Nobel Prize for Physiology or Medicine, 1980: Studies in histocompatibility. *Scand. J Immunol* 36: 513-526.

5. Dausset, J. 1981. The Major Histocompatibility Complex in Man. *Science* 213: 1469-1474.

6. Benacerraf, B. 1981. Role of MHC gene products in immune regulation. *Science* 212: 1229-1238.

7. Doherty, P. C. 1997. The Nobel Lectures in Immunology. The Nobel Prize for Physiology or Medicine, 1996. Cell mediated immunity in virus infections. *Scand. J Immunol* 46: 527-540.

8. Zinkernagel, R. M. 1997. The Nobel Lectures in Immunology. The Nobel Prize for Physiology or Medicine, 1996. Cellular immune recognition and the biological role of major transplantation antigens. *Scand. J Immunol* 46: 421-436.

9. Chien, Y. h., N. R. J. Gascoigne, J. Kavaler, N. E. Lee, and M. M. Davis. 1984. Somatic recombination in a murine T-cell receptor gene. *Nature* 309: 322-326.

10. Tonegawa, S. 1988. Nobel lecture in physiology or medicine--1987. Somatic generation of immune diversity. *In Vitro Cell Dev. Biol* 24: 253-265.

11. Kelley, J., L. Walter, and J. Trowsdale. 2005. Comparative genomics of major histocompatibility complexes. *Immunogenetics* 56: 683-695.

12. Schook, L. B., and S. J. Lamont. 1996. *The major histocompatibility complex region of domestic animal species*. CRC Press, Boca Raton.

13. Horton, R., L. Wilming, V. Rand, R. C. Lovering, E. A. Bruford, V. K. Khodiyar, M. J. Lush, S. Povey, C. C. Talbot, Jr., M. W. Wright, H. M. Wain, J. Trowsdale, A. Ziegler, and S. Beck. 2004. Gene map of the extended human MHC. *Nat Rev Genet.* 5: 889-899.

14. Driscoll, J., M. G. Brown, D. Finley, and J. J. Monaco. 1993. MHC-linked LMP gene products specifically alter peptidase activities of the proteasome. *Nature* 365: 262-264.

15. McCluskey, J., J. Rossjohn, and A. W. Purcell. 2004. TAP genes and immunity. *Current Opinion in Immunology* 16: 651-659.

16. Heath, W. R., and F. R. Carbone. 2001. Cross-presentation in viral immunity and self-tolerance. *Nat Rev Immunol* 1: 126-134.

17. Rodgers, J. R., and R. G. Cook. 2005. MHC class Ib molecules bridge innate and acquired immunity. *Nat Rev Immunol* 5: 459-471.

18. Narayan, K., C. L. Chou, A. Kim, I. Z. Hartman, S. Dalai, S. Khoruzhenko, and S. Sadegh-Nasseri. 2007. HLA-DM targets the hydrogen bond between the histidine at position [beta]81 and peptide to dissociate HLA-DR-peptide complexes. *Nat Immunol* 8: 92-100.

19. Kronenberg, M., G. Siu, L. E. Hood, and N. Shastri. 1986. The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *Annu. Rev Immunol* 4: 529-591.

20. Vyas, J. M., A. G. Van der Veen, and H. L. Ploegh. 2008. The known unknowns of antigen processing and presentation. *Nat Rev Immunol* 8: 607-618.

21. Goulder, P. J. R., and D. I. Watkins. 2008. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat Rev Immunol* 8: 619-630.

22. Jones, E. Y., L. Fugger, J. L. Strominger, and C. Siebold. 2006. MHC class II proteins and disease: a structural perspective. *Nat Rev Immunol* 6: 271-282.

23. Kennedy, L. J., H. J. Huson, J. Leonard, J. M. Angles, L. E. Fox, J. W. Wojciechowski, C. Yuncker, and G. M. Happ. 2006. Association of hypothyroid disease in Doberman Pinscher dogs with a rare major histocompatibility complex DLA class II haplotype. *Tissue Antigens* 67: 53-56.

24. Kennedy, L. J., L. J. Davison, A. Barnes, A. D. Short, N. Fretwell, C. A. Jones, A. C. Lee, W. E. Ollier, and B. Catchpole. 2006. Identification of susceptibility and protective major histocompatibility complex haplotypes in canine diabetes mellitus. *Tissue Antigens* 68: 467-476.

25. Kennedy, L. J., A. Barnes, W. E. Ollier, and M. J. Day. 2006. Association of a common dog leucocyte antigen class II haplotype with canine primary immune-mediated haemolytic anaemia. *Tissue Antigens* 68: 502-508.

26. Kennedy, L. J., S. Quarmby, G. M. Happ, A. Barnes, I. K. Ramsey, R. M. Dixon, B. Catchpole, C. Rusbridge, P. A. Graham, N. S. Hillbertz, C. Roethel, W. J.

Dodds, N. G. Carmichael, and W. E. Ollier. 2006. Association of canine hypothyroidism with a common major histocompatibility complex DLA class II allele. *Tissue Antigens* 68: 82-86.

27. Deeg, C. A., E. Marti, C. Gaillard, and B. Kaspers. 2004. Equine recurrent uveitis is strongly associated with the MHC class I haplotype ELA-A9. *Equine Vet. J* 36: 73-75.

28. Kaufman, J., S. Milne, T. W. Gobel, B. A. Walker, J. P. Jacob, C. Auffray, R. Zoorob, and S. Beck. 1999. The chicken B locus is a minimal essential major histocompatibility complex. *Nature* 401: 923-925.

29. Miller, M. M., L. D. Bacon, K. Hala, H. D. Hunt, S. J. Ewald, J. Kaufman, R. Zoorob, and W. E. Briles. 2004. 2004 Nomenclature for the chicken major histocompatibility (B and Y) complex. *Immunogenetics* 56: 261-279.

30. Shiina, T., W. E. Briles, R. M. Goto, K. Hosomichi, K. Yanagiya, S. Shimizu, H. Inoko, and M. M. Miller. 2007. Extended gene map reveals tripartite motif, C-type lectin, and Ig superfamily type genes within a subregion of the chicken MHC-B affecting infectious disease. *J. Immunol.* 178: 7162-7172.

31. Salomonsen, J., D. Marston, D. Avila, N. Bumstead, B. Johansson, H. Juul-Madsen, G. D. Olesen, P. Riegert, K. Skjodt, O. Vainio, M. V. Wiles, and J. Kaufman. 2003. The properties of the single chicken MHC classical class II alpha chain ( B-LA) gene indicate an ancient origin for the DR/E-like isotype of class II molecules. *Immunogenetics* 55: 605-614.

32. Miller, M. M., R. Goto, S. Young, J. Chirivella, D. Hawke, and C. G. Miyada. 1991. Immunoglobulin variable-region-like domains of diverse sequence within the major histocompatibility complex of the chicken. *Proc. Natl. Acad. Sci. U. S. A* 88: 4377-4381.

33. Ruby, T., B. Bed'Hom, H. Wittzell, V. Morin, A. Oudin, and R. Zoorob. 2005. Characterisation of a cluster of TRIM-B30.2 genes in the chicken MHC B locus. *Immunogenetics* 57: 116-128.

34. Miller, M. M., C. Wang, E. Parisini, R. D. Coletta, R. M. Goto, S. Y. Lee, D. C. Barral, M. Townes, C. Roura-Mir, H. L. Ford, M. B. Brenner, and C. C. Dascher. 2005. Characterization of two avian MHC-like genes reveals an ancient origin of the CD1 family. *Proc. Natl. Acad. Sci. U. S. A* 102: 8674-8679.

35. Salomonsen, J., M. R. Sorensen, D. A. Marston, S. L. Rogers, T. Collen, H. A. van, A. L. Smith, R. K. Beal, K. Skjodt, and J. Kaufman. 2005. Two CD1 genes map to the chicken MHC, indicating that CD1 genes are ancient and likely to have been present in the primordial MHC. *Proc. Natl. Acad. Sci. U. S. A* 102: 8668-8673.

36. Hosomichi, K., M. M. Miller, R. M. Goto, Y. Wang, S. Suzuki, J. K. Kulski, M. Nishibori, H. Inoko, K. Hanzawa, and T. Shiina. 2008. Contribution of Mutation, Recombination, and Gene Conversion to Chicken Mhc-B Haplotype Diversity. *J Immunol* 181: 3393-3399.

37. Shiina, T., S. Shimizu, K. Hosomichi, S. Kohara, S. Watanabe, K. Hanzawa, S. Beck, J. K. Kulski, and H. Inoko. 2004. Comparative genomic analysis of two avian (quail and chicken) MHC regions. *J. Immunol.* 172: 6751-6763.

38. Hosomichi, K., T. Shiina, S. Suzuki, M. Tanaka, S. Shimizu, S. Iwamoto, H. Hara, Y. Yoshida, J. K. Kulski, H. Inoko, and K. Hanzawa. 2006. The major histocompatibility complex (Mhc) class IIB region has greater genomic structural flexibility and diversity in the quail than the chicken. *BMC. Genomics* 7: 322.

39. Briles, W. E., R. M. Goto, C. Auffray, and M. M. Miller. 1993. A polymorphic system related to but genetically independent of the chicken major histocompatibility complex. *Immunogenetics* 37: 408-414.

40. Miller, M. M., R. Goto, A. Bernot, R. Zoorob, C. Auffray, N. Bumstead, and W. E. Briles. 1994. Two Mhc class I and two Mhc class II genes map to the chicken Rfp-Y system outside the B complex. *Proc. Natl. Acad. Sci. U. S. A* 91: 4397-4401.

41. Miller, M. M., R. Goto, R. Zoorob, C. Auffray, and W. E. Briles. 1994. Regions of homology shared by Rfp-Y and major histocompatibility B complex genes. *Immunogenetics* 39: 71-73.

42. Afanassieff, M., R. M. Goto, J. Ha, M. A. Sherman, L. Zhong, C. Auffray, F. Coudert, R. Zoorob, and M. M. Miller. 2001. At least one class I gene in restriction fragment pattern-Y (Rfp-Y), the second MHC gene cluster in the chicken, is transcribed, polymorphic, and shows divergent specialization in antigen binding region. *J. Immunol.* 166: 3324-3333.

43. Rogers, S., I. Shaw, N. Ross, V. Nair, L. Rothwell, J. Kaufman, and P. Kaiser. 2003. Analysis of part of the chicken Rfp-Y region reveals two novel lectin genes, the first complete genomic sequence of a class I alpha-chain gene, a truncated class II beta-chain gene, and a large CR1 repeat. *Immunogenetics* 55: 100-108.

44. Wakenell, P. S., M. M. Miller, R. M. Goto, W. J. Gauderman, and W. E. Briles. 1996. Association between the Rfp-Y haplotype and the incidence of Marek's disease in chickens. *Immunogenetics* 44: 242-245.

45. Vallejo, R. L., G. T. Pharr, H. C. Liu, H. H. Cheng, R. L. Witter, and L. D. Bacon. 1997. Non-association between Rfp-Y major histocompatibility complex-like genes and susceptibility to Marek's disease virus-induced tumours in 6(3) x 7(2) F2 intercross chickens. *Anim Genet.* 28: 331-337.

46. Lakshmanan, N., and S. J. Lamont. 1998. Rfp-Y region polymorphism and Marek's disease resistance in multitrait immunocompetence-selected chicken lines. *Poult Sci* 77: 538-541.

47. Fillon, V., R. Zoorob, M. Yerle, C. Auffray, and A. Vignal. 1996. Mapping of the genetically independent chicken major histocompatibility complexes B@ and RFP-Y@ to the same microchromosome by two-color fluorescent in situ hybridization. *Cytogenet. Cell Genet.* 75: 7-9.

48. Miller, M. M., R. M. Goto, R. L. Taylor, Jr., R. Zoorob, C. Auffray, R. W. Briles, W. E. Briles, and S. E. Bloom. 1996. Assignment of Rfp-Y to the chicken major histocompatibility complex/NOR microchromosome and evidence for high-frequency recombination associated with the nucleolar organizer region. *Proc. Natl. Acad. Sci. U. S. A* 93: 3958-3962.

49. Bloom, S. E., and L. D. Bacon. 1985. Linkage of the major histocompatibility (B) complex and the nucleolar organizer in the chicken. Assignment to a microchromosome. *J. Hered.* 76: 146-154.

50. Kaufman, J., and J. Salomonsen. 1997. The "minimal essential MHC" revisited: both peptide-binding and cell surface expression level of MHC molecules are polymorphisms selected by pathogens in chickens. *Hereditas* 127: 67-73.

51. Kaufman, J. 2000. The simple chicken major histocompatibility complex: life and death in the face of pathogens and vaccines. *Philos. Trans. R. Soc. Lond B Biol. Sci.* 355: 1077-1084.

52. Briles, W. E., H. A. Stone, and R. K. Cole. 1977. Marek's disease: effects of B histocompatibility alloalleles in resistant and susceptible chicken lines. *Science* 195: 193-195.

53. Taylor, R. L., Jr. 2004. Major histocompatibility (B) complex control of responses against Rous sarcomas. *Poult Sci* 83: 638-649.

54. Lamont, S. J., M. G. Kaiser, and W. Liu. 2002. Candidate genes for resistance to Salmonella enteritidis colonization in chickens as detected in a novel genetic cross. *Veterinary Immunology and Immunopathology* 87: 423-428.

55. Liu, W., M. M. Miller, and S. J. Lamont. 2002. Association of MHC class I and class II gene polymorphisms with vaccine or challenge response to Salmonella enteritidis in young chicks. *Immunogenetics* 54: 582-590.

56. Palmer, D. K., and A. W. Nordskog. 1980. The major histocompatibility complex in the turkey: do erythrocytes bear serologically detectable H antigens? *Poult Sci* 59: 2158-2159.

57. Emara, M. G., K. E. Nestor, D. N. Foster, and S. J. Lamont. 1992. The turkey major histocompatibility complex: identification of class II genotypes by

restriction fragment length polymorphism analysis of deoxyribonucleic acid. *Poult Sci* 71: 2083-2089.

58. Emara, M. G., K. E. Nestor, and L. D. Bacon. 1993. The turkey major histocompatibility complex: characterization by mixed lymphocyte, graft-versus-host splenomegaly, and skin graft reactions. *Poult Sci* 72: 60-66.

59. Zhu, J., K. E. Nestor, and S. J. Lamont. 1995. Survey of major histocompatibility complex class II haplotypes in four turkey lines using restriction fragment length polymorphism analysis with nonradioactive DNA detection. *Poult Sci* 74: 1067-1073.

60. Zhu, J., K. E. Nestor, and Y. Tang. 1996. Frequencies and genetic diversity of major histocompatibility complex class II haplotypes in commercial turkey lines. *Poult Sci* 75: 954-958.

61. Sacco, R. E., R. B. Rimler, X. Ye, and K. E. Nestor. 2001. Identification of new major histocompatibility complex class II restriction fragment length polymorphisms in a closed experimental line of Beltsville Small White turkeys. *Poult Sci* 80: 1109-1111.

62. Ahmed, K. A., V. K. Saxena, M. Saxena, A. Ara, A. B. Pramod, M. L. Rajaram, K. S. Dorman, S. Majumdar, and T. J. Rasool. 2007. Molecular cloning and sequencing of MHC class II beta 1 domain of turkey reveals high sequence identity with chicken. *Int. J Immunogenet.* 34: 97-105.

63. Sironi, L., B. Lazzari, P. Ramelli, C. Gorni, and P. Mariani. 2006. Single nucleotide polymorphism discovery in the avian Tapasin gene. *Poult Sci* 85: 606-612.

64. Sironi, L., B. Lazzari, P. Ramelli, A. Stella, and P. Mariani. 2008. Avian TAP genes: detection of nucleotide polymorphisms and comparative analysis across species. *Genet. Mol Res.* 7: 1267-1281.

65. Nestor, K. E., Y. M. Saif, J. Zhu, D. O. Noble, and R. A. Patterson. 1996. The influence of major histocompatibility complex genotypes on resistance to Pasteurella multocida and Newcastle disease virus in turkeys. *Poult Sci* 75: 29-33.

66. Cockett, N. E. 2008. *Genome mapping and genomics in domestic animals*. Springer, New York.

67. Nestor, K. E., Y. M. Saif, J. Zhu, and D. O. Noble. 1996. Influence of growth selection in turkeys on resistance to Pasteurella multocida. *Poult Sci* 75: 1161-1163.

68. Jirjis, F. F., S. L. Noll, D. A. Halvorson, K. V. Nagaraja, and D. P. Shaw. 2002. Pathogenesis of avian pneumovirus infection in turkeys. *Vet. Pathol.* 39: 300-310.

69. Munir, S., and V. Kapur. 2003. Regulation of host cell transcriptional physiology by the avian pneumovirus provides key insights into host-pathogen interactions. *J Virol.* 77: 4899-4910.

70. Munir, S., and V. Kapur. 2003. Transcriptional analysis of the response of poultry species to respiratory pathogens. *Poult Sci* 82: 885-892.

71. Panigrahy, B., D. A. Senne, J. C. Pedersen, T. Gidlewski, and R. K. Edson. 2000. Experimental and serologic observations on avian pneumovirus (APV/turkey/Colorado/97) infection in turkeys. *Avian Dis.* 44: 17-22.

72. Chaves, L. D., T. P. Knutson, S. B. Krueth, and K. M. Reed. 2006. Using the chicken genome sequence in the development and mapping of genetic markers in the turkey (Meleagris gallopavo). *Anim Genet.* 37: 130-138.

73. Chaves, L. D., B. J. Ostroski, and K. M. Reed. 2003. Myosin light chain genes in the turkey (Meleagris gallopavo). *Cytogenet. Genome Res.* 102: 340-346.

74. Griffin, D. K., L. B. Robertson, H. G. Tempest, A. Vignal, V. Fillon, R. P. Crooijmans, M. A. Groenen, S. Deryusheva, E. Gaginskaya, W. Carre, D. Waddington, R. Talbot, M. Volker, J. S. Masabanda, and D. W. Burt. 2008. Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. *Bmc Genomics* 9: 168-184.

75. Reed, K. M., L. D. Chaves, and K. M. Mendoza. 2007. An integrated and comparative genetic map of the turkey genome. *Cytogenet. Genome Res.* 119: 113-126.

76. Romanov, M. N., and J. B. Dodgson. 2006. Cross-species overgo hybridization and comparative physical mapping within avian genomes. *Anim Genet.* 37: 397-399.

77. Kaiser, V. B., M. van Tuinen, and H. Ellegren. 2007. Insertion Events of CR1 Retrotransposable Elements Elucidate the Phylogenetic Branching Order in Galliform Birds. *Mol Biol Evol* 24: 338-347.

78. Schook LB, and Lamont SJ. 1996. ***The major histocompatibility complex region of domestic animal species.*** **CRC Press**, Boca Raton.

79. Ross MT, LaBrie S, McPherson J, and Stanton VP. 1999. Screening large-insert libraries by hybridization. In *Current Protocols in Human Genetics*. Dracopoli NC, Haines JL, Korf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, and Smith DR, eds. John Wiley and Sons, New York. **5.6.1 -5.6.52**.

80. Reed, K. M., L. D. Chaves, J. R. Garbe, Y. Da, and D. E. Harry. 2003. Allelic variation and genetic linkage of avian microsatellites in a new turkey population for genetic mapping. *Cytogenet. Genome Res.* 102: 331-339.

81. Garbe J, and Da Y. 2003. Locusmap user manual Version 1.1. Department of Animal Science, University of Minnesota, St. Paul MN.

82. Reed, K. M., L. R. Sullivan, L. K. Foster, L. D. Chaves, and F. A. Ponce de Leon. 2006. Assignment of linkage groups to turkey chromosome 1 (MGA1). *Cytogenet. Genome Res.* 115: 176-178.

83. Howell, W. M., and D. A. Black. 1980. Controlled silver-staining of nucleolus organizer regions with a protective colloidal developer: a 1-step method. *Experientia* 36: 1014-1015.

84. Delany, M. E., and A. B. Krupkin. 1999. Molecular characterization of ribosomal gene variation within and among NORs segregating in specialized populations of chicken. *Genome* 42: 60-71.

85. Delany ME, Daniels LM, Gessaro TM, and Rodrigue KM. 2006. A cytogenomic approach to understanding the organization and function of the ultra-long (aka mega-) telomere arrays of the chicken genome.

86. Axelsson, E., M. T. Webster, N. G. C. Smith, D. W. Burt, and H. Ellegren. 2005. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. *Genome Res.* 15: 120-125.

87. Reed, K. M., L. D. Chaves, M. K. Hall, T. P. Knutson, and D. E. Harry. 2005. A comparative genetic map of the turkey genome. *Cytogenet. Genome Res.* 111: 118-127.

88. Ashley, T., and D. C. Ward. 1993. A "hot spot" of recombination coincides with an interstitial telomeric sequence in the Armenian hamster. *Cytogenet. Cell Genet.* 62: 169-171.

89. LePage, K. T., M. M. Miller, W. E. Briles, and R. L. Taylor, Jr. 2000. Rfp-Y genotype affects the fate of Rous sarcomas in B2B5 chickens. *Immunogenetics* 51: 751-754.

90. Rogers, S., I. Shaw, N. Ross, V. Nair, L. Rothwell, J. Kaufman, and P. Kaiser. 2003. Analysis of part of the chicken Rfp-Y region reveals two novel lectin genes, the first complete genomic sequence of a class I alpha-chain gene, a truncated class II beta-chain gene, and a large CR1 repeat. *Immunogenetics* 55: 100-108.

91. Jacob, J. P., S. Milne, S. Beck, and J. Kaufman. 2000. The major and a minor class II beta-chain (B-LB ) gene flank the Tapasin gene in the B-F /B-L region of the chicken major histocompatibility complex. *Immunogenetics* 51: 138-147.

92. Shaw, I., T. J. Powell, D. A. Marston, K. Baker, H. A. van, P. Riegert, M. V. Wiles, S. Milne, S. Beck, and J. Kaufman. 2007. Different evolutionary histories

of the two classical class I genes BF1 and BF2 illustrate drift and selection within the stable MHC haplotypes of chickens. *J. Immunol.* 178: 5744-5752.

93. Rogers, S. L., T. W. Gobel, B. C. Viertlboeck, S. Milne, S. Beck, and J. Kaufman. 2005. Characterization of the chicken C-type lectin-like receptors B-NK and B-lec suggests that the NK complex and the MHC share a common ancestral region. *J. Immunol.* 174: 3475-3483.

94. Liu, W., M. M. Miller, and S. J. Lamont. 2002. Association of MHC class I and class II gene polymorphisms with vaccine or challenge response to Salmonella enteritidis in young chicks. *Immunogenetics* 54: 582-590.

95. Kim, D. K., H. S. Lillehoj, Y. H. Hong, D. W. Park, S. J. Lamont, J. Y. Han, and E. P. Lillehoj. 2008. Immune-Related Gene Expression in Two B-Complex Disparate Genetically Inbred Fayoumi Chicken Lines Following Eimeria maxima Infection. *Poult Sci* 87: 433-443.

96. Chaves, L. D., S. B. Krueth, and K. M. Reed. 2007. Characterization of the turkey MHC chromosome through genetic and physical mapping. *Cytogenet. Genome Res.* 117: 213-220.

97. Pereira, S. L., and A. J. Baker. 2006. A molecular timescale for galliform birds accounting for uncertainty in time estimates and heterogeneity of rates of DNA substitutions across lineages and sites. *Mol. Phylogenet. Evol.* 38: 499-509.

98. Yip, S. S., and R. A. Coulombe, Jr. 2006. Molecular cloning and expression of a novel cytochrome p450 from turkey liver with aflatoxin b1 oxidizing activity. *Chem. Res. Toxicol.* 19: 30-37.

99. Dyrlov Bendtsen, J., H. Nielsen, G. von Heijne, and S. Brunak. 2004. Improved Prediction of Signal Peptides: SignalP 3.0. *Journal of Molecular Biology* 340: 783-795.

100. Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27: 573-580.

101. Lowe, T. M., and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucl. Acids Res.* 25: 955-964.

102. Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418-426.

103. Korber B. 2000. HIV Signature and Sequence Variation Analysis. Computational Analysis of HIV Molecular Sequences. Allen G.Rodrigo and Gerald H.Learn, eds. Kluwer Academic Publishers, Dordrecht, Netherlands. 55-72.

104. Yang, Z., and W. J. Swanson. 2002. Codon-Substitution Models to Detect Adaptive Evolution that Account for Heterogeneous Selective Pressures Among Site Classes. *Mol Biol Evol* 19: 49-57.

105. Kelley, J., and J. Trowsdale. 2005. Features of MHC and NK gene clusters. *Transpl. Immunol.* 14: 129-134.

106. Kroemer, G., R. Zoorob, and C. Auffray. 1990. Structure and expression of a chicken MHC class I gene. *Immunogenetics* 31: 405-409.

107. van den Elsen, P. J., T. M. Holling, H. F. Kuipers, and N. van der Stoep. 2004. Transcriptional regulation of antigen presentation. *Current Opinion in Immunology* 16: 67-75.

108. Shimizu, S., T. Shiina, K. Hosomichi, S. Takahashi, T. Koyama, T. Onodera, J. K. Kulski, and H. Inoko. 2004. MHC class IIB gene sequences and expression in quails (Coturnix japonica) selected for high and low antibody responses. *Immunogenetics* 56: 280-291.

109. Reith, W., S. LeibundGut-Landmann, and J. M. Waldburger. 2005. Regulation of MHC class II gene expression by the class II transactivator. *Nat. Rev. Immunol.* 5: 793-806.

110. Chen, Y., S. L. Carpenter, and S. J. Lamont. 2000. A functional role for the Y box in regulating an MHC class II B gene promoter in chicken lymphocytes. *Immunogenetics* 51: 882-886.

111. Bonneaud, C., G. Sorci, V. Morin, H. Westerdahl, R. Zoorob, and H. Wittzell. 2004. Diversity of Mhc class I and IIB genes in house sparrows (Passer domesticus). *Immunogenetics* 55: 855-865.

112. Westerdahl, H., H. Wittzell, and S. T. von. 1999. Polymorphism and transcription of Mhc class I genes in a passerine bird, the great reed warbler. *Immunogenetics* 49: 158-170.

113. Westerdahl, H., H. Wittzell, and S. T. von. 2000. Mhc diversity in two passerine birds: no evidence for a minimal essential Mhc. *Immunogenetics* 52: 92-100.

114. Westerdahl, H., H. Wittzell, S. T. von, and S. Bensch. 2004. MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. *Heredity* 92: 534-542.

115. Wittzell, H., S. T. von, R. Zoorob, and C. Auffray. 1994. Molecular characterization of three Mhc class II B haplotypes in the ring-necked pheasant. *Immunogenetics* 39: 395-403.

116. Strand, T., H. Westerdahl, J. Hoglund, V. Alatalo, and H. Siitari. 2007. The Mhc class II of the Black grouse (Tetrao tetrix) consists of low numbers of B and Y genes with variable diversity and expression. *Immunogenetics* 59: 725-734.

117. Strunnikov, A. V., and R. Jessberger. 1999. Structural maintenance of chromosomes (SMC) proteins: conserved molecular properties for multiple biological functions. *Eur. J Biochem.* 263: 6-13.

118. James, L. C., A. H. Keeble, Z. Khan, D. A. Rhodes, and J. Trowsdale. 2007. Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function. *Proceedings of the National Academy of Sciences* 104: 6200-6205.

119. Short, K. M., and T. C. Cox. 2006. Subclassification of the RBCC/TRIM Superfamily Reveals a Novel Motif Necessary for Microtubule Binding. *J. Biol. Chem.* 281: 8970-8980.

120. DiBenedetto, A., J. Guinto, T. Ebert, K. Bee, M. Schmidt, and T. Jackman. 2008. Zebrafish brd2a and brd2b are paralogous members of the bromodomain-ET (BET) family of transcriptional coregulators that show structural and expression divergence. *BMC Developmental Biology* 8: 39.

121. LeRoy, G., B. Rickards, and S. J. Flint. 2008. The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol Cell* 30: 51-60.

122. Briles, W. E., R. W. Briles, R. E. Taffs, and H. A. Stone. 1983. Resistance to a malignant lymphoma in chickens is mapped to subregion of major histocompatibility (B) complex. *Science* 219: 977-979.

123. Viertlboeck, B. C., A. Wortmann, R. Schmitt, J. Plachy, and T. W. Gobel. 2008. Chicken C-type lectin-like receptor B-NK, expressed on NK and T cell subsets, binds to a ligand on activated splenocytes. *Mol. Immunol.* 45: 1398-1404.

124. Rogers, S. L., and J. Kaufman. 2008. High allelic polymorphism, moderate sequence diversity and diversifying selection for B-NK but not B-lec, the pair of lectin-like receptor genes in the chicken MHC. *Immunogenetics* 60: 461-475.

125. Moon, D. A., S. M. Veniamin, J. A. Parks-Dely, and K. E. Magor. 2005. The MHC of the duck (Anas platyrhynchos) contains five differentially expressed class I genes. *J Immunol* 175: 6702-6712.

126. Briles, W. E., R. W. Briles, R. E. Taffs, and H. A. Stone. 1983. Resistance to a malignant lymphoma in chickens is mapped to subregion of major histocompatibility (B) complex. *Science* 219: 977-979.

127. Koch, M., S. Camp, T. Collen, D. Avila, J. Salomonsen, H. J. Wallny, H. A. van, L. Hunt, J. P. Jacob, F. Johnston, D. A. Marston, I. Shaw, P. R. Dunbar, V. Cerundolo, E. Y. Jones, and J. Kaufman. 2007. Structures of an MHC class I

molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity.* 27: 885-899.

128. Zoorob, R., G. Behar, G. Kroemer, and C. Auffray. 1990. Organization of a functional chicken class II B gene. *Immunogenetics* 31: 179-187.

129. Thoraval, P., M. Afanassieff, D. Bouret, G. Luneau, E. Esnault, R. M. Goto, A. M. Chausse, R. Zoorob, D. Soubieux, M. M. Miller, and G. Dambrine. 2003. Role of nonclassical class I genes of the chicken major histocompatibility complex Rfp-Y locus in transplantation immunity. *Immunogenetics* 55: 647-651.

130. Hunt, H. D., R. M. Goto, D. N. Foster, L. D. Bacon, and M. M. Miller. 2006. At least one YMHCI molecule in the chicken is alloimmunogenic and dynamically expressed on spleen cells during development. *Immunogenetics* 58: 297-307.

131. Lehner, B., J. I. Semple, S. E. Brown, D. Counsell, R. D. Campbell, and C. M. Sanderson. 2004. Analysis of a high-throughput yeast two-hybrid system and its use to predict the function of intracellular proteins encoded within the human MHC class III region. *Genomics* 83: 153-167.

132. Shen, H., X. Zheng, J. Shen, L. Zhang, R. Zhao, and M. R. Green. 2008. Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. *Genes & Development* 22: 1796-1803.

133. Kota, K. P., S. R. Wagner, E. Huerta, J. M. Underwood, and J. A. Nickerson. 2008. Binding of ATP to UAP56 is necessary for mRNA export. *J Cell Sci* 121: 1526-1537.

134. Sasaki, T., E. C. Gan, A. Wakeham, S. Kornbluth, T. W. Mak, and H. Okada. 2007. HLA-B-associated transcript 3 (Bat3)/Scythe is essential for p300-mediated acetylation of p53. *Genes & Development* 21: 848-861.

135. Rogers, S. L., B. C. Viertlboeck, T. W. Gobel, and J. Kaufman. 2008. Avian NK activities, cells and receptors. *Semin. Immunol* 20: 353-360.

136. Sun, X., H. D. Le, J. M. Wahlstrom, and G. H. Karpen. 2003. Sequence analysis of a functional Drosophila centromere. *Genome Res.* 13: 182-194.

137. Carone, D., M. Longo, G. Ferreri, L. Hall, M. Harris, N. Shook, K. Bulazel, B. Carone, C. Obergfell, M. O'Neill, and R. O'Neill. 2008. A new class of retroviral and satellite encoded small RNAs emanates from mammalian centromeres. *Chromosoma*.

138. Childers, C. P., H. L. Newkirk, D. A. Honeycutt, N. Ramlachan, D. M. Muzney, E. Sodergren, R. A. Gibbs, G. M. Weinstock, J. E. Womack, and L. C. Skow. 2006. Comparative analysis of the bovine MHC class IIb sequence identifies inversion breakpoints and three unexpected genes. *Anim Genet.* 37: 121-129.

139. Lunney, J. K., C. S. Ho, M. Wysocki, and D. M. Smith. 2009. Molecular genetics of the swine major histocompatibility complex, the SLA complex. *Developmental & Comparative Immunology* 33: 362-374.

140. Sambrook, J. G., F. Figueroa, and S. Beck. 2005. A genome-wide survey of Major Histocompatibility Complex (MHC) genes and their paralogues in zebrafish. *Bmc Genomics* 6: 152.

141. Sambrook, J. G., R. Russell, Y. Umrania, Y. J. Edwards, R. D. Campbell, G. Elgar, and M. S. Clark. 2002. Fugu orthologues of human major histocompatibility complex genes: a genome survey. *Immunogenetics* 54: 367-380.

142. Ohta, Y., W. Goetz, M. Z. Hossain, M. Nonaka, and M. F. Flajnik. 2006. Ancestral Organization of the MHC Revealed in the Amphibian Xenopus. *J Immunol* 176: 3674-3685.

143. Belov, K., J. E. Deakin, A. T. Papenfuss, M. L. Baker, S. D. Melman, H. V. Siddle, N. Gouin, D. L. Goode, T. J. Sargeant, M. D. Robinson, M. J. Wakefield, S. Mahony, J. G. Cross, P. V. Benos, P. B. Samollow, T. P. Speed, J. A. Graves, and R. D. Miller. 2006. Reconstructing an ancestral mammalian immune supercomplex from a marsupial major histocompatibility complex. *PLoS. Biol* 4: e46.

144. Trowsdale, J. 2001. Genetic and functional relationships between MHC and NK receptor genes. *Immunity.* 15: 363-374.

145. Dijkstra, J., T. Katagiri, K. Hosomichi, K. Yanagiya, H. Inoko, M. Ototake, T. Aoki, K. Hashimoto, and T. Shiina. 2007. A third broad lineage of major histocompatibility complex (MHC) class I in teleost fish; MHC class II linkage and processed genes. *Immunogenetics* 59: 305-321.

146. Dohm, J., E. Tsend-Ayush, R. Reinhardt, F. Grutzner, and H. Himmelbauer. 2007. Disruption and pseudoautosomal localization of the major histocompatibility complex in monotremes. *Genome Biology* 8: R175.

147. Kaufman, J., J. Jacob, I. Shaw, B. Walker, S. Milne, S. Beck, and J. Salomonsen. 1999. Gene organisation determines evolution of function in the chicken MHC. *Immunol. Rev.* 167: 101-117.

148. Marsh, S. G., E. D. Albert, W. F. Bodmer, R. E. Bontrop, B. Dupont, H. A. Erlich, D. E. Geraghty, J. A. Hansen, C. K. Hurley, B. Mach, W. R. Mayr, P. Parham, E. W. Petersdorf, T. Sasazuki, G. M. Schreuder, J. L. Strominger, A. Svejgaard, P. I. Terasaki, and J. Trowsdale. 2005. Nomenclature for factors of the HLA system, 2004. *Int. J Immunogenet.* 32: 107-159.

149. Harry, D. E., D. Zaitlin, P. J. Marini, and K. M. Reed. 2003. A first-generation map of the turkey genome. *Genome* 46: 914-924.

150. Stephens, M., and P. Donnelly. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am. J Hum. Genet.* 73: 1162-1169.

151. Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.

152. Chaves, L. D., J. A. Rowe, and K. M. Reed. 2005. Survey of a cDNA library from the turkey (Meleagris gallopavo). *Genome* 48: 12-17.

153. Green P, Falls K, and Crooks S. 1990. Documentation for CRIMAP, version 2.4. Washington University School of Medicine, St. Louis.

154. Reed, K. M., M. K. Hall, L. D. Chaves, and T. P. Knutson. 2006. Single nucleotide polymorphisms for integrative mapping in the Turkey (Meleagris gallopavo). *Anim Biotechnol.* 17: 73-80.

155. Horton, R., R. Gibson, P. Coggill, M. Miretti, R. J. Allcock, J. Almeida, S. Forbes, J. G. Gilbert, K. Halls, J. L. Harrow, E. Hart, K. Howe, D. K. Jackson, S. Palmer, A. N. Roberts, S. Sims, C. A. Stewart, J. A. Traherne, S. Trevanion, L. Wilming, J. Rogers, P. J. de Jong, J. F. Elliott, S. Sawcer, J. A. Todd, J. Trowsdale, and S. Beck. 2008. Variation analysis and gene annotation of eight MHC haplotypes: the MHC Haplotype Project. *Immunogenetics* 60: 1-18.

156. Takagi, Y., T. Kuramoto, B. Voigt, T. Tsurumi, S. Nakanishi, T. Mashimo, N. Masui, and T. Serikawa. 2008. An informative set of SSLP markers and genomic profiles in the rat MHC, the RT1 complex. *Immunogenetics*.

157. Fulton, J., H. Juul-Madsen, C. Ashwell, A. McCarron, J. Arthur, N. O'Sullivan, and R. Taylor. 2006. Molecular genotype identification of the Gallus gallus major histocompatibility complex. *Immunogenetics* 58: 407-421.

158. Jacob, S., M. K. McClintock, B. Zelano, and C. Ober. 2002. Paternally inherited HLA alleles are associated with women's choice of male odor. *Nature Genetics* 30: 175.

159. Havlicek, J., and S. C. Roberts. MHC-correlated mate choice in humans: A review. *Psychoneuroendocrinology* In Press, Corrected Proof.

160. Zelano, B., and S. Edwards. 2002. An Mhc Component to Kin Recognition and Mate Choice in Birds: Predictions, Progress, and Prospects. *The American Naturalist* 160: S225-S237.

161. Steiger, S. S., A. E. Fidler, M. Valcu, and B. Kempenaers. 2008. Avian olfactory receptor gene repertoires: evidence for a well-developed sense of smell in birds? *Proc. Biol Sci* 275: 2309-2317.

162. Muir, W. M., G. K.-S. Wong, Y. Zhang, J. Wang, M. A. M. Groenen, R. P. M. A. Crooijmans, H. J. Megens, H. Zhang, R. Okimoto, A. Vereijken, A. Jungerius, G. A. A. Albers, C. T. Lawley, M. E. Delany, S. MacEachern, and H. H. Cheng. 2008. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proceedings of the National Academy of Sciences* 105: 17312-17317.

163. Richardson, D. S., and H. Westerdahl. 2003. MHC diversity in two Acrocephalus species: the outbred Great reed warbler and the inbred Seychelles warbler. *Mol Ecol.* 12: 3523-3529.

164. Freeman-Gallant, C. R., E. M. Johnson, F. Saponara, and M. Stanger. 2002. Variation at the major histocompatibility complex in Savannah sparrows. *Mol Ecol.* 11: 1125-1130.

165. Alcaide, M., S. V. Edwards, J. J. Negro, D. Serrano, and J. L. Tella. 2008. Extensive polymorphism and geographical variation at a positively selected MHC class II B gene of the lesser kestrel (Falco naumanni). *Mol Ecol.* 17: 2652-2665.

166. Dickson, J. G., National Wild Turkey Federation (U.S.), United States, and Forest Service. 1992. *The wild turkey: biology and management*. Stackpole Books, Harrisburg, PA.

167. Mock, K. E., T. C. Theimer, O. E. Rhodes, Jr., D. L. Greenberg, and P. Keim. 2002. Genetic variation across the historical range of the wild turkey (Meleagris gallopavo). *Mol Ecol.* 11: 643-657.

168. Latch, E. K., E. J. Smith, and O. E. Rhodes. 2002. Isolation and characterization of microsatellite loci in wild and domestic turkeys (Meleagris gallopavo). *Molecular Ecology Notes* 2: 176-178.

169. Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin: A software for population genetics data analysis. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

170. Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263-265.

171. Jones, D., C. C. Locklear, R. Buchholz, S. J. Hecht, and A. M. Findley. 2003. Production of an MHC class II B molecular probe in the turkey, Meleagris gallopavo. *J Appl. Genet* 44: 369-373.

172. Lenz, T. L., and S. Becker. 2008. Simple approach to reduce PCR artefact formation leads to reliable genotyping of MHC and other highly polymorphic loci -- Implications for evolutionary analysis. *Gene* 427: 117-123.

173. Edwards, S. V., J. Gasper, and M. March. 1998. Genomics and polymorphism of Agph-DAB1, an Mhc class II B gene in red-winged blackbirds (Agelaius phoeniceus). *Mol Biol Evol* 15: 236-250.

174. Bollmer, J. L., F. H. Vargas, and P. G. Parker. 2007. Low MHC variation in the endangered Galapagos penguin (Spheniscus mendiculus). *Immunogenetics* 59: 593-602.

175. Aguilar, A., S. V. Edwards, T. B. Smith, and R. K. Wayne. 2006. Patterns of variation in MHC class II beta loci of the little greenbul (Andropadus virens) with comments on MHC evolution in birds. *J Hered* 97: 133-142.

176. Buchholz R., Jones Dukes M.D., Hecht S., and Findley A.M. 2004. Investigating the turkey's 'snood' as a morphological marker of heritable disease resistance. *Journal of Animal Breeding and Genetics* 121: 176-185.

177. Andreescu, C., S. Avendano, S. R. Brown, A. Hassen, S. J. Lamont, and J. C. M. Dekkers. 2007. Linkage Disequilibrium in Related Breeding Lines of Chickens. *Genetics* 177: 2161-2169.

178. Rao, Y. S., Y. Liang, M. N. Xia, X. Shen, Y. J. Du, C. G. Luo, Q. H. Nie, H. Zeng, and X. Q. Zhang. 2008. Extent of linkage disequilibrium in wild and domestic chicken populations. *Hereditas* 145: 251-257.

179. Lindblad-Toh, K., C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe, M. Kamal, M. Clamp, J. L. Chang, E. J. Kulbokas, M. C. Zody, E. Mauceli, X. Xie, M. Breen, R. K. Wayne, E. A. Ostrander, C. P. Ponting, F. Galibert, D. R. Smith, P. J. deJong, E. Kirkness, P. Alvarez, T. Biagi, W. Brockman, J. Butler, C. W. Chin, A. Cook, J. Cuff, M. J. Daly, D. DeCaprio, S. Gnerre, M. Grabherr, M. Kellis, M. Kleber, C. Bardeleben, L. Goodstadt, A. Heger, C. Hitte, L. Kim, K. P. Koepfli, H. G. Parker, J. P. Pollinger, S. M. J. Searle, N. B. Sutter, R. Thomas, C. Webber, and E. S. Lander. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438: 803-819.

180. International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432: 695-716.

181. Chowdhary, B., and T. Raudsepp. 2008. The Horse Genome Derby: racing from map to whole genome sequence. *Chromosome Research* 16: 109-127.

182. Reed, K. M., L. D. Chaves, M. K. Hall, T. P. Knutson, J. A. Rowe, and A. J. Torgerson. 2003. Microsatellite loci for genetic mapping in the turkey (Meleagris gallopavo). *Anim Biotechnol.* 14: 119-131.

183. DeWoody, Y. D., and J. A. DeWoody. 2005. On the Estimation of Genome-wide Heterozygosity Using Molecular Markers. *J Hered* 96: 85-88.

184. Burt, D. W., D. R. Morrice, A. Sewalem, J. Smith, I. R. Paton, E. J. Smith, J. Bentley, and P. M. Hocking. 2003. Preliminary linkage map of the turkey (Meleagris gallopavo) based on microsatellite markers. *Anim Genet* 34: 399-409.

185. Reed, K. M., T. P. Knutson, S. B. Krueth, L. R. Sullivan, and L. D. Chaves. 2005. In silco mapping of ESTs from the turkey (Meleagris gallopavo). *Anim Biotechnol.* 16: 81-102.

186. Dranchak, P. K., L. D. Chaves, J. A. Rowe, and K. M. Reed. 2003. Turkey microsatellite loci from an embryonic cDNA library. *Poult. Sci.* 82: 526-531.

187. Abdulrahman, L., and H. M. Hafez. 2009. Susceptibility of different turkey lines to Histomonas meleagridis after experimental infection. *Parasitol. Res.*

**Appendix**

Dear Dr. Chaves

I would like to refer to your request mentioned below, which has been forwarded to our department (Rights and Permissions) by Mr. M. Schmid. As to it, permission herewith is granted to use the article

Chaves, L.D. et al: Cytogenet Genome Res 117;213:220

to be republished in your PhD thesis provided that full credit is given to the original source and that S. Karger AG, Basel is mentioned.

Please note that this permission is valid for one edition only. Any further use or distribution requires written permission again and may be subject to a permission fee.

Hopefully, I have been of assistance to you with the above.

Best regards,

Silvia Meier
Rights and Permissions

S. Karger AG
Medical and Scientific Publishers
Allschwilerstrasse 10
CH - 4009 Basel
Tel +41 61 306 14 75
Fax +41 61 306 12 34
E-Mail permission@karger.ch