Krüppel-Like Factors, Chromatin, And Epigenetics: New Frontiers In Understanding The Regulation Of Complex Diseases

A Thesis
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Adrienne Lucille Grzenda

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Master of Science

Claudia Neuhauser, Ph.D

May 2014

ACKNOWLEDGEMENTS

I wish to thank the numerous individuals who have made this thesis possible. To my Ph.D. mentor, Dr. Raul Urrutia, who first enlightened me as to the enormous potential for a new type of hybrid scientist, one steeped in systems biology, who couples the power of bioinformatics and computational modeling of high throughput data to basic bench science to answer complex hypotheses. To my M.S. mentor, Dr. Claudia Neuhauser, who warmly welcomed me into the BICB program and structured an academic program that permitted rigorous, flexible training to accelerate my current and future research. To the University of Minnesota-Rochester, for its vision and excitement in partnering with Mayo Clinic, making this sort of unique academic opportunity feasible. To Drs. Grazia Isaya and Scott Kaufmann, as well as the entire Medical Scientist Training Program, who supported this venture as a critical complement to my M.D./Ph.D. training. To all of the members of the Urrutia lab, past and present, with whom working has been an absolute pleasure. In particular, I wish to thank Dr. Ezequiel Calvo of Centre Hospitalier de l'Universite Laval (CHUL) Research Center, an extraordinarily patient teacher who never failed to patiently and thoroughly answer my endless questions.

DEDICATION

This thesis is dedicated to Raul Urrutia, M.D., my mentor and friend, who provided a home to explore, challenge, experiment, and even fail with unfailing support and encouragement. Quand on a le droit de se tromper impunément, on est toujours sûr de réussir.

ABSTRACT

The discovery of Krüppel-Like Factor (KLF) proteins has contributed more than one could have imagined regarding the mechanisms that underscore complex physiological and pathological phenotypes. From the start, our laboratory hypothesized that elucidating the intricate network of interactions between KLF transcription factors and epigenetic machinery would provide significant insight into the mechanisms of human diseases. Drawing from nearly two decades of data and utilizing a novel system biology approach, we propose the novel hypothesis that KLF proteins function as master transcription factors to deliver epigenetic information to an orchestra of gene promoters, influencing chromatin dynamics and global patterns of gene expression. Multidimensional comparison and visualization of robust experimental datasets using a paradigm transcription factor, KLF11, illustrate the complicated, yet delightful nature of the relationship between KLF proteins and their chromatin cofactors, providing exciting revelations into the causes, prognosis, and therapeutic management of complex diseases.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 2:

CHAPTER 3:

LIST OF FIGURES

.

CHAPTER 1:

Introduction

Uncovering the environmental and genetics factors underlying the etiology of complex, non-Mendelian human diseases remains an elusive goal of biomedical research. Despite the herculean efforts of genome wide association studies (GWASs) and high throughput genomic technologies, alterations in DNA sequence alone have failed to provide complete causality for the highly variable presentation and therapeutic responsiveness of common complex pathologies. Sequencing and re-sequencing retain their promise of unlocking further critical associations between genetic variation and disease. However, non-genetic determinants—particularly epigenetic factors—have garnered increasing attention as probable culprits in the origination, susceptibility, or progression of many complex diseases.

Epigenetics refers to the regulation of gene expression mediated by heritable, yet theoretically reversible, changes in chromatin structure that influence the accessibility of transcription factors to the DNA template. Chromatin is comprised of DNA and proteins that permit the compaction of duplex DNA into the fixed volume of the nucleus. The first layer of chromatin organization occurs at the base unit, the nucleosome, where approximately 146 base pairs of DNA are wrapped around an octamer of basic histone proteins (H2A, H2B, H3, and H4) and stabilized by a single molecule of H1. Each primary array of nucleosomes, termed the "beads on a string" conformation, is further folded and compacted into tight fibers of arrays, eventually condensing into individual chromosomes. Euchromatin denotes the open, transcriptionally accessible state of the chromatin, while heterochromatin refers to the closed, transcriptionally inaccessible state. Within the nucleoplasm, chromatin is further subdivided into membrane-free compartments with heterochromatin distributed along the nuclear periphery, proximal to the nuclear lamina, and euchromatin located more centrally. Other, smaller compartments exist that are organized toward a common function, e.g. rRNA transcription in the nucleolus. Reconfiguration of chromatin state may occur through a number of mechanisms, including (i) methylation of DNA at cytosine/guanine (CpG) dinucleotides via DNA methyltransferases, (ii) the action of ATP-dependent chromatin remodeling complexes to alter or remove nucleosomes, (iii) posttranslational modifications

(acetylation, methylation, phosphorylation, etc.) of residues on the amino terminal tails of histone proteins by histone modifying enzymes (histone code hypothesis), and (iv) the incorporation of variant histones into nucleosomes. All modifications to chromatin are potentially reversible, a fact that holds clear therapeutic potential to permit primary treatment options for diseases currently only treated at the level of symptom management (1).

Epigenetic variation: "New" culprit in complex disease?

Epigenetics mechanisms are highly conserved among all eukaryotic species and knockout of epigenetic regulators is frequently embryonic lethal, demonstrating that both the fidelity of DNA sequence and its epigenetic status are critically important to the proper function of an organism's genetic program. The number of currently identified epigenetic molecules, modifications, and mechanisms is enormous, illustrating the dynamic flexibility inherent in the system to respond rapidly to internal programs or environmental stresses. Organism development, where many epigenetic molecules and processes were first characterized, provides dramatic examples of the profound consequences of epigenetic modifications on gene expression networks. The monarch butterfly possesses the same genetic information throughout its metamorphosis yet epigenetic regulation of gene expression manifests in two drastically different phenotypes, caterpillar and butterfly. Furthermore, while the lifespan of a monarch is typically only 4-5 weeks, a unique subset of monarchs is born each year, the "Methuselah generation," that survive for up to 7-8 months and undergo a roundtrip migration from Canada/United States to Mexico. Differentially expressed microRNAs were uncovered between the summer and migratory populations to partially explain this phenomena (2). While genetic in origin, regulation in the expression and timing of these microRNAs is mostly likely epigenetic. Thus, a single genome may possess effectively infinite epigenetic states and corresponding patterns of gene expression.

Epigenetic variations arise as either a cause or consequence of disease. The need to determine pre-existing versus acquired epigenetic status as well as the advent of high

throughput epigenomic screens (e.g. whole genome DNA methylation assays) has fueled the development of epigenome-wide association studies (EWASs) (3). Pre-existing epigenetic variations may be transgenerational, inherited from patterns established in the parental gametes and disseminated to all or selective tissues in the subsequent generations (4). Other variations may arise stochastically (5). The penetrance of the variation is then dictated by the chronological timing of the imprinting event with early developmental events possessing a stronger likelihood of dispersed effects. Finally, variation may be directly linked to environmental exposure (e.g. smoking, alcohol) or lifestyle choices (e.g. diet) (6). The lifestyle choices of an ancestral population may result in the generation of heritable epialleles disseminated to their offspring. Inherited epigenetic variation, or "soft inheritance," is a prime candidate for explaining the observed phenotypic variability of complex diseases insufficiently accounted for by "hard inheritance," i.e. genetic and environmental variation (7).

The overall accessibility of the DNA template to transcriptional machinery is highly dependent on the chromatin landscape. Until recently, the relationship between chromatin and transcription factors appeared largely passive, with chromatin confirmation, histone modifications, and DNA methylation governing the affinity between *trans*-acting factors and their binding sites on gene promoters. However, the evidence that overexpression of four transcription factors—Oct4, Sox2, m-Myc, and KLF4—is sufficient to reprogram fully differentiated fibroblasts into pluripotent ES-like cells suggests an active role for transcription factors in the epigenetic regulation of chromatin state (8). KLF proteins, in particular, have demonstrated a multitude of feed-forward effects on epigenetic regulation of gene expression. Here we propose KLF proteins as a model for understanding the critical role of sequence-specific factors in epigenetic regulation of a number of complex diseases, including diabetes, obesity, and cancer.

Discovery of Sp1 and Sp1-like sequence-specific DNA-binding transcription factors

Prior to 1980, the manner in which genes responded to specific stimuli to execute programs of protein expression was unclear. Between 1983-1987, Robert Tijan and

colleagues at the University of California purified and cloned the first mammalian transcription factor, named Specificity Protein 1 (Sp1), and demonstrated that Sp1 could activate the simian virus 40 (SV40) promoter by binding GC-rich elements via three contiguous C_2H_2 zinc finger motifs in its C-terminus (9,10). Subsequently, ~2% of the human genome has been found encode C_2H_2 zinc finger proteins, making it the second largest subset of transcription factors after the odorant receptor family (11). Early homology analysis pointed to the *D. melanogaster* "gap" segmentation gene Krüppel as a potential ancestral gene given the high degree of conservation of both the zinc finger region as well as a seven-amino acid spacer region between the fingers, the H/C link (12).

The discovery of Sp1 and momentum of the cloning era fueled a decades-long hunt for structurally and functionally related mammalian transcription factors. In 1992, Fujii-Kuriyama cloned two regulatory proteins bound to a GC-rich basic transcription element (BTE) in the rat—Sp1 and basic transcription element B1 (BTEB1/KLF9), a novel C_2H_2 zinc finger protein with high similarity to Sp1. Interestingly, the two proteins displayed antagonistic effects on the *CYP1A1* promoter (13). Identification of Sp2-4 followed (14,15) as well as Bieker reporting the cloning and characterization of human erythroid Krüppel-like factor (EKLF/KLF1), a factor capable of activating the β-globin promoter through an Sp1-like target site (16). During the same period, our laboratory cloned and contributed two transforming growth factor-β (TGF-β)-inducible Sp1-like proteins, TIEG1 (KLF10) and TEIG2 (KLF11) (17,18). The "Krüppel-like factor (KLF)" designation ultimately triumphed in the nomenclature. In total, 17 different human proteins with similarity to Sp1/Krüppel within their zinc finger region have been identified that bind to GC-rich regions to regulate gene expression, forming the KLF family of proteins (19). Unlike the ubiquitously expressed Sp1, tissue expression patterns of individual KLFs vary widely and often hold organ-specific functions in development and function. Several excellent reviews have been published that detail the complete history and biochemical characterization of individual factors (19-21).

Characteristic structural features of KLF family proteins

Site-sequence transcription factors must possess three defining characteristics: (i) a DNA-binding domain, (ii) a nuclear localization signal (NLS), and (iii) transcriptional regulatory domain. The C-terminal DNA-binding domain houses three C_2H_2 zinc fingers of invariant length between KLF members (Figure 1A). The first two zinc fingers are precisely 23 amino acids in length with the third zinc finger spanning 21 amino acids (22,23). Separating each zinc finger is a highly conserved seven amino acid spacer, the H/C link (TGE(R/K)(P/k/r)(F/y)X) (12,24). The overall physical structure of the factor is divided into an N-terminus comprised of two β-pleated sheets and a α-helical C-terminus with each zinc finger chealating a single zinc ion. Each individual zinc finger makes contact with a trinucleotide repeat within the major groove of DNA (25,26). Although KLF members typically recognize GC-rich sequences containing either a CGCCC or CACCC core sequence, recognition sites have been observed to vary from those that would be predicted from the amino acid composition of a particular zinc finger protein (14,15,27-29). Variation may attributed to co-operative binding between zinc fingers altering the overall selectivity of the C-terminus, a "wobble" effect akin to that observed in protein translation, or altered composition of the H/C link.

Sequence identity at the carboxyl terminus among KLF family members is greater than >65% (Figure 1B), suggesting the regulation of similar types of gene promoters and the potential for synergistic or antagonistic regulation by family members, with competition between factors functioning as "on" or "off" switches (19). Despite the high degree of sequence identity in their DNA-binding ability, the functional activities of KLF members differ widely. For instance, while Sp1 is one of the most potent activators of transcription, KLF11 is an equally powerful transcriptional repressor (17). Furthermore, several KLF family members have been observed to perform duel functions as activators in certain cellular contexts and repressors in others. The diversification displayed by family members is embedded in the high degree of variability within the N-terminal domain of each protein. The location of the NLS is one criterion to delineate family members, with one group possessing placement of the NLS within the zinc finger and the other directly upstream of this region (30,31). Far more diverse, however, are the

transcriptional regulatory domains of each family member that contain specific activation and repression elements that interact with distinct co-activators and co-repressors, thus permitting each factor to mediate the chromatin dynamics and transcription of individual gene promoters in a highly specific fashion (32-34). Although the highly conserved DNA-binding domain denotes their familial origins, the ultimate functional identity of each KLF protein rests in its N-terminal character.

***cabut*: evolutionarily conserved root for the KLF mediation of complex phenotypes**

KLF proteins are highly conserved among species, from *C. elegans* to human (35). The recent availability of genome sequence from many organisms has permitted the unprecedented cataloguing of KLF homologs/orthologs, permitting advanced understanding in the evolution of these factors (Figure 1C). Unlike other important transcription factors, such as the Kruppel-associated (KRAB) box containing C_2H_2 zinc finger proteins, which underwent significant cluster expansion, the genes encoding KLF proteins are dispersed through the genome with the exception of KLF1 and KLF2, which share the same locus (11,21).

In 2005, well after the identification of the human KLFs, Paricio reported the discovery of *cabut*, a novel gene involved in Drosophila embryogenesis, which encodes a transcription factor possessing a C_2H_2 zinc finger motif (36). *cbt* mutants suffer lethal defects in dorsal closure with *cbt* expression occurring downstream of the JNK signaling cascade to positively regulate expression of *dpp* and alter cytoskeletal dynamics along the leading edge (36). Structural analysis of *cbt* revealed that the closest human homologs are KLF10 and KLF11 (37,38). KLF10/11 possesses three repressive domains in their N-terminal region responsible for mediating interactions between co-activators and co-repressors. *cbt* displays complete conservation of the serine-and proline-rich R3 domain of KLF10/11 but incomplete R1 and R2 domains (38). Examination of effects of *cbt* mutation revealed that the protein is capable of positively regulating the activity of different TGF-β signals and mediating crosstalk between different pathways in the control of cellular differentiation and proliferation, a function completely conserved in its

human counterparts (37). Additionally, *cbt* is downstream of *forkhead* (*fkh*), target of rapamycin (TOR) signaling, and, potentially, insulin/PI3K signaling. Knockdown of *cbt* via RNAi resulted in increases in cells size and proliferation, indicating a potential growth suppressive function for the protein (39). Under conditions of dietary protein depravation, *fkh* translocates to the nucleus, activating growth-inhibiting *cbt*, which subsequently executes a growth suppressive program through a regulation of a network of genes (40). A similar growth suppressive function has been described by KLF11 in human pancreatic epithelial cells in response to TGF-β induction, a process disrupted in neoplastic transformation (41). The conservation of structure and function from invertebrate to vertebrate provides compelling evidence of the critical function of KLF proteins in the regulation of expansive, non-linear, highly interdependent gene networks in metabolism, growth, and cellular differentiation; functions exquisitely sensitive to environmental input.

Cbt homologs are found across a wide variety of vertebrate and invertebrate species, but are notably absent from nematodes (38). *C. elegans*, however, possesses three known KLF homologs, suggesting that the current family of human KLF homologs has multiple ancestral genes, indicating early expansion events. *Drosophilia* possesses six KLF homologs (42). Luna, CG12029, CG9895 correspond to KLF6/7 (43), KLF5, and KLF3/8/12, respectively. Bteb2 and cabut correspond to KLF15 and KLF9/10/11/13/14/16, respectively. The remaining two homologs, CG3065 and khb, appear to be unique to *Drosophila*. KLF17 similarly appears as a unique post-speciation event. Examination of the *Daphnia* genome reveals homology between five KLF homologs, KLF1A-E, and human KLF1/2/4, suggesting an independent vertebrate expansion. Despite rapidly divergent amino acid identities, function is highly conserved within the paralog clusters in vertebrates descended from their invertebrate orthologs (44). These data indicate high evolutionary pressure for strict conservation of DNA-binding capability with selective conservation of amino terminal domains to retain primary function while diverging rapidly post-speciation to adapt to the increasing complex regulatory demands of higher order organisms.

KLF proteins target epigenetic information to gene promoters to regulate gene expression

No single classification scheme may accurately characterize the function of any KLF member in its entirety. However, the classification of family members by virtue of their co-factors provides a framework to understand the functional differences between KLF subfamilies and the manner in which they engage chromatin to activate or repress transcription in a dynamic and reversible manner. KLF proteins deliver epigenetic information to gene promoters through three primary mechanisms: (i) sensing and translating environmental stimuli into a program of gene expression, (ii) sequence-specific targeting of chromatin remodeling complexes to gene promoters, and (iii) transactivation of other transcription factors to assist in the regulation of large networks of interdependent genes. Transient regulation of gene expression occurs through association with histone acetyltransferases (HATs) and histone deacetyltransferases (HDACs), while long-term gene silencing is enacted by interactions with histone methyltransferases (HMTs) and DNA methyltransferases (DNMTs).

Mechanism I: Translating internal and external input into gene regulation (Figure 2A)

KLF/Co-activator interactions: HATs, p300, CBP, and PCAF

Several members of the KLF family interact with co-activators to promote transcription. As described above, the chromatin landscape largely governs the accessibility of DNA sequences to transcription factors. As such, besides physical interaction between KLF protein and DNA sequence, the recruitment of co-activators to gene promoters is critical to the remodeling of chromatin to a permissive state. KLF proteins primarily restructure chromatin at target binding sites via recruitment of proteins that possess histone acetyltransferase (HAT) activity, such as E1A binding protein (p300) or CREB-binding protein (CBP). Together with acetylase p300/CBP associated factor (PCAF), p300 and CBP interact with numerous transcription factors (45). Histone acetylation at target promoters provides docking sites for “reader” molecules via a bromodomain designed to

recognize acetylation modifications (46). For instance, the ATPase unit of the SWI/SNF chromatin-remodeling complex contains one bromodomain. Action of bound SWI/SNF results in physical remodeling of nucleosome position along a stretch of template, permitting entry of transcriptional factors. p300/CBP and PCAF also acetylate other transcription factors, the effect of which is postulated to influence their stability and interaction with other protein in a manner similar to protein phosphorylation (47,48).

KLF/Co-repressor interactions: HDACs, Sin3, and CtBP

Concurrently in 1998, two groups codified the primary functional subfamilies of KLF repressors utilizing the divergent interactions with its then identified co-repressors as a system for division. Our laboratory described the TGF-β-inducible (TIEG) subfamily of repressors that operates through the Sin3/HDAC (histone deacteylase) system (17,32,34), while Turner and Crossley reported a subfamily that utilizes the C-terminal-binding protein (CtBP) co-repressor (49). Our group further characterized an extended family of Sin3/HDAC repressors—the basic transcription element-binding proteins (BTEBSs) (50,51)—based on structural similarity, although the TIEG and BTEB subgroups are functional indistinguishable.

The CtBP-dependent KLF repressors include KLF3, KLF8, and KLF12, which contain a five amino acid motif, PXDLS, which mediates interaction with CtBP (49,52,53). The degree of dissimilarity in the N-terminus between these three factors is high, such that the CtBP binding motif is the only point of conservation. The mechanisms by which CtBP proteins effect transcriptional repression in concert with their KLF partners are not fully characterized. One method of action may be through the recruitment of HDACs (54,55), although HDAC-independent repression through the recruitment remodeling complexes, such as Ikaros and members of the Polycomb complexes has been observed (54,56-58). The context in which CtBP recruits additional co-factors is unknown. As abolishment of CtBP binding through mutation of the PXDLS motif only partially abrogates the repressive function of KLF3 and KLF8, suggesting additional critical co-repressors (53,59). However, gene silencing appears to occur through rearrangement or alteration of

nucleosomes as nearly all identified interactions are with critical components of chromatin-remodeling complexes designed toward chromatin compaction and transcriptional repression.

The Sin3-dependent KLF repressors substantially characterized by our group are comprised of the TIEG and BTEB subfamilies, namely KLF9, KLF10, KLF11, KLF13, and KLF16, which utilize the HDAC system to execute gene repression through direct interaction with the scaffold co-repressor protein Sin3 (60). Sin3 proteins constitute large scaffolds with multiple protein interaction domains to permit assembly of large, often heterogeneous complexes with the ability to mediate repression through its associated subunits, which include HDAC1, HDAC2, RBAP46/18 (retinoblastoma protein (Rb)-associated proteins 46/48), among others (61). Sin3 mediates interaction between KLF and other transcription factors through four evolutionarily conserved imperfect repeats of ~100 residues each which are predicted to form a four-helix-bundle fold, known as a paired amphipathic helix (PAH) region (60,61). KLF16 is capable of executing repression through interactions with all three Sin3 isoforms and the HDAC system, suggesting even higher levels of plasticity in chromatin co-factor selection than previously realized (60).

Initial biochemical characterization of KLF10 and KLF11 revealed that the N-terminus of each possess three distinct transcriptional repressors domains, R1, R2, and R3 (32). Secondary structure predictions indicated that the R1 domain would adopt an α-helical conformation, which was later confirmed by circular dichroism analysis (34). Proline mutations within this domain significantly disrupted its repressive activity and subsequent binding studies uncovered Sin3a as the high affinity binding protein at this site (32,34). The core R1 domain was characterized as the SID (Sin3-interacting domain) motif and found to bind to the PAH2 domain of Sin3a to effect gene repression. KLF9, KLF13, KLF14, and KLF16 display conservation of the SID and execute gene repression through remodeling of chromatin marks via HDAC-mediated alterations to surrounding nucleosomes (32,34,50,51). KLF1 also interacts with Sin3a but lack an N-terminal SID,

interacting instead through its zinc finger domain (62). For instance, TGF-β pathway activation leads to the recruitment of KLF14-Sin3a-HDAC repressors complex to the TGFRII promoter, resulting in reductions in activating histone acetylation marks and increases in marks associated with transcriptional silencing, i.e. methylated H4-K20 (63).

Efforts to determine the interacting partners for the R2 and R3 domains of KLF11 have revealed additional interactions in transient gene regulation. For example, structural analysis of the R2 domains predicts the formation of a second, SID-like α-helical patch. Yeast two-hybrid results indicate strong binding affinity for RNA-recognition-motif (RRM) proteins responsible for splicing of primary mRNA following transcription (Urrutia, unpublished data). This data places KLF in intimate connect with transcriptional machinery and suggests a host of other novel functions in the regulation of gene expression that await characterization.

KLF/HMTS and DNMTs

Two systems predominate for instituting a repressive chromatin landscape in long-term gene silencing: Polycomb and Heterochromatin-associated Protein 1 (HP1). As opposed to the dynamic regulation of transcription by the activity of HATS and HDACs, HP1 and Polycomb proteins lead to permanent silencing through the activity of histone methyltransferases (HMTs). HP1 is one of the epigenetic "reader" molecules that serve as "gatekeepers" of silencing by recognizing methylated lysine 9 marks on histone 3 (H3-K9me) via its chromodomain and chromoshadow domains (64-66). Current knowledge supports a paradigm whereby HP1 proteins repress gene expression by binding to H3K9me marks and interacting to H3K9 HMTs, such as G9a or SUV39H1, which methylate this residue on adjacent nucleosomes thereby compacting chromatin and silencing gene expression (67,68). Thus far, this phenomenon has primarily been assumed to occur independently of sequence-specific DNA sites. HP1 remodels chromatin through interactions with HP1-binding proteins containing a consensus sequence, PxVxL, or in a PxVxL-independent manner (69-71). Our data demonstrated that the HP1α isoform interacts with KLF11 through its PxVxL domain within the

extreme C-terminus of the protein (72). HP1 recruitment to the CXCR4 promoter occurred in the presence of wild type KLF11 but not with KLF11 with deletion of the C-terminal HP1-binding site (KLF11ΔHP1). Failure to recruit Notably, the recruitment of SUV39H1 is contingent on the integrity of the KLF11-HP1 complex since this HMT was absent with KLF11ΔHP1. These results demonstrate that HP1α-SUV39H1 recruitment is dependent on KLF11 sequence-specific promoter binding and underscores the functionality of this recruitment in long-term gene silencing of gene targets.

Polycomb proteins may be divided into two categories, the components of Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2). The enzymatic engine of PRC2 is EZH2 (enhancer of zeste homologue 2), which possesses the SET domain that confers the complex with its methyltransferase activity (73-76). The H3-K27me3 mark imprinted by PRC2 subsequently serves to recruit PRC1, a large, heterogeneous complex that contains chromodomain-containing CBX proteins responsible for recognition of trimethylated H3-K27 (77). PRC1 functions to maintain the repressive state initiated by PRC2, although the mechanisms by which PRC1 accomplishes this task are unclear. Repression may be mediated through the recruitment of other chromatin-remodeling enzymes, specifically DNMTs and HDACs (78-82). PRC1 can restrict access of chromatin remodeling or transcriptional machinery to the DNA template, as evidenced by the ability of PRC1 to hinder ATP-dependent SWI/SNF remodeling (83,84). In *Drosophila*, Spps, a KLF family protein is critical to the recruitment of Polycomb complexes to gene promoters (85). Interactions between CtBP and vertebrate Polycomb homologs (related to PRC1 components) have also been observed in *Xenopus* through a six amino acid conserved motif in CtBP (58). Furthermore, our collaborative group has described an antagonistic relationship between the KLF10-PCAF pathway and PRC2 function in the regulation of the FOXP3 promoter. In the absence of KLF10, PCAF is not recruited to the FOXP3 promoter, resulting in a lack of H4 acetylation and a persistent block in FOXP3 transcription, resulting in a failure of adaptive Treg cells (86). The recruitment of PRC2 through binding of another KLF family member, similar to recruitment via Spps in *Drosophila*, is hypothesized, as is

the involvement of DNMTs to solidify the repressed chromatin state. These data describe for the first time a role for KLF proteins in guiding long-term gene repression and transcriptional memory.

Regulation of KLF proteins and pleiotropic effects

A wealth of evidence indicates that a large measure of ambiguity exists within the current classification of KLF repressors and activators, with the majority KLF protein displaying some degree of pleiotropic effects, i.e. activation when a factor typically mediates repression and vice versa. For instance, KLF1, the well-characterized erythroid-specific activator of the β-globin gene displays a repressive function under certain contexts (16). However, in its repressive function, KLF1 is not recruited via its DNA-binding domain to gene promoters but through association with other DNA-binding proteins and is capable of association with the Sin3-HDAC system (62). Another example is KLF13, which activates several promoters, including SV40, RANTES, and γ-globin, but represses others, such as CYP1A1 (50,87-89). Given our incomplete understanding of the entire array of co-repressor and co-activator interactions, it is reasonable to speculate that all KLF family members are capable of activation or repression in a context-dependent fashion.

The mechanisms by which these factors selectively recognize associated co-repressors or co-activators to determine their function remains purely conjecture. Promoter context, to be discussed in the following section, is likely to play a large role. Additionally, KLF proteins are capable of undergoing a variety of posttranslational modifications, including acetylation, phosphorylation, ubiquitination, and SUMOylation, that alters their transcriptional activity through modulation of binding affinity downstream of signaling pathways (reviewed in detail in (90)). Our laboratory and collaborators described a novel pathway in which phosphorylation of KLF11 by extracellular signal-regulated kinase (ERK)/mitogen-activated protein kinase (MAPK) abrogates the interaction with Sin3a and inhibits repression of *Smad7* (91). Furthermore, we demonstrated that KLF11 inhibits prostaglandin E(2) synthesis through transcriptional silencing of its biosynthesis enzyme,

cytosolic phospholipase A2alpha. KLF11 accomplishes this repression through recruitment of the Sin3 remodeling complex, an interaction interrupted by epidermal growth factor receptor-AKT-mediated phosphorylation at T56, within the SID (92). Phosphorylation of Y10 of KLF16, potentially by Src tyrosine-kinase, also disrupts its binding with Sin3a, disruption repression of *CYP1A1*, affecting estrogen metabolism and uterine pathobiology (60). Like KL11, KLF16 regulation is directed downstream of cellular signaling rather than being constitutive (60,93). The diverse array of pos-translational modifications and the complex network of protein-protein interactions permit precise, yet dynamic regulation of gene promoters channeled through KLF. Regulation of KLF affinity through posttranslational modifications provides a direct link from environmental induction of signaling cascades and the delivery of epigenetic information to gene promoters by influencing interactions with co-factor chromatin remodeling complexes. Disruption of these associations is a prime mechanism in the development of complex disease phenotypes.

Mechanism II: Sequence-specific targeting of chromatin remodeling complexes (Figure 2B)

Posttranslational modifications of KLF proteins only partially resolves the conundrum of what determines the selective interactions with co-activators or co-repressors that defines their function at a given gene promoter. In describing one of the first known KLF factors, Imataka observed that KLF9 could activate transcription if a promoter element contained multiple GC boxes but behaved as a repressor on promoters containing only a single copy of the sequence (13). For decades, the transcriptional field focused on characterizing and cataloguing the consensus sequences of binding sites, holding tight to the paradigm of a single binding site as the ultimate, passive determinant in the binding of a transcription factor. Two exciting discoveries of our laboratory support the notion that regulation of gene expression by KLF proteins is highly dependent on promoter context. In this paradigm, KLF binding and function is determined by both the KLF binding sites as well as neighboring transcription factor binding sites, forming a *cis*-regulatory element or block.

Polycomb Response Elements (PREs)

In *Drosophilia*, Polycomb remodeling complexes are recruited to gene promoters through specific *cis*-regulatory domains known as Polycomb Response Elements (PREs), which contains clusters of DNA binding sites for Polycomb proteins. The relationship between KLF proteins and Polycomb in Drosophila is extensive. Early studies in *Drosophil*a identified the importance of both Krüppel and Polycomb as critically relevant to regulation of the homeotic gene *Scr* (Sex combs reduced) (94). Subsequent work on a homeotic gene within the bithorax complex *Abdominal-B* provided evidence for a generalized model of gap gene products (Krüppel) promoting stable silencing through Polycomb in homeotic *Drosophila* genes (94). Perhaps the greatest revelation, however, has been the report that KLF binding sties are present in the majority of Drosophila PREs and that binding of Spps, a KLF protein, is necessary for the full activity of PREs (85,95).

Identifying mammalian PREs has proven difficult, as the majority of the recruitment machinery is not conserved from *Drosophila*. KLF10 knockout mice display impaired *FOXP3* activation with impaired adaptive T regulatory cell differentiation (96). Chromatin immunoprecipitation (ChIP) studies revealed that in $KLF10^{-/-}$ mice, the FOXP3 promoter is occupied by EZH2 with subsequent increases in H3-K27 trimethylation (86). We hypothesized that silencing of *FOXP3* by Polycomb is a default mechanism that must be overcome by KLF10 to achieve an inducible state. Bioinformatics analysis of the mammalian FOXP3 promoter element revealed binding sites for KLF proteins, Yin Yang 1 (YY1), the vertebrate ortholog of *Drosophila* Pho, and GAGA factor (GAF), the ortholog of Pipsqueak. Introduction of this promoter element fused with luciferase into an epithelial cell line resulted in progressive silencing of luciferase expression. Subsequent studies detailed a mechanism by which KLF10/PCAF antagonizes the function of PRC2, resulting in increased acetylation and induction of *FOXP3* expression (86). The potential for other KLF factors to bind and modulate the activity of mammalian PREs is an area of active investigation.

HP1 Response Elements (HREs)

Elegant studies have previously demonstrated KRAB (Krüppel-Associated Box)-containing zinc finger transcription factors recruit KAP1 (KRAB-ZFP Associated Protein 1), which binds to HP1 through a transcriptional intermediate, KAP1. On the other hand, KLF11 binds to its target DNA sequence regardless of whether it is bound or not to HP1. The two obvious differences indicate that the biophysical and biochemical properties that regulate these two mechanisms for recruiting HP1 to promoters are readily distinct. In an elegant biochemical study, Lomberk demonstrated that an extended KLF11 binding site (CCGCCCCGCCCC) mediates the sequence-specific recruitment of HP1-SUV39H1 to promoters with an associated increased in H3K9me3 marks (72). HP1 triggers this function by being recruited to promoters in a sequence-specific manner rather than its well-characterized binding to methylated chromatin, indicative of how this system can function in the regulation of gene expression with a higher degree of specificity. Sequence analysis of a subset of gene targets affected by the uncoupling of KLF11 from HP1 reveals that in addition to conservation of the extended KLF11 binding site, the gene promoters share 100-200bp regions of similarity centering around the KLF11 binding site (Lomberk and Urrutia, unpublished data). Similar to the KLF-PRE described here, the KLF HP1 Recruitment Elements (HREs) are comprised of binding sites for other transcription factors, indicating that additional combinatorial effects, besides those critically mediated by KLF proteins, may modulate the recruitment of both HMT systems to promoters.

Disruption of the interaction between KLF11 and HP1 via mutation results in the deregulation of a large network of genes related to the regulation of cell death, proliferation, and senescence, consistent with KLF11's previously characterized role as a tumor suppressor (72). A subsequent study probed further into the relationship between KLF11 and HP1 in the regulation of neuronal differentiation networks. We demonstrate that KLF11 binds to a distinct KLF site within the Drd2 promoter and recruits p300 histone acetyltransferase to activate its expression (97). Disrupting the interaction between KLF11 and HP1 further enhances Drd2 transcriptional activation, indicating that

the interplay between these two pathways functions as a rheostat to maintain appropriate levels of Drd2.

Overall, the discovery of this KLF11-mediated HP1 recruitment mechanism, together with the knowledge of KLF-PREs, suggests that KLF sites have undergone evolutionary pressure to maintain their ability to recruit potent HMT-based silencing complexes. Furthermore, the observation that KLF can work via histone acetylation (p300), deacetylation (Sin3a), and histone methylation (HP1), as described above, indicates that this sequence-specific transcription factor has many functions which occur via different pathways or alternatively, that several pathways are required for the same function. In addition, it is interesting to consider the convergence of short-term (Sin3a) and long-term (HP1) repression mechanisms on the same KLF molecule. Dysfunction of mechanisms of long-term repression precipitates the permanent silencing of critical genes through aberrant histone methylation or DNA methylation. For instance, in pancreatic cancer, the p16 tumor suppressor is effectively knocked out due to hypermethylation of the CDKN2A promoter (98). The possibility that KLF proteins may effect long-term gene silencing of gene targets opens a previously unexplored pathway by which these factor may lead to constitutively activated or repressed genes, blurring the line between "soft" and "hard" inheritance even further.

Mechanism III: Transactivation of other transcription factors and cross-regulation within the KLF family (Figure 2C)

Amplification of membrane-to-nucleus signaling requires multistep induction of cascades of gene expression. The activation of one transcription factor induces the expression of other transcription factors, resulting in exponential amplification of the initial signal as well as a high level of control in fine-tuning expression of large, highly interdependent networks of genes downstream of the initial activation point. Recent evidence suggests that some if not the majority of KLF proteins occupy a position at the apex of signal transduction cascades, modeling the activity of potent downstream transcription factors to significantly expand and diversify its repertoire of gene targets.

Instances of modulation of downstream transcription factor pathways through KLF with "local" effects abound. Transforming growth factor-β (TGF-β) is a pleiotropic cytokine that regulates a large number of cellular processes, including cell growth, apoptosis, differentiation, migration, and metastasis (99). TGFβ binds to its transmembrane receptor kinase complex to induce Smad signaling molecules, which translocate into the nucleus to regulation transcription. KLF10 and KLF11 were TGFβ-inducible factors cloned and characterized in our laboratory to function as effector proteins in the control of cell growth and differentiation (18,32,41,100). Nuclear interaction of Smad3 and KLF11 results in the repression of the c-Myc oncogenic transcriptional factor, attenuating downstream cascades initiated by Myc. Furthermore, KLF11 inhibits Smad7, breaking the negative feedback loop on the expression of TGFβ and amplifying Smad-mediates signaling. A similar function has been observed in the ability of KLF4 and KLF5 to modulate immune response in the gastrointestinal tract, where both transcriptional factors are highly enriched. KLF5 is activated through the MAPK and ERK1/2 pathways downstream of lipopolysaccharide (LPS) exposure, a bacterially-derived endotoxin (101-103). LPS activation of transmembrane toll-like receptors (TLRs) leads to induction of MAPK signaling, activation of KLF5, and the subsequent KLF-dependent activation of the NF-κB transcription factor expression, eliciting a proinflammatory cascade response (103). KLF has been demonstrated to regulate other transcription factors, including PPARγ (104,105). Furthermore, high-throughput screening methods (microarray, RNAi, ChIP-seq, GWAS, among other) have revealed the ability for KLFs to regulate much larger cascades of transcription factors than previously imaged (Figure 3).

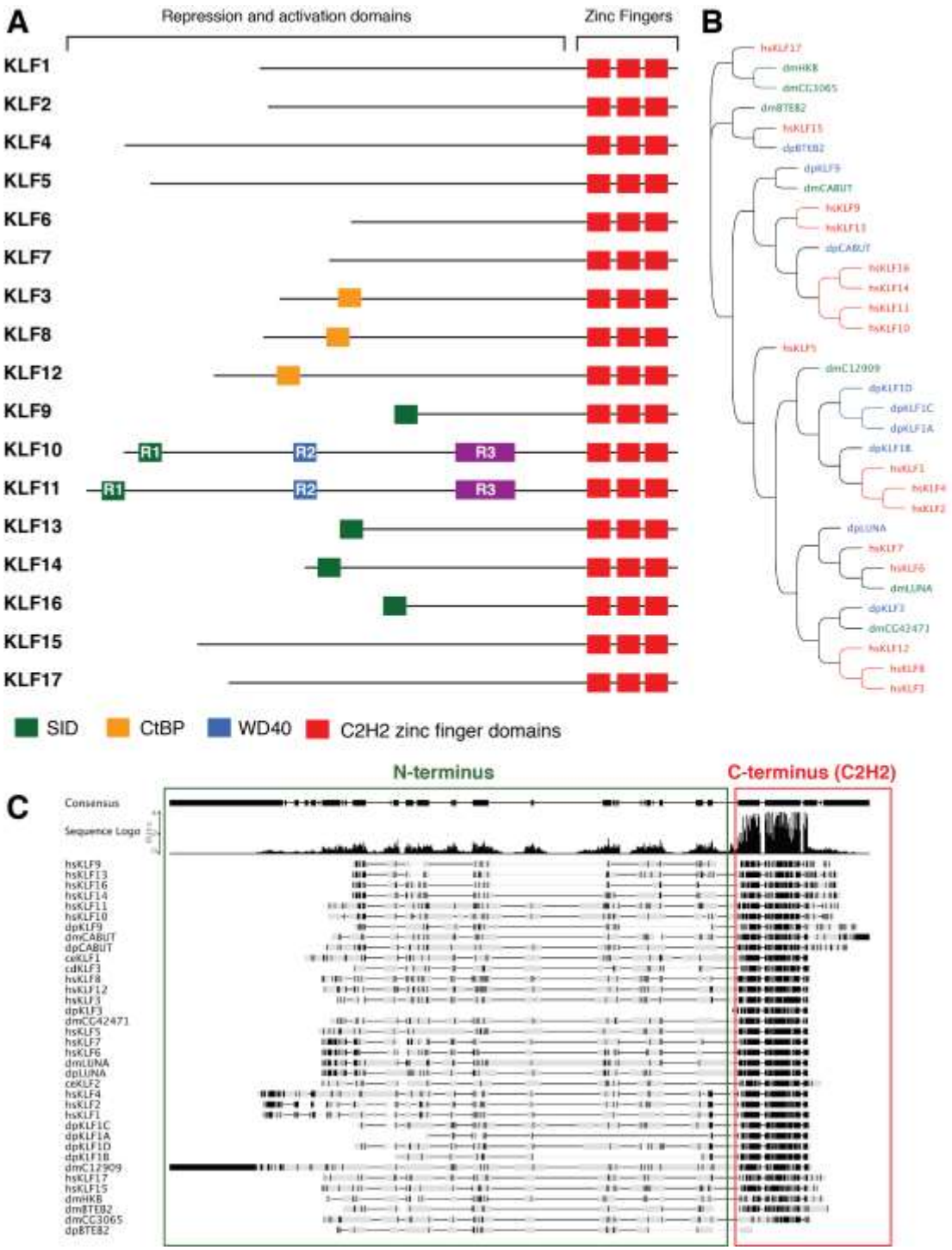


FIGURE 1. Structure and evolution of the KLF family of transcription factors. A. Structural domains of the human KLF protein family. B. Differential evolution of the variable C and invariant N terminals of KLF proteins from fly to human. C. Evolutionary conservation of ancesteral gene *cabut* to KLF descendants.

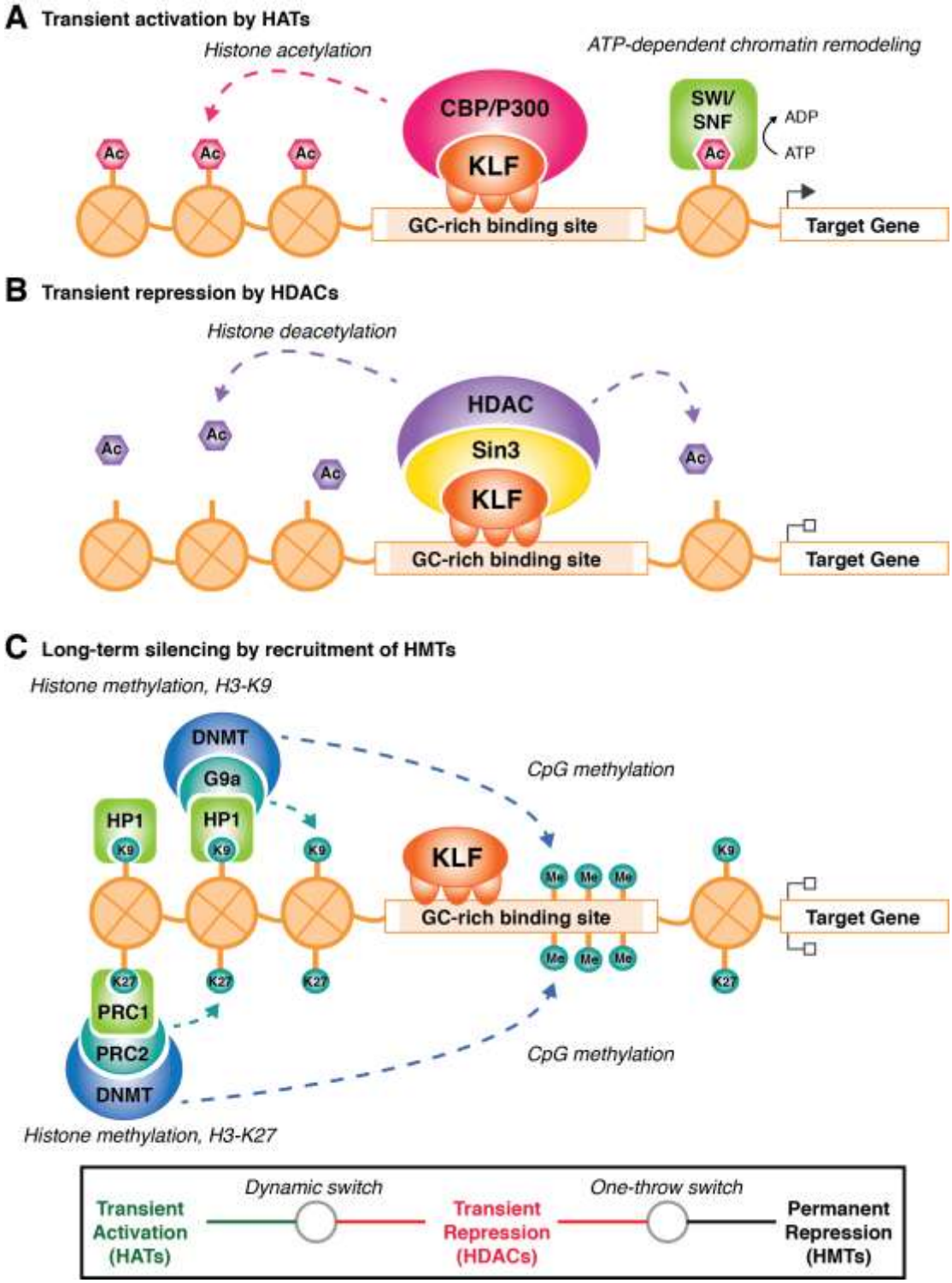


FIGURE 2. Mechanisms of the delivery of epigenetic information to gene promoters by KLF proteins. A. Sequence-specific recruitment of chromatin-remodeling complexes by KLF proteins by coupling of KLF11 to HATs (A) in transient activation, HDACs in transient repression (B), and HMTs in long-term gene silencing (C).

CHAPTER 2:

Krüppel-like Factor 11 Regulates the Expression of Metabolic Genes *via* an Evolutionarily Conserved Protein-Interaction Domain Functionally Disrupted in Juvenile Diabetes

Gwen Lomberk[1], Adrienne Grzenda[1], Angela Mathison[1], Jin-San Zhang[2], Ezequiel Calvo[3], Laurence J. Miller[4], Juan Iovanna[5], Martin E. Fernandez-Zapico[2], and Raul Urrutia[1]

[1]Laboratory of Epigenetics and Chromatin Dynamics; Epigenomics Translational Program, Mayo Clinic Center for Individualized Medicine; Division of Gastroenterology and Hepatology;
Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, United States; [2]Schulze Center for Novel Therapeutics, Division of Oncology Research, Mayo Clinic, Rochester, MN, United States; [3]Molecular Endocrinology and Oncology Research Center, CHUL Research Center, Quebec, Canada; [4]Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Scottsdale, AZ, USA; [5]Centre de Recherche en Carcérologie de Marseille (CRCM), INSERM UMR 1068, CNRS UMR 7258, Aix-Marseille University and Institut Paoli-Calmettes, Parc Scientifique et Technologique de Luminy, Marseille, France

J Biol Chem. 2013 Jun 14;288(24):17745-58

SUMMARY

The function of Krüppel-like Factor 11 (KLF11) in the regulation of metabolic pathways is conserved from flies to human. Alterations in KLF11 function result in MODY VII and neonatal diabetes, however, the mechanisms underlying the role of this protein in metabolic disorders remain unclear. Here, we seek to investigate how the MODY VII A347S genetic variant, present in patients with juvenile diabetes, modulates KLF11 transcriptional activity. A347S affects a previously identified transcriptional regulatory domain 3 (TRD3) for which co-regulators remain unknown. Structure-oriented sequence analyses, described here, predict that the KLF11 TRD3 represents an evolutionarily conserved protein domain. Combined, yeast two-hybrid and protein arrays demonstrate that the TRD3 binds WD40, WWI, WWII, and SH3-domain containing proteins. Using one of these proteins as a model, guanine nucleotide binding protein beta 2 (Gβ2), we investigate the functional consequences of KLF11 coupling to a TRD3 binding partner. Combined, immunoprecipitation and biomolecular fluorescence complementation assays confirm that activation of three different metabolic GPCR receptors (□-adrenergic, secretin and CCK) induce nuclear translocation of Gβ2 to directly bind KLF11 in a manner that is disrupted by the MODY VII A347S variant. Using genome-wide expression profiles, we identify metabolic gene networks impacted upon disruption of the TRD3. Thus, this study characterizes a novel protein-protein interaction domain disrupted in a KLF gene variant that associates to juvenile diabetes, contributing to our understanding of gene regulation events in complex metabolic diseases.

INTRODUCTION

KLF11, a human ortholog of the *Drosophila* gene *cabut*, belongs to the Krüppel-Like Family (KLF) of transcription factors. Members of this family regulate GC-promoters in organisms ranging from flies to humans (106). Rapidly emerging evidence demonstrates that these *cabut*/KLF pathways regulate important metabolic processes conserved in organisms ranging from flies to humans (38). For instance, *cabut* is a transcriptional regulator of metabolic gene pathways in *Drosophila* (40,107). Disruption of KLF pathways leads to biochemical alterations and metabolic impairment, often resulting in lethality. In fact, human variants in the KLF11 protein (MODY VII) as well as on its DNA binding site within the insulin promoter cause juvenile and neonatal diabetes, respectively (108,109). Moreover, extensive studies have demonstrated KLF11 binds and regulates many promoters of genes involved in cholesterol, prostaglandin, neurotransmitter, fat, and sugar metabolism (108-114). Thus, the medical significance of this knowledge led us to study how alterations in KLF11 proteins impact the regulation of metabolic gene networks of relevance toward better understanding of complex human diseases.

KLF11 is a well-characterized protein in terms of its ability to couple to several chromatin pathways and epigenetic regulators, including Sin3a, histone acetyltransferase (HATs), and HP1 (34,115,116). The following study has been designed to further our mechanistic understanding of how disease-associated alterations that affect the membrane-to-nucleus coupling of KLF11 to the regulation of metabolic genes. Contrary to the variants in KLF11 DNA binding sites, which have the power to uncouple a single gene promoter (e.g. c.-331 INS variants) (108), alterations in chromatin coupling can disrupt complex functions, such as metabolism, by altering the expression of entire gene networks. Interestingly, variants that fall near the Sin3a domain alter the ability of KLF11 to regulate metabolic target genes and associate with MODY VII (109). The KLF11-mediated HAT pathways appear to impact the regulation of insulin in neonatal diabetes (108). However, the mechanisms and function of other KLF11 mutations that associate with diseases remain to be characterized. Thus, studying these types of alterations can aid

in better defining the identity of KLF11-mediated metabolic gene networks, the mechanisms of their regulation, and the mechanisms that inactivate these pathways in humans.

The KLF11/MODY VII A347S variant found in juvenile diabetic patients (109) maps to a previously characterized transcriptional regulatory domain (TRD3), for which functional cofactors have remained unknown (117). In the current study, we demonstrate that the KLF11 TRD3 functions as a novel protein-protein interaction domain, which function is altered in the A347S diabetes variant. We further show that activation of cell surface receptors involved in metabolism (adrenergic, secretin, and CCK receptors) can induce the nuclear translocation of a TRD3 binding protein (Gβ2) to bind KLF11 in order to regulate metabolic gene targets. Combined, these studies reveal the existence of a novel KLF-mediated pathway for coupling extracellular signals to the regulation of gene expression related to metabolism and diabetes. These findings extend our understanding of transcriptional regulatory events, which are disrupted in human metabolic diseases.

EXPERIMENTAL PROCEDURES

Tissue culture and reagents - Cell lines were obtained from the American Type Culture Collection (ATCC, Rockville, MD). Cells were cultured as described previously (118,119). CHO cells stably transfected with the Cholecystokinin (CCK) Receptor A (CHO-CCKAR) were grown as described previously (120). Isoproterenol, secretin and CCK were purchased from Sigma-Aldrich (St. Louis, MO).

Plasmids and recombinant adenovirus - Standard molecular biology techniques were used to clone full length KLF11, KLF11-A347S, Gβ2, as well as specified deletions into pcDNA3.1/His (Invitrogen, Carlsbad, CA), pCMV-Tag 2B (Stratagene, La Jolla, CA), pGEX (GE Healthcare, Piscataway, NJ), and EYFP vectors as previously described (34,70). QuickChange® Site-Directed Mutagenesis was performed as suggested by the manufacturer (Agilent Technologies, Inc, Santa Clara, CA). All constructs were verified by sequencing at the Mayo Clinic Molecular Biology Core Facility. Epitope-tagged

(6XHis-Xpress™) KLF11 and KLF11 A347S variant as well as empty vector (Ad5CMV) were generated as recombinant adenovirus in collaboration with the Gene Transfer Vector Core at the University of Iowa.

GST pulldown assays, Immunoprecipitation and Western Blot - GST and GST fusion protein purification, *in vitro* translation, GST pulldown assays, immunoprecipitation and western blot were all done as previously described (70). Antibodies were used against the FLAG (Sigma) or His-Tag (OMNI D8; Santa Cruz Biotechnology, Santa Cruz, CA) to detect recombinant expression of KLF11, KLF11 A347S variant, and Gβ2, endogenous Gβ1/2 (Santa Cruz Biotechnology), HP1α (Millipore, Billerica, MA), Sin3a (Santa Cruz Biotechnology), or CBP (Abcam, Cambridge, MA).

Algorithm For Identifying TRD3 - To identify putative evolutionarily conserved domains, the N-termini prior to the conserved zinc finger regions of cabut (Drosophila) and KLF1-17 (Homo sapiens) were analyzed using a sliding window of 15 amino acids. Windows were scored for overall hydrophobicity using the Kyte-Doolite scale (121); polyproline helix conformations using experimental derived values (122); and overall proline, glycine, and glutamine content (PGQ index). Regions were screened via pairwise alignment with KLF10 and KLF11 for R1 and R2 domains to determine the N-terminal boundary. Pairwise alignments were performed using the MUltiple Sequence Comparison by Log-Expectation Program (MUSCLE) (123). Phylogenetic tree was constructed using Geneious Tree Maker (Biomatters Ltd, Auckland, New Zealand) with the following settings: BLOSUM62 Matrix, Global alignment with free gap ends, outgroup set to cabut.

Genome-wide expression profiles - Pancreatic cells (Panc1) were plated at a density of 10^6 cells and transduced with empty vector, KLF11, or KLF11-A347S (Ad5CMV). RNA was prepared as previously described (124). Global gene expression profiling was carried out at the Microarrays Facility of the Research Center of Laval University CRCHUL utilizing the Affymetrix Human Gene 1.0 ST arrays (28,869 well-annotated genes and 764,885 distinct probes). Intensity files were generated by Affymetrix GCS 3000 7G and

the Gene-Chip Operating Software (Affymetrix, Santa Clara, CA). Data analysis, background subtraction and intensity normalization was performed using Robust Multiarray Analysis (RMA) (125). Genes that were differentially expressed along with false discovery rate were estimated from t test (>0.005) and corrected using Bayes approach (126,127). Data analysis, hierarchical clustering, and ontology were performed with the OneChanelGUI to extend affylmGUI graphical interface capabilities (128), and Partek Genomics Suite, version 6.5 (Partek Inc., St. Louis, MO) with ANOVA analysis. The criteria of log2 fold change ±1.5 and a p-value of <0.05 compared to empty vector control or wild type KLF11 levels were used to determine significant gene targets. A subset of genes was validated by qPCR as previously described (124,129). The criteria of log2 fold change ±1.5 and a p-value of <0.05 compared to empty vector control was used to determine significant gene targets. Hela cells were plated at a density of 10^6 cells and transfected with empty vector, Gβ, or Gβ fused in frame to 3 copies of the SV40 NLS (130). RNA was prepared as previously described above. Global gene expression profiling was carried out by the Microarray core of Mayo Clinic utilizing the Affymetrix Human Genome U133 Plus 2.0 Array. The criteria of log2 fold change ±1.5 and a p-value of <0.05 compared to empty vector control was used to determine significant gene targets.

Chromatin immunoprecipitation (ChIP) - ChIP was performed as previously described (70,116,129,131) using antibodies against His-Tag (OMNI D8; Santa Cruz Biotechnology) to detect recombinant expression of KLF11 or endogenous Gβ1-2 (Santa Cruz Biotechnology). Binding activity was derived using the NimbleGen human promoter hybridization system (Madison, WI). Peaks were detected by searching for >4 probes where signals were above the specified cutoff values (90% to 15%) using a 500bp sliding window along 5kb upstream of the transcriptional start site in human promoters. Each peak was assigned a score that is the log2 ratio of the fourth highest probe in each peak. If multiple peaks are present, the peak nearest the TSS is reported. Ratio data was

then randomized 20 times to evaluate the false discovery rate (FDR). Only peaks with FDR scores <0.2 were deemed high confidence binding sites and reported.

Identification of TRD3 Binding Proteins using yeast two-hybrid system - As bait, the coding region of the KLF11 TRD3 (amino acids 273-351) was cloned into the pGBKT7 plasmid of the Matchmaker system (Clontech). The integrity and expression of the fusion construct was confirmed by sequencing. No autoactivation of the reporters was associated with this bait construct, as determined by cotransformation of the bait with prey library vector in host AH109 cells. Sequential library scale transformations were performed from a normal bone marrow cDNA library. A total of 2.5 million clones were screened and selected on high-stringency plates (SD/–Ade/–His/–Leu/–Trp) coated with X-β-Gal. After incubation for a period of 72–96 hours at 30°C, colonies were recovered and DNA from each colony was extracted and sequenced. The cDNA inserts from yeast clones were amplified by PCR using primers 5′CTATTCGATGATGAAGATACCCCACCA (forward) and 5′GTGAACTTGCGGGGTTTTTCAGTATCTACGA-3′ and sequenced at the Mayo Molecular Biology Core Facility. In order to eliminate false positives, isolated library prey plasmids were transformed into Y187 yeast and crossed with AH109 yeast carrying either the empty plasmid or the bait plasmid; activation of the reporter gene was assessed by growth in SD -Trp/-Leu/-His, plus 3-AT.

Identification of TRD3 Binding Proteins using Solid Phase Binding Assays - To screen for binding between the KLF11 TRD3 and SH3 or WW domain containing proteins, solid phase SH3 and WW domain arrays were obtained from Panomics (Fremont, CA) and processed according to manufacturer's instructions utilizing 10-20μg/ml of purified His-KLF11 TRD3 (amino acids 273-351) for hybridization.

Immunofluorescence and Bimolecular fluorescence complementation (BiFC) - CHO-CCKAR cells were transfected with Flag-tagged or His-tagged Gβ2 and KLF11 expression constructs. The fluorescence was imaged using an LSM510 microscope (Zeiss, Heidelberg, Germany). For the BiFC analysis of individual cells by confocal

microscopy, HeLa and Capan2 cells were transfected with KLF11-EYFP1 and Gβ2-EYP2 expression vectors as previously described (70,116).

RESULTS

The KLF11 A347S variant maps to a protein interaction domain conserved in the KLF family of metabolic transcription factors. Previous studies in human populations have demonstrated that defined changes in the sequence of the KLF11 protein (Q62R, T220M and A347S) associate to early-onset type II juvenile diabetes mellitus **(OMIM:** MODY VII) (109). Functional studies demonstrated that these three **KLF11** variants have impaired transcriptional activity compared to wild type. Subsequent studies focused on characterizing the mechanism of function for KLF11Q62R demonstrated that this protein is defective in binding to Sin3a, resulting in lower levels of insulin expression and promoter activity (109). In contrast, here, our structure-oriented bioinformatics analyses demonstrate that the A347S variant falls within the regulatory domain 3 of KLF11 (TRD3) and has the potential to disrupt the function of this domain for which cofactors have long remained to be identified (Figure 1A). Analysis of the KLF11 TRD3 sequence within the region affected in MODY VII (A347S) as well as the correponding region of its fly ortholog *cabut*, identified several proline-rich motifs (PRM, Figure 1B). PRMs are present in many transcription factors, chromatin regulators, and epigenetic proteins and facilitate protein-protein interactions (132). Interestingly, we find a high degree of conservation of these PRMs in all KLF proteins which are orthologues of the fly cabut gene, a known metabolic regulator (38). These human proteins form two KLF subfamilies, namely the TIEGs (KLF10 and KLF11) and BTEBs (KLF9, KLF13, KLF14, and KLF16) (106). The PRM motif identified in these studies displays key characteristics, falling into distinct structural types, two of them mark both the N-terminal (PRM1: PPØPØØØQØØP) and C-teminal (PRM4: GØXXØØPØØPØP) boundaries for the TRD3. Two other PRM domains are located centrally (PRM3: QØØPØPQPØØØGP and PRM3: ØØPPPØPØØØ). All of the PRMs are embedded within a biochemical-biophysical environment entirely provided by hydrophophic amino acids and few flexible residues such as G and Q. Notably, analyses of extensive available data from NMR, x-ray

crystallography, phage display experiments, and liquid-phase binding assays demonstrate that PRM of this type are highly enriched within bonafide protein-protein interaction modules (132-134). These repetitive PRM motifs are shared by several protein-protein interaction modules. Binding of PRMs to their partners is required to form structures that have more stable thermodynamic properties. Therefore, to guide further biochemical studies, we developed a computational model that defines this KLF protein-protein interaction domain (Figure 2A), which reveals that the TRD3 is located immediately upstream of the KLF DNA binding domain in the region affected by MODY VII A347S. Evolutionary analysis of the TRD3 region further demonstrates the high pressure for preservation of the domain between fly and human across the cabut-KLF family (Figure 2B). Thus, combined these results support the prediction that the TRD3 domain functions as a proline-rich module that likely mediates protein-protein interactions, a hypothesis that we subsequently tested experimentally.

In order to define potential binding partners for this protein interaction domain we utilized both, the yeast two-hybrid system and domain-specific protein binding arrays. The bait for yeast two hybrid studies as well as the recombinant protein for hybridization to protein arrays were derived from the region of KLF11 encompassing the TRD3, consisting of amino acids 273-351. The advantage of utilizing these two distinct assays was to perform both unbiased (yeast two hybrid) and candidate-based (protein array) approaches to identify potential TRD3-interacting proteins. For protein array experiments, we used a high TRD3/bound array protein ratio (5/1) so that proteins from the array would capture different amounts of TRD3 according to their binding affinities, as detected by different intensities. Arrays also included both positive and negative controls. Thus, the intentities in the authoradiographs gave an indication of relative binding strengths. These experiments show that, the KLF11 TRD3 binds to several WD40-containing proteins, including Gβ1, Gβ2, WD40 Repeat Domain 6, and echinoderm microtubule associated protein like 2, in addition to SH3- (8 positives from a total of 38), WWI- (10 positives from a total of 34), and WWII-containing proteins (4 positives from a total of 33) (Table 1). When analyzed, the potential TRD3 interactome

defined by both methods reveals a network of proteins which, similar to KLF11, participate the regulation of transcription. Furthermore, some of these factors, such as ITCH, have previously identified roles in the functional regulation of cabut proteins (KLF10) (135), thus providing internal validation for our experiments. Interestingly, though with less conservation in overall domain structure, TRD3-like domains are present in a large number of KLF proteins outside of the cabut subfamily (Supplemental Table 1), some of which have been shown to interact with proline-rich binding proteins such as Nedd4-like protein, WWP1, for KLF2 (136), as well as YAP and WWP1 for KLF5 (137,138). Thus, from these experiments, we conclude that cabut members, as well as other KLF proteins, appear to have undergone evolutionary pressure to conserve proline-rich domains that harbor the potential to regulate their function. Combined, these structural relationships led us to perform subsequent studies aimed at clarifying the biochemical properties and functional impact of the TRD3 interactions in the transcriptional regulation of metabolic gene networks focused on isolating those which are disrrupted by the MODY VII A347S variant.

Membrane-to-nucleus signaling pathways regulate the KLF TRD3 binding domain

For subsequent functional studies, we choose to study Gβ2 as a model for a TRD3-interacting protein since this interaction was identified as our most represented candidate via several clones from the yeast two hybrid system. This is functionally relevant since in this system protein-protein interaction occurs in the eukaryotic nucleus (yeast), the compartment in which KLF11 performs its function. Additionally, the interaction of a few KLF proteins with SH3 and WW domains have been previously characterized (135-139) but those with WD40-containing proteins, such as Gβ (140), remain unknown, increasing the potential novelty of our findings. Furthermore, Gβ is known to regulate cell signaling cascades that are critical for maintaining metabolic homeostasis, including in diabetes (141-143). In fact, genetic variants of these proteins have also been assocaited with the development of this disease (141,144). Previous reports have also shown that the Gβ subunits translocate to the nucleus and interact with other transcription factors, including fos and the glucorticoid receptor (GR) (145,146). Therefore, in order to

confirm the interaction between KLF11 and Gβ2 detected by our yeast two hybrid experiments, we initially performed in vitro binding assays using ^{35}S-labeled Gβ2 and GST fusion protein of the KLF11 TRD3 (amino acids 273-351). We demonstrate that Gβ2 indeed binds the KLF11 TRD3 *in vitro* (Figure 3A). Furthermore, based upon the sequences recovered from the positive yeast two hybrid clones, we utilized deletions of Gβ2 to determine the region of this protein required for interaction with KLF11. These investigations reveal that binding with KLF11 is maintained by Gβ2 deletions containing a minimum of amino acids 209-340, as observed by positive GST-KLF11/^{35}S-Gβ2 interaction for Gβ2Δ1-105, Δ1-120, and Δ1-208 (Figure 3B). However, interaction was not detected with the N-terminal regions encompassing Gβ2□1-214) or Gβ2□106-214), as well as Gβ2Δ1-221 (Figure 3B), which presumably disrupts the protein in WD40 repeat 5. Therefore, Gβ2 with intact WD40 repeats 5-7 is necessary for interaction with KLF11.

We next investigated this interaction in cultured cells and whether the KLF11-Gβ2 can transduce signals from the membrane to the nucleus. For this purpose, HeLa cells under basal were co-transfected with His-tagged full length KLF11 and either empty Flag control vector or Flag-tagged Gβ2. Upon immunoprecipitation of Flag-tagged Gβ2, KLF11 was detected (Figure 3C). This result was further confirmed by immunoprecipitation of Flag-tagged Gβ1, as this isoform of Gβ was also recovered by our yeast two hybrid assay, which corroborated binding to KLF11 (Figure 3C). To examine the influence of G protein-coupled receptor (GPCR) activation at the plasma membrane by a physiological stimulus, we first utilized isoproterenol, an agonist of □-adrenergic GPCR signaling, which regulates many functions in HeLa cells (147). Interestingly, a more robust immunoprecipitation of KLF11 was observed with Gβ2 upon isoproterenol treatment (Figure 3C). In order to visualize this interaction in living cells, we used BiFC by fusing the EYFP N terminus (EYFP(1)) to KLF11 and the EYFP C-terminal portion (EYFP(2)) to Gβ2. These constructs were co-transfected in both HeLa and Capan2 cells and subsequently stimulated with isoproterenol or secretin (148), respectively, for 30min prior to fixation. Secretin is an agonist for the GPCR family

generally with high levels in Capan2 cells (148). Interaction was detected in both cell types upon their respective physiological stimulus as evidenced by yellow fluorescence reconstitution. Fluorescence resulting from the KLF11-Gβ2 interaction was localized, as expected, in the nucleus as determined by nuclear DAPI co-stain (Figure 3D). Co-expression of a negative control leucine zipper protein with either KLF11 or Gβ2 fused to their respective EYFP halves did not reconstitute fluorescence, confirming the fluorescence obtained is specific and devoid of background (data not shown). Representative panels of untreated HeLa or untreated Capan2 cells are also shown (Figure 3D). To complement these studies, we also utilized CHO cells stably transfected with the GPCR, CCK Receptor A (CHO-CCKAR) (120), and transiently transfected epitope-tagged Gβ2 and KLF11. Similar to the adrenergic and secretin receptors, the CCKAR is a well-known regulator of metabolic functions (149). In agreement with previous reports, we find that Gβ2 localizes to the membrane and cytoplasm in unstimulated cells (Figure 3E). However, upon CCK treatment of the CHO-CCKAR cells, Gβ2 translocates to the nucleus where it localizes with KLF11, suggesting that under these conditions they function in close proximity (Figure 3E). Therefore, several physiological stimuli of distinct GPCRs (isoproterenol, secretin and CCK), which trigger dissociation of beta/gamma heterodimers from the trimeric G-protein alpha-s/beta/gamma complex at the plasma membrane to allow downstream signaling, result in interaction between Gβ and KLF11 in the nucleus of three different cell types. Combined, these results not only confirm our yeast two hybrid experiments, but more importantly indicate that activation of GPCRs induce the translocation of the Gβ2 subunit, which by forming a complex with KLF11, likely regulates responses (e.g. gene expression).

Metabolic gene networks requiring coupling of KLF11 to cofactors via the PRBD are deregulated by the A347S variant

Since the region of interaction between KLF11 and Gβ2 encompasses the A347S KLF11 variant, previously identified in a family of individuals affected by MODY-diabetes (109), we examined whether this naturally occurring variant interferes with binding between these two proteins. Immunoprecipitation of KLF11-A347S demonstrates that

this variant disrupts its interaction with Gβ2 (Figure 4A). Notably, this effect is isolated to binding between KLF11 and Gβ2, as the variant does not affect the interaction of KLF11 with its other known co-factors, Sin3a, HP1α or CBP/p300 (Figure 4B). These results reveal that the KLF11-A347S variant is useful to define whether protein-protein interactions influence the function that KLF11 has in the regulation of metabolic gene expression pathways. Thus, we performed a genome-wide query using a whole genome Affymetrix expression profile (RaGene 2.0) in primary pancreatic beta cells transduced with empty vector control, wild type KLF11, or the KLF11-A347S variant to define genes, and in particular, metabolic gene networks that are regulated by KLF11 but altered via the A347S variant that interrupts the function of the KLF11 TRD3 (Figure 5A). Clustering of all significantly altered probes ($p<0.05$) revealed distinct clusters of genes regulated by KLF11. Detailed statistical and bioinformatics analysis identified 2275 unique genes that significantly ($p<0.05$, fold change ± 2) associate with KLF11 expression, while 1205 genes significantly associate with KLF11-A347S expression.

Visually, clustering analysis reveals that in beta cells, expression of the KLF11-A347S variant results in a blunted activating or repressive effect compared to the effects of wild type KLF11, suggesting loss of function. To statistically confirm the disruptive effect of the A347 variant in beta cell function, Gene Ontology ANOVA analysis of the dataset was performed. KLF11 expression regulates clusters of genes related to the entire insulin response pathway, including glucose import, glycolysis, fatty acid metabolism, insulin secretion, calcium and potassium channels, MAPK activity, and PI3K activity, among others (Supplemental Table 2). Figure 5B represents the overall pattern of expression of 32 genes related to the regulation of insulin secretion following 48 hours of adenoviral mediated overexpression of KLF11, which mimics glucose exposure. A given ontological cluster, is must be noted, may be comprised of both positive and negative regulators of insulin secretion, as well as associated co-factors and upstream/downstream effectors. As such, the relative relationships between conditions hold more informative value than the direction of net effect. Compared to empty vector control, expression of KLF11 results in a net, significant ($p<0.05$) repression of the targets within the group.

In the presence of the KLF-A347S variant, however, the expression of these genes is significantly and globally disrupted. A small subset of other glucose and insulin related biological processes significantly deregulate din the presence of the A347S mutant are also shown (Figure 5B). These insights into the transcriptional effects of the MODY VII A347S variant are congruent with the defects in insulin secretion and response observed in the patient population.

To validate the regulatory effects of KLF11 and the KLF11-A347S variant observed in the genome-wide array, we performed qPCR on beta cells (Figure 4C) for a panel of insulin secretion and response pathway markers identified as significantly regulated from the global expression analysis Targets such as X and Y, related to X and Y function, respectively, are significantly de-repressed in the presence of the KLF11-A347 variant. This validation indicates the overall reliability of the array in identifying patterns of gene regulation by KLF11 or its variant. The effects of KLF11 and its A347S variant were also examined in human exocrine pancreatic epithelial cells (Supplemental Figure 1), which also revealed disruption of metabolic pathways associated to insulin response Additionally, we validated the regulatory effect of Gβ2 overexpression on a subset of identified KLF11-regulated targets in human exocrine pancreatic epithelial cells. Targets significantly regulated by KLF11 overexpression demonstrate similar activation or repression with overexpression of Gβ2 fused to a 3X SV40 nuclear localization signal (Figure 6A), indicating co-operative regulation of the targets by these two molecules. Lastly, using ChIP assay we confirm that human gene promoters may indeed be co-occupied by both KLF11 and Gβ within the first 5kb of their promoters (Figure 6B). Combined, these data support co-occupancy and regulation of gene targets by this newly identified KLF11-Gβ2 transcriptional complex.

DISCUSSION

While the strength of modern genetics is to discover association of genes and their variants to disease which high level of resolution, significantly less is known on how the protein products work to give rise to diseases. This has been the case with those genes

within the KLF family of transcriptional regulators, such as KLF11 and KLF14, which have been strongly associated to diabetes, obesity, and the insulin resistance/metabolic syndrome (108,109,150-152). For instance, very little is known as how these genes influence metabolic function. The current study provides mechanistic insights into how human diabetes-associated genetic mutations and variants that occur in these metabolic transcription factors, such as the KLF11 A347 variant, impair the regulation of metabolic gene networks. In addition, we provide biochemical and genetic evidence for the existence of a novel transcriptional regulatory pathway that appears to translate membrane-to-nuclear signals such as those imparted by signaling proteins, in particular the Gβ subunit, to the KLF11-mediated regulation of gene expression patterns relevant to metabolism and diabetes.

The functional importance of the region corresponding to the KLF11 TRD3 was suggested in 1999 through a careful mapping of the N-terminal domain of this protein (117), though the identity of a potential cofactor for this domain has remained elusive. Thus, the identification of the TRD3 as a protein interaction domain has increased our mechanistic understanding of how KLF11 regulates its functions and how the MODY VII A347S variant impairs it. The TRD3 functions as a protein-interaction module as demonstrated by yeast two hybrid, protein array, immunoprecipitation, and BiFC experiments. In addition, we confirm that the Gβ subunit of the heterotrimeric G protein, the most represented candidate from our yeast two hybrid screening, interacts with KLF11 via its TRD3 domain. Furthermore, we demonstrate that the MODY VII A347S variant disrupts binding to this protein. In addition, using a solid-phase in vitro binding assay (protein array), we show that the TRD3 domain can bind to several types of P-rich binding modules such as WWI, WWII, and SH3 domain containing proteins. Interestingly, some of these proteins, such as ITCH and Nedd4-like proteins, have been previously shown to bind to KLF transcription factors (135-137), at their C-terminal domain, providing a cross-validation of our experiments. While these regions show less homology to the corresponding cabut-like TRD3 domain, nevertheless they display the structural and functional characteristics that suggest that binding to members

of these families of proteins immediately upstream of the KLF DNA binding domain is a wider mechanism of regulation for KLF transcription factors.

The TRD3 binding proteins identified here are found in a wide variety of eukaryotic proteins that have a range of functions including adaptor/regulatory modules in signal transduction, pre-mRNA processing, cytoskeleton assembly, gene transcriptional activation and cell cycle control. Of further interest, Gβ is a WD40 repeat protein, a large protein family for which members are increasingly being recognized as key regulators of chromatin dynamics, transcription, and epigenetics (153,154). Among the diversity of WD40 domain function, a common theme has emerged that they collude as repeats to form β-propeller structures that acts as a platform for the stable or reversible association of binding partners. For instance, WD40 domains are overrepresented in HDAC and HMT complexes with well-known examples found in Groucho, Polycomb group protein EED, and WDR5 (153,154). The Gβ subunits, which have been here associated to KLF11, have also been reported to interact with other transcription factors, including fos and the glucorticoid receptor (GR) (145,146). In fact, analyses of the GR (a.a: 263-419) and fos (159-238), the regions from both of these transcription factors which are known to interact with Gβ, reveal remarkable similarities to the KLF11 TRD3, in that they contains proline-rich domains embedded within a sequence that is rich in hydrophobic (Φ) as well as flexible amino acid (G and Q). The WD-40 motif of Gβ may be essential for the GR interaction, as a Gβ structural analog RACK1 (receptor for activated C-kinase 1) also binds GR (155). This sequence feature adds to the similarities that Gβ works with this steroid receptor and KLF11. For example, activation of Gi-coupled somatostatin receptor in rat pituitary GH3 cells induces nuclear translocation, which works as a corepressor through the direct binding to the AF-2 domain of GR (145). Similarly, we show that the activation of the β2 adrenergic receptor by isoproterenol or the secretin receptor by secretin induces the formation of the KLF11-Gβ2 complex in the cell nucleus. Together, these results suggest that members of very distinct families of transcription factors can utilize the Gβ2 subunit in order to mediate their function in gene expression.

The observations described here are also of significant biological importance since the role of Gβ2 in the nucleus has remained elusive for many decades. Fortunately, the working model that emerges by the congruent association of all data described in this study as well as those available for the GR is that extracellular signals induce the binding of WD40 proteins to these transcription factors in order to regulate metabolic gene networks. Consequently, disruption of this mechanism should lead to alteration in the regulation of these networks, an effect that we observed experimentally and that is congruent with a role of MODY VII A347S in human diabetes.

CONCLUSION

In summary, the current study extends the available knowledge on KLF proteins and identifies a novel evolutionarily conserved protein-protein interaction domain, which is involved in the transcriptional regulation of metabolic gene networks. The biochemical contribution of this work lies in the extensive characterization of this domain, identification of Gβ as a binding partner, and the role of this domain in the regulation of metabolic gene networks. Lastly, the fact that mutations in this pathway, which are known to cause juvenile diabetes, impact on the regulation of metabolic gene networks should be taken into consideration for understanding potential molecular mechanisms that contribute to disease phenotypes.

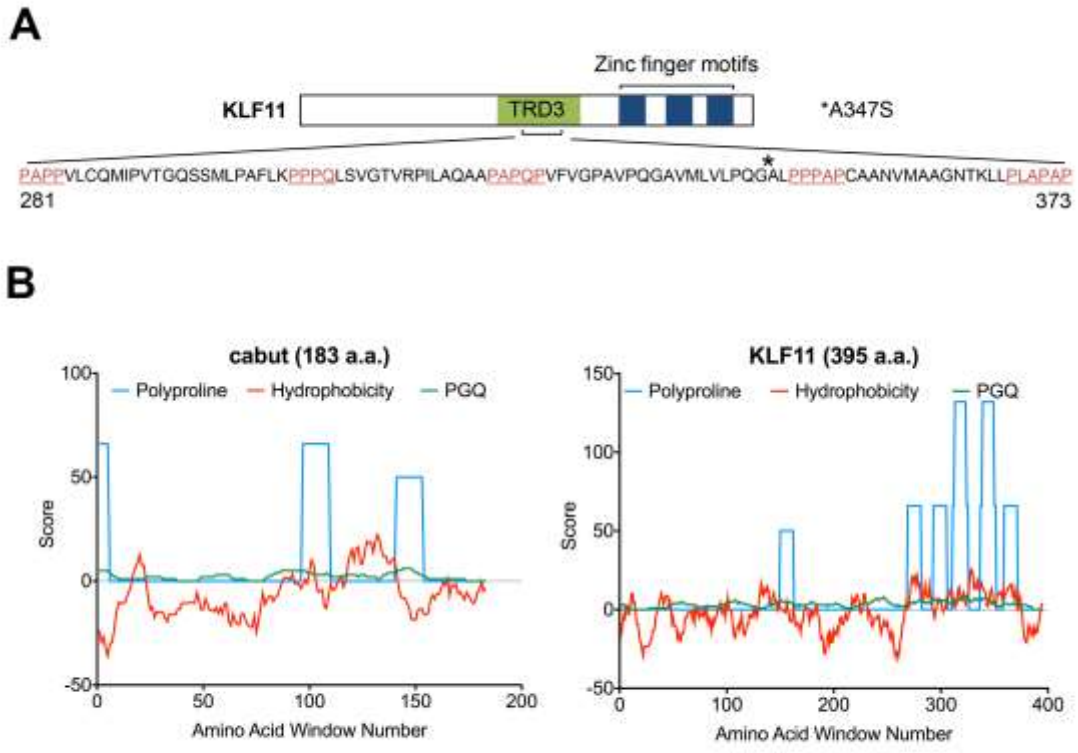


FIGURE 1. KLF11 A347S variant maps to a novel hydrophobic-glycine-glutamine-proline rich domain that is observed in the correponding region of its fly ortholog cabut. A. Diagramatic representation of the location and composition of a conserved protein-to-protein interaction domain present in the human ortholog of the Drosophila protein cabut, KLF11. Polyproline-rich regions are underlined and denoted in red. The A347S variant found in a subset of MODY VII patients is identified by an asterisk (*). B. Cabut and KLF11 were compared utilizing 15 amino acid sliding windows across the N-terminus of each protein prior to the conserved zinc finger domain and scoring for overall hydrophobicity using the Kyte-Doolite scale (20); polyproline helix conformations using experimental derived values (21); and overall proline, glycine, and glutamine content (PGQ index). The analysis reveals the presence of polyproline-enriched islands amid stretches of strongly hydrophobic residues.

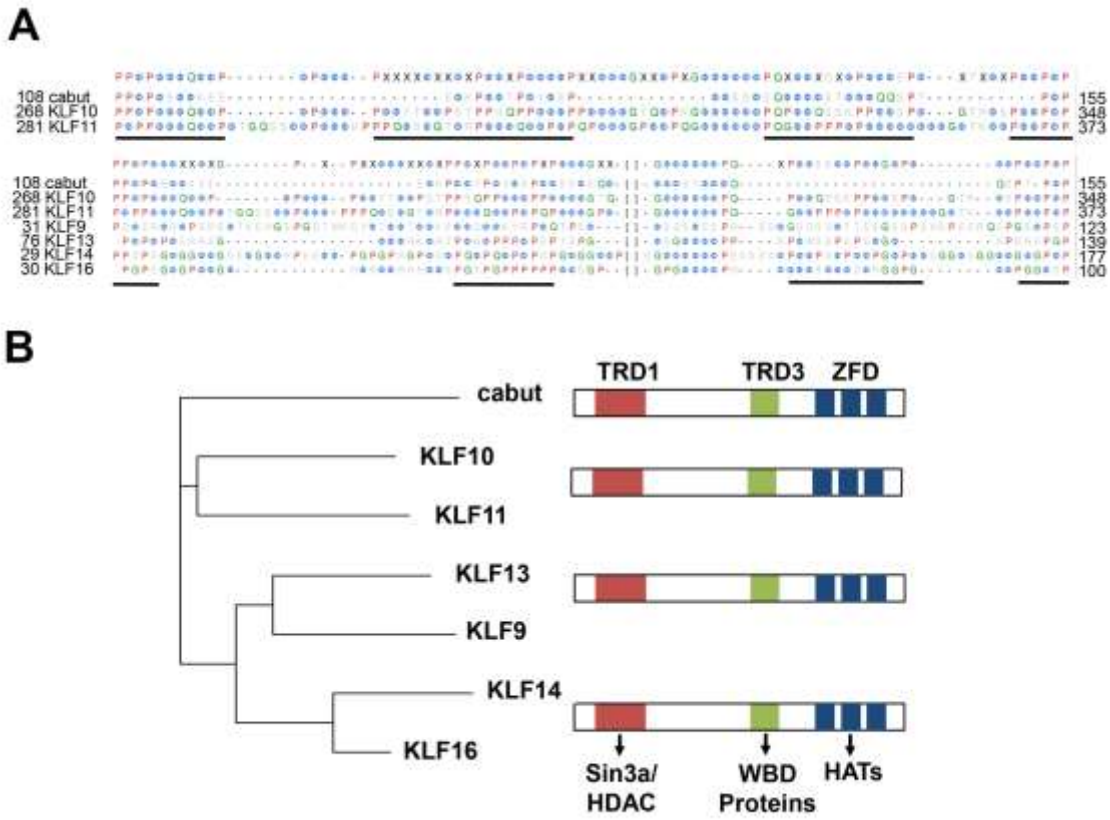


FIGURE 2. The KLF11 TRD3 protein domain is conserved in the cabut family of metabolic transcription factors. A. Comparison of biochemical-biophysical structure of the proline-rich domain between cabut and KLF10/11, its two closest orthologs, along with their consensus, is shown at the top. A similar comparision is shown below, which also includes the related BTEB subfamily of KLF proteins (KLF9, KLF13, KLF14, and KLF16) and the corresponding consensus. Regions with high polyproline content with either high hydrophobic character or high PQG index were deemed putative TRD3 domains. KLF11 contains four distinct PRM domains which are embedded within an environment of hydrophophic amino acids (Φ) and few flexible residues such as G and Q (green residues). Proline residues are denoted in red. Represented regions include amino acids 108-155 for cabut, 268-348 for KLF10, 281-373 for KLF11, 31-123 for KLF9, 76-139 for KLF13, 29-177 for KLF14 and 30-100 for KLF16. B. Evolutionary analysis reveals a high pressure for conservation of this protein domain from Drosophila cabut to human, with the closest orthologs that of the TIEG family of KLF proteins (KLF10 and KLF11). Members of the BTEB family of KLF proteins (KLF9, KLF13, KLF14, and KLF16) are also closely related.

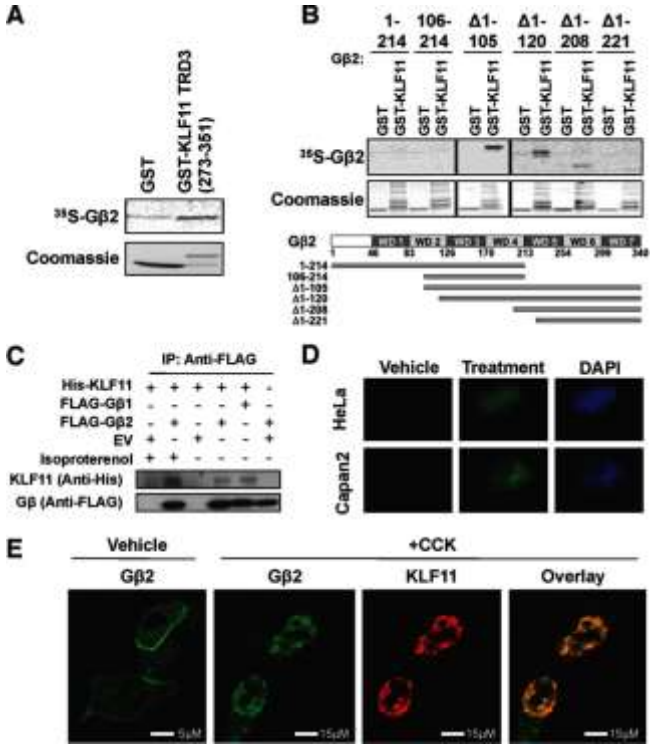


FIGURE 3. Membrane-to-Nuclear coupling of the KLF11 TRD3 with the Gβ subunit of the heterotrimeric G protein. A. Gβ2 binds the KLF11 TRD3 in vitro. GST-KLF11 TRD3 (a.a. 273-351) and GST alone were utilized for GST pulldown assays with 35S-labeled in vitro translated Gβ2. Phosphorimage results of 35S-Gβ2 binding (upper) and Coomassie staining of GST and GST-KLF11 TRD3 proteins as a loading control (lower) are shown. GST alone was included as a negative control. B.
KLF11 binding requires Gβ2 WD40 repeats 5-7 in vitro. GST-KLF11 full length and GST alone were utilized for GST pulldown assays with various 35S-labeled in vitro translated Gβ2 deletions. Interaction between KLF11 and Gβ2 was not detected with the N-terminal regions encompassing Gβ2Δ1-214) or Gβ2Δ106-214). Positive interaction was detected for Gβ2Δ1-105, Δ1-120, and Δ1-208. Gβ2Δ1-221, which disrupts WD40 repeat 5, also does not interact with KLF11, indicating that Gβ2 needs intact WD40 repeats 5-7 for binding. Upper panel shows phosphorimage results of 35S-Gβ2 binding (upper) and Coomassie staining of GST and GST-KLF11 proteins as a loading control (lower). Lower panel depicts a cartoon representation of Gβ2 with location of its WD40 repeats, as well as the various deletions used for mapping of the Gβ2-KLF11 interaction. C. Gβ binds to KLF11 in cells, which is enhanced by activation of GPCR signaling. Upon co-transfection of epitope-tagged Gβ2 and KLF11 in HeLa cells, Flag-tagged Gβ2 immunoprecipitates His-tagged KLF11. Empty vector (EV) was utilized as a negative control in combination with either Flag-Gβ2 or His-KLF11, separately. Immunoprecipitation of Flag-tagged Gβ1 also detected a complex with KLF11. Experiments with Flag-tagged Gβ2 or EV and His-tagged KLF11 were repeated in the presence of isoproterenol, which detected a more robust interaction. D. BiFC of KLF11 and Gβ2 demonstrates nuclear interaction. Cotransfection of N-terminal EYFP protein (EYFP(1)) fused to KLF11 with the C-terminal EYFP (EYFP(2)) fused to Gβ2 in both HeLa and Capan2 cells was performed and subsequently, cells were stimulated with isoproterenol or secretin, respectively. While control treatment (vehicle) does not show positive interaction, activation of GPCRs in both cell types (treatment) demonstrates interaction in the nucleus through fluorescence reconstitution and nuclear DAPI co-stain. E. GPCR stimulation translocates Gβ2 from the cell membrane to the nucleus where it co-localizes with KLF11. CHO-CCKAR cells were transiently transfected epitope-tagged Gβ2 and KLF11. Without stimulation of CCKAR (vehicle), Gβ2 localizes to the membrane and cytoplasm. Upon CCK treatment (+ CCK), Gβ2 (green) translocates to the nucleus where it localizes with KLF11 (red). Overlay shows extensive co-localization (yellow) of these two proteins, confirming their close proximity.

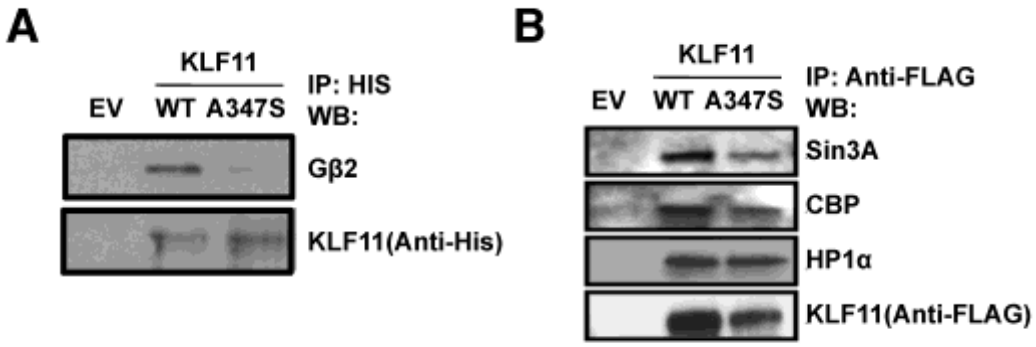


FIGURE 4. The KLF11 A347S MODY VII variant disrupts its ability to bind the Gβ subunit. A. The KLF11-Gβ interaction in cells is altered by the A347S variant. Immunoprecipitation of His-tagged KLF11 wild type (WT) demonstrates binding with endogenous Gβ2 in Panc1 cells. However, immunoprecipitation of His-tagged KLF11 A347S shows that this variant disrupts its interaction with Gβ2. EV was used as a negative control. B. KLF11 A347S does not impair binding to other KLF11 co-factors. Although KLF11 A347S does not interact with Gβ2, immunoprecipitation of the His-tagged variant does not affect interaction with other known KLF11 co-factors, namely Sin3a, HP1gamma and CBP/p300.

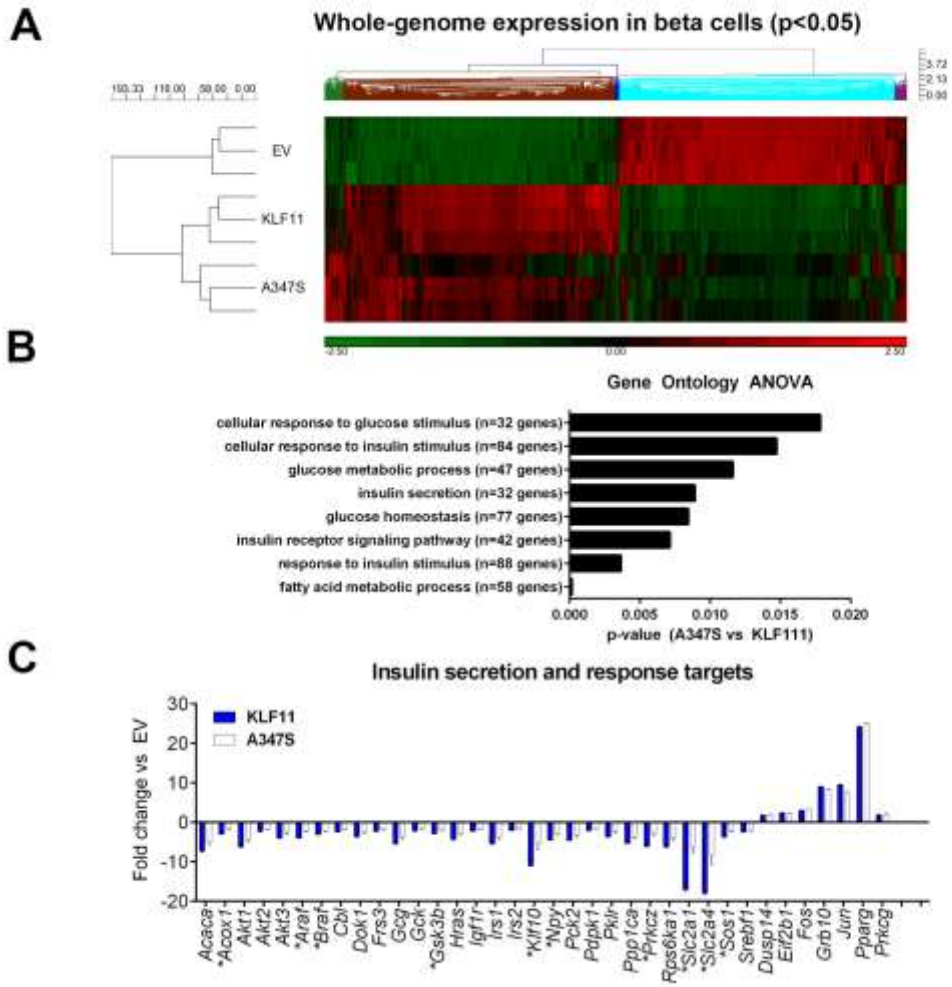


FIGURE 5. The KLF11 A347S MODY VII variant deregulates metabolic gene networks that require coupling of KLF11 to cofactors via the TRD3. A. Genome-wide query using a whole genome Affymetrix expression profile (ReGene 2.0) in pancreatic beta cells transduced with empty vector, KLF11, or the A347S variant. Significant ($p<0.05$) genes were clustered to compare the effects of the A347S variant compared to wild type. Detailed statistical and bioinformatics analysis identified unique genes that significantly associate with KLF11 expression ($p<0.05$, log2 fold change +/-2). This analysis reveals that the targets significantly repressed or activated under wild type stimulation are de-repressed or de-activated in the presence of the A347S variant. B. Gene ontology ANOVA analysis of KLF11 and A347S targets reveals a distinct and enrichment of insulin secretion and response pathway ontological clusters that are significantly disrupted in the presence of the A347S variant. C. qPCR validation of a panel of insulin secretion and response genes again reveals the disruptive effect of the A347S variant identified by global expression profiling in Part A. Asterisks equals gene targets where the A347S variant deregulation is significant compared to wild type ($p<0.05$).

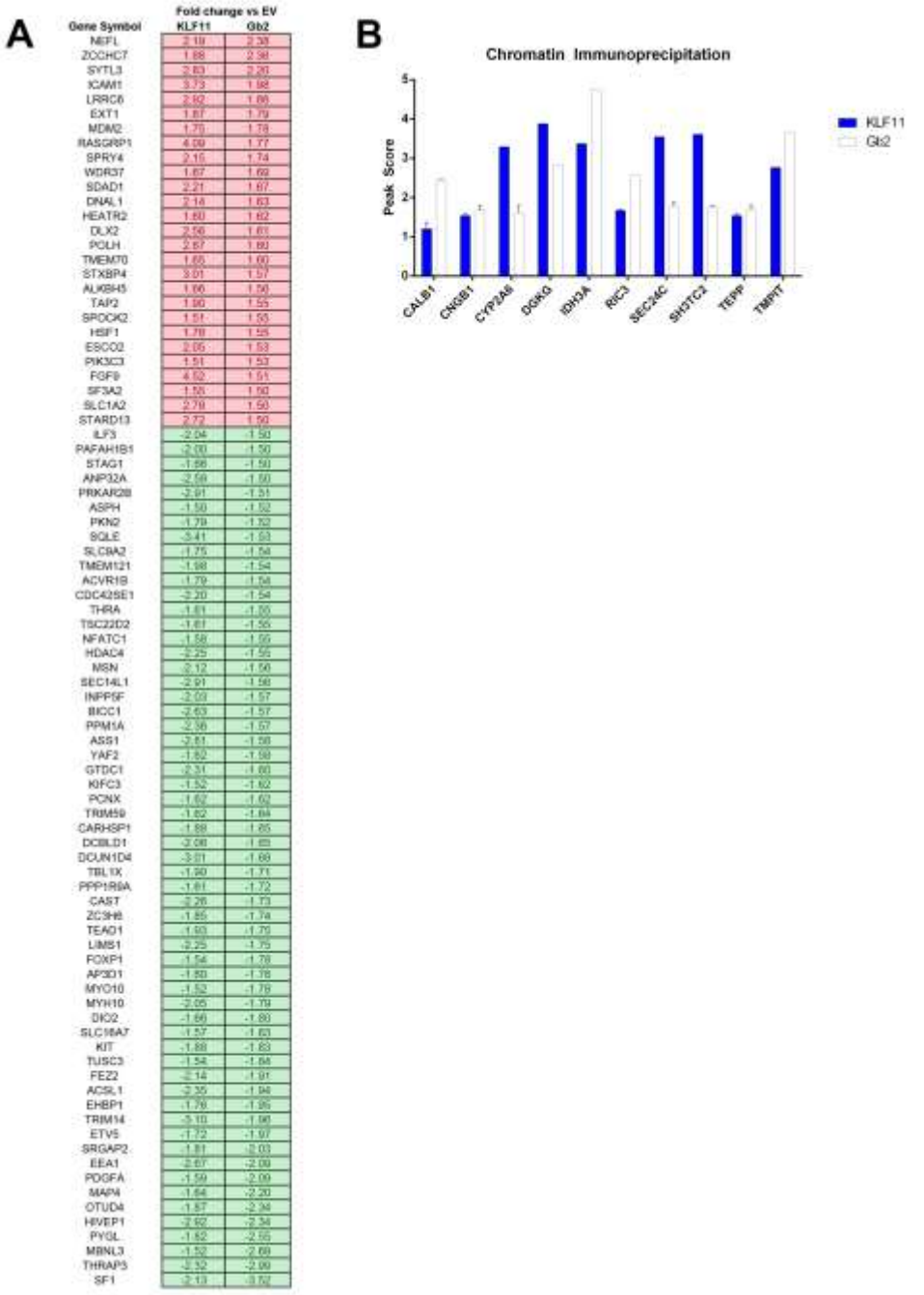


FIGURE 6. Regulatory effect of Gβ2 overexpression on a subset of identified KLF11 target genes. A. Targets significantly regulated by KLF11 demonstrate similar activation or repression in the presence of Gβ2 overexpression with a 3X SV40 nuclear localization signal. Significance was assumed at $p<0.05$, log2 fold change +/-1.5 for either KLF11 vs EV or Gβ2 vs EV. B. Co-occupation of KLF11 and Gβ at human gene promoters. ChIP was utilized to examine co-occupation of KLF11 and Gbeta on regions of human gene promoters up to 5kb upstream of the transcriptional start site in epithelial cells. False discovery rate is presented here as error on reported peak scores.

Supplemental Figure 1: Comparative effects of KLF11, A347S, and Gβ2 expression in human epithelial cells. A. Genome-wide query using a whole genome Affymetrix expression profile (Human Gene 1.0 ST) in exocrine epithelial cells transduced with empty vector, KLF11, or the A347S variant. Significant ($p<0.05$) genes were clustered to compare the effects of the A347S variant compared to wild type. B. Gene ontology ANOVA analysis of glucose metabolism and insulin response pathways reveals disruption by the A347S variant, a subset of which are presented here.

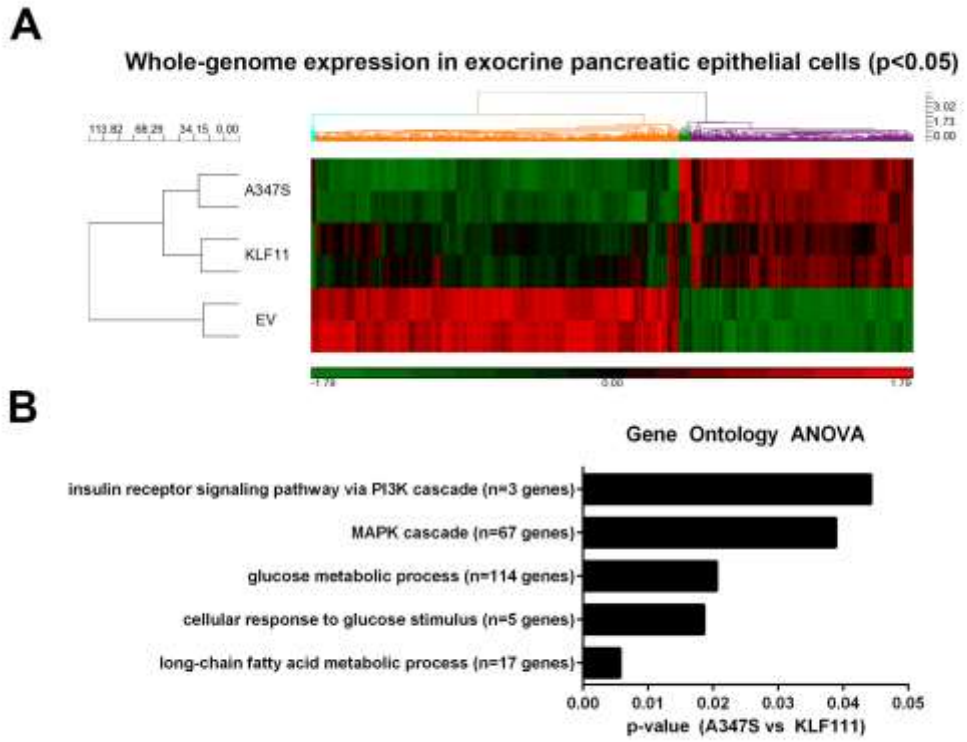
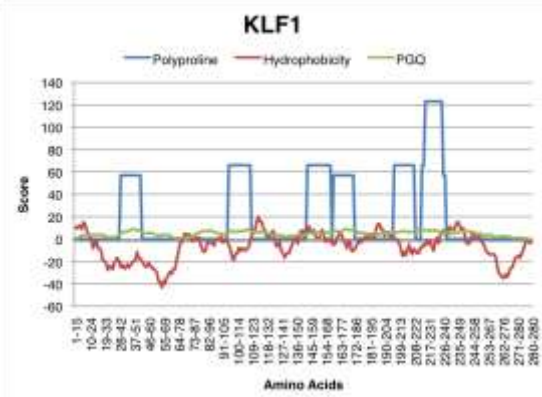
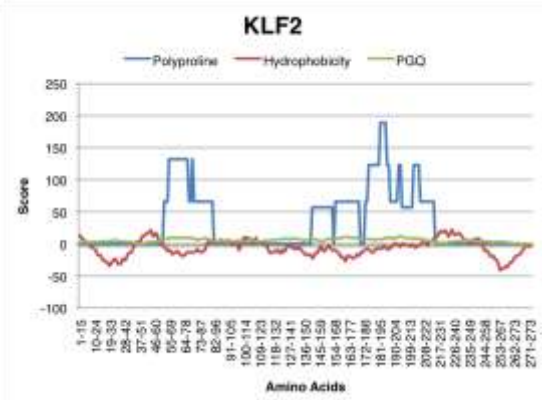
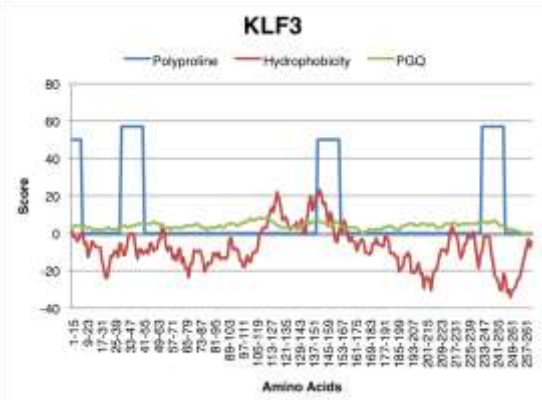


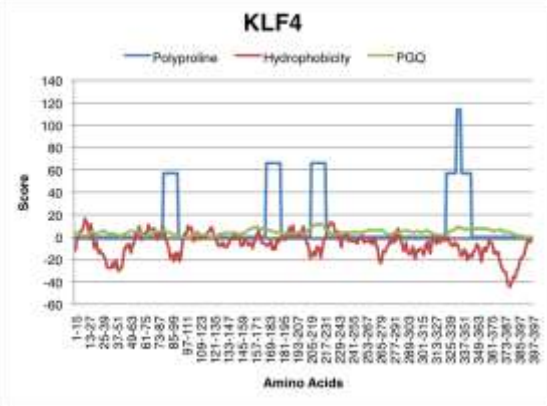
TABLE 1. List of potential KLF11 TRD3-interacting proteins identified by yeast two hybrid and protein array.

Gene Symbol	Gene Name	Assay	Unigene
AHCY	Adenosylhomocysteinase	Y2H	Hs.388004
AMPH	Amphiphysin	SH3	Hs.592182
ARHGAP12	Rho GTPase activating protein 12	WWII	Hs.499264
AZGP1	Alpha-2-glycoprotein 1, zinc-binding	Y2H	Hs.546239
AZGP1	Alpha-2-glycoprotein 1, zinc-binding	Y2H	Hs.546239
BTK	Bruton agammaglobulinemia tyrosine kinase	SH3	Hs.159494
CSK	C-src tyrosine kinase	SH3	Hs.77793
CTBP1	C-terminal binding protein 1	Y2H	Hs.208597
CTTN	Cortactin	SH3	Hs.596164
EEF2	Eukaryotic translation elongation factor 2	Y2H	Hs.515070
EML2	Echinoderm microtubule associated protein like 2	Y2H	Hs.24178
ENPP3	Ectonucleotide pyrophosphatase/phosphodiesterase 3	Y2H	Hs.486489
FBXL5	F-box and leucine-rich repeat protein 5	Y2H	Hs.643433
FYN	FYN oncogene related to SRC, FGR, YES	SH3	Hs.390567
GAA	Glucosidase, alpha; acid	Y2H	Hs.1437
GAS7	Growth arrest specific 7	WWI	Hs.462214
GGT5	Gamma-glutamyltransferase 5	Y2H	Hs.437156
GNB1	guanine nucleotide binding protein (G protein), beta polypeptide 1	Y2H	Hs.430425
GNB1	guanine nucleotide binding protein (G protein), beta polypeptide 1	Y2H	Hs.430425
GNB2	guanine nucleotide binding protein, beta polypeptide 2	Y2H	Hs.185172
GNB2	guanine nucleotide binding protein, beta polypeptide 2	Y2H	Hs.185172
GNB2	guanine nucleotide binding protein, beta polypeptide 2	Y2H	Hs.185172
GNB2	guanine nucleotide binding protein, beta polypeptide 2	Y2H	Hs.185172
GNB2	guanine nucleotide binding protein, beta polypeptide 2	Y2H	Hs.185172
GNB2	guanine nucleotide binding protein, beta polypeptide 2	Y2H	Hs.185172
HYPB/SETD2	Huntingtin interacting protein HYPB, SET domain containing 2	WWII	Hs.517941
IFITM1	Interferon induced transmembrane protein 1	Y2H	Hs.458414
IRF1	Interferon regulatory factor 1	Y2H	Hs.436061
ITCH	Itchy E3 ubiquitin protein ligase	WWII	Hs.632272
ITK	IL2-inducible T-cell kinase	SH3	Hs.558348
LCK	Lymphocyte-specific protein tyrosine kinase p56 LCK	SH3	Hs.470627
MAGI3	Membrane associated guanylate kinase-related MAGI3	WWI	Hs.486189
METTL20	Methyltransferase like 20	Y2H	Hs.740628
NEDD4	NEDD4, E3 ubiquitin protein ligase	WWI	Hs.1565
NEDD4L	NEDD4-like, E3 ubiquitin protein ligase	WWI	Hs.185677
PCSK2	Proprotein convertase subtilisin/kexin type 2	Y2H	Hs.315186
PEX13	Peroxisomal biogenesis factor 13	SH3	Hs.161377
POLR2G	Polymerase (RNA) II (DNA directed) polypeptide G	Y2H	Hs.14839
POLR2G	Polymerase (RNA) II (DNA directed) polypeptide G	Y2H	Hs.14839
POLR2G	Polymerase (RNA) II (DNA directed) polypeptide G	Y2H	Hs.14839
RBM47	RNA binding motif protein 47	Y2H	Hs.518727
RPS3	Ribosomal protein S3	Y2H	Hs.546286
SMURF1	SMAD specific E3 ubiquitin protein ligase 1	WWI	Hs.189329
SMURF2	SMAD specific E3 ubiquitin protein ligase 2	WWI	Hs.741194
SRCAP	Snf2-related CREBBP activator protein	Y2H	Hs.647334
TAP1	Transporter 1, ATP-binding cassette, sub-family B	Y2H	Hs.352018
WDR6	WD repeat domain 6	Y2H	Hs.654815
WWP2	WW domain containing E3 ubiquitin protein ligase 2	WWI	Hs.408458
ZBTB4	Zinc finger and BTB domain containing 4	Y2H	Hs.741193

Y2H=yeast two hybrid; SH3= SH3 protein array; WWI= WWI protein array; WWII= WWII protein array

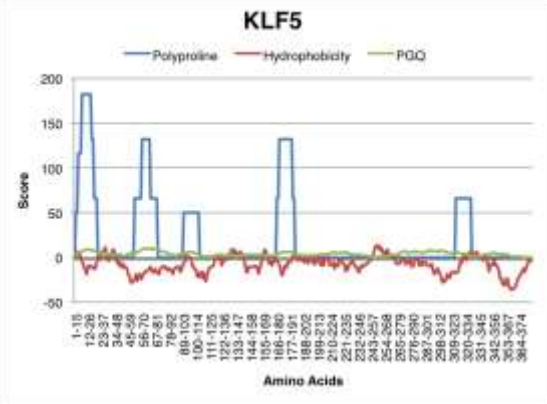
Supplementary Table 1: Analysis of the entire subset of KLF proteins predicts the presence of the TRD3 in a large number of KLF proteins outside of the cabut subfamily. Examination of hydrophobic-glycine-glutamine-proline enrichment across all members of the KLF family. Φ denotes hydrophobic residues.

	**KLF1 (1-280aa)** ΦΦTΦETΦΦPSΦSTΦTΦΦGPΦPDTQDDΦΦK ΦΦRSEEΦQDΦGPGPPDPTEPPΦHΦKSEDQ PGEEEDDERGΦDΦTΦDΦDΦΦΦTNΦSGPEP GGΦPQTΦΦΦΦPSEΦSGΦQYPPPPETΦGΦY ΦGGPGΦΦΦGΦΦGSEDHSGΦΦRPΦΦRΦR ΦPDΦΦΦGPΦΦΦPΦPΦPEPKΦΦΦΦQPΦYP GPGΦGSSGGYΦPRTGΦSΦPΦΦSGΦPYGΦ ΦSGYPΦΦYPΦPQYQGHΦQΦΦRGΦQGPΦP GPΦTSPSΦΦSΦΦGPGTΦGTGΦGGTΦEDPG ΦΦΦETΦPSKRGRRSΦΦRKRQΦΦHT
	**KLF2 (1-273aa)** ΦΦΦSEPΦΦPSΦSTΦΦSPΦRERGΦQERΦP RΦEPESGGTDDDΦNSΦΦDΦΦΦSΦGΦDGΦ GΦEΦΦPEPPPPPPPPPΦΦYYPEPGΦPPPYSΦ PΦGGΦΦSEΦΦRPEΦDΦPΦGPΦΦHGRΦΦΦ ΦPPGRΦΦKΦEPPEΦDGGGGYGΦΦPGΦTR GPRGΦKREGΦPGPΦΦSΦΦRGPGGRPPPPP DTPPΦSPDGPΦRΦPΦPGPRΦSΦPPPΦGGP GΦGΦPGPGΦHYΦPPΦPPΦΦGΦΦDDΦΦΦΦ ΦΦΦΦGΦΦPPΦΦRGΦΦTPPΦSPΦEΦΦEΦK PKRGRRSΦPRKRTΦTHT
	**KLF3 (1-261aa)** ΦΦΦΦDPΦPΦKQEΦΦDPΦSΦSYPSNYΦES ΦKPNKYGΦΦYSTPΦPEKΦΦQTPEGΦSHGΦ QΦEPΦDΦTΦNKRSSPPSΦGNSPSSΦKΦPS SHRRΦSPGΦSΦPSSSPPΦKKYSPPSPGΦQP ΦGΦPΦSΦPPΦΦΦΦΦΦSRHGΦRSPGΦΦPΦ ΦQPΦΦΦQPΦPΦΦYTSHΦQQPΦΦΦSΦSEE ΦENSSSSΦQΦPΦΦESYEKPΦSQKKΦKΦEP GΦEPQRTDYYPEEΦSPPΦΦNSΦSPPQΦΦΦ QENHPSΦΦΦQPGKRPΦPΦESPDTQRKRRΦ HR



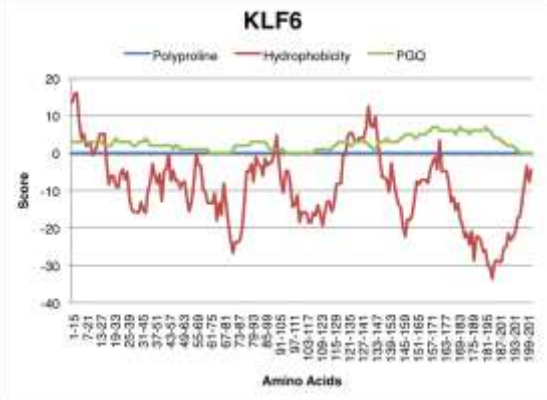
KLF4 (1-397aa)

ΦRQPPGESDΦΦΦSDΦΦΦPSΦSTΦΦSGPΦ
GREKTΦRQΦGΦPNNRΦREEΦSHΦKRΦPPΦ
ΦPGRPYDΦΦΦΦTΦΦTDΦESGGΦGΦΦΦGG
SNΦΦPΦPRRETEEΦNDΦΦDΦDΦΦΦSNSΦT
HPPESΦΦΦTΦSSSΦSΦSSSSSPSSSGPΦSΦ
PSTΦSΦTYPΦRΦGNDPGΦΦPGGTGGGΦΦY
GRESΦPPPTΦPΦNΦΦDΦNDΦSPSGGΦΦΦ
EΦΦRPEΦDPΦYΦPPQQPQPPGGGΦΦGKΦ
ΦΦKΦSΦSΦPGSEYGSPSΦΦSΦSKGSPDGS
HPΦΦΦΦPYNGGPPRTΦPKΦKQEΦΦSSΦTH
ΦGΦGPPΦSNGHRPΦΦHDΦPΦGRQΦPSRTT
PTΦGΦEEΦΦSSRDΦHPΦΦPΦPPGΦHPHPG
PNYPSΦΦPDQΦQPQΦPPΦHYQEΦΦPPGSΦ
ΦPEEPKPKRGRRSΦPRKRTΦTHT



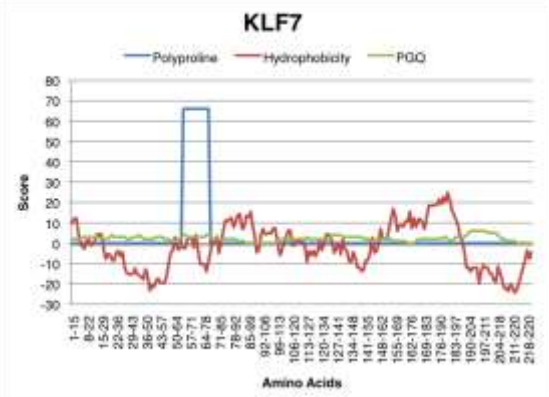
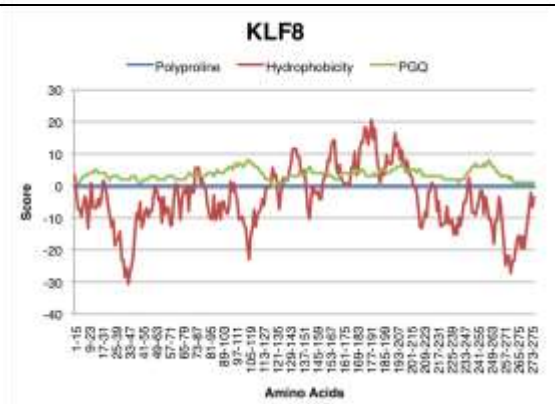
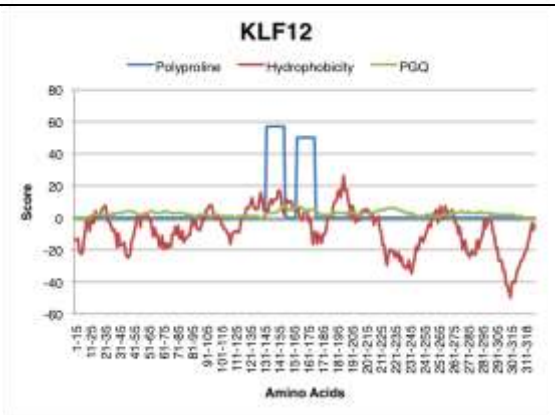
KLF5 (1-374aa)

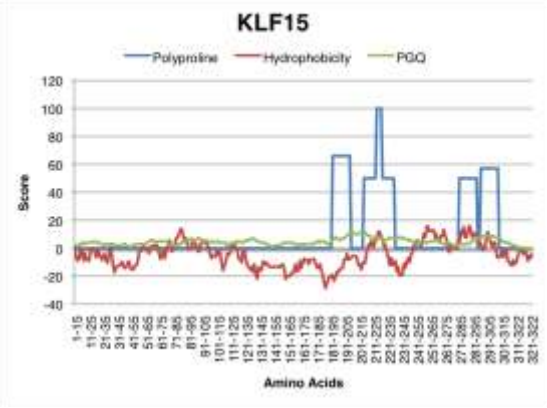
ΦΦTRΦΦSΦSΦRΦGPΦPQPPΦPQDEPΦΦΦ
QΦKPΦΦGΦΦNPΦRDΦΦΦΦPGEEΦKHΦHH
RPQΦQPΦPΦQΦPQPΦQPPΦTGPRΦPPEDΦ
ΦQTRΦEΦEKYΦTPQΦPPΦPΦΦPEHKKYRR
DSΦSΦΦDQΦΦTDTEGΦPYSΦNΦNΦΦΦPD
ΦTHΦRTGΦYKSQRPΦΦTHΦKTEPΦΦΦΦSH
QSETTΦPPPΦPTQΦΦPEΦTSΦΦSSHQTΦΦ
PEΦNNΦΦΦKQEΦPTPDΦHΦSΦPTQQGHΦY
QΦΦNTPDΦDΦPSSTNQTΦΦΦDTΦNΦSΦSΦ
ΦΦΦGΦNTHTSΦΦPQTΦΦKQΦQGΦPPΦTY
TΦPSQΦΦPQQΦTYΦPPSPPSSEPGSPDRQ
ΦEΦΦQNΦTPPPSYΦΦTΦΦSKΦΦΦHNPNΦ
PTTΦPΦNSQNΦQPΦRYNRRSNPDΦEKRRΦ
HY



KLF6 (1-201aa)

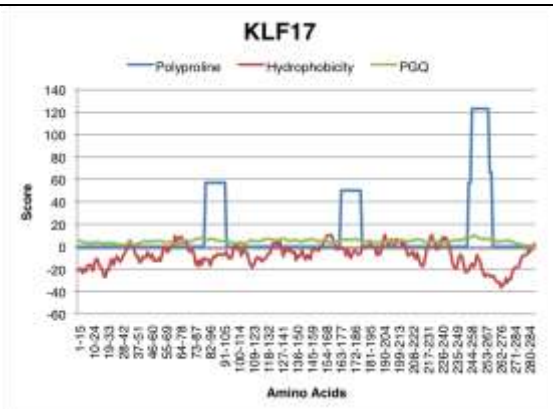
ΦDΦΦPΦΦSΦΦQEΦQΦΦHETGYΦSΦΦPSΦ
EEYΦQQTΦΦEΦERYΦQSEPΦYΦSΦSEΦKΦ
DSQEDΦΦTKΦΦΦΦREKKEESEΦKΦSSSPP
EDTΦΦSPSΦΦYNΦETNSΦNSDΦSSESSDSS
EEΦSPTΦKΦTSDPΦGEΦΦΦSSGKΦSSSΦT
STPPSSPEΦSREPSQΦΦGΦΦPGEΦPSPGK
ΦRSGTSGKPGDKGNGDΦSPDGRRRΦHR

	**KLF7 (1-220aa)** ΦDΦΦΦSYSΦΦQEΦQΦΦHDTGYΦSΦΦPSΦ EETΦQQTΦΦEΦERYΦQTEPRRΦSETΦGED ΦDΦΦΦHΦSPPPΦΦEESΦRRΦDPΦΦΦPΦE ΦΦΦΦEKSSΦΦDΦΦΦSRDKΦΦSETΦΦSΦQ PΦSSSΦDSYTΦΦNQΦQΦNΦΦTSΦTPPSSP EΦSRHΦΦKTSQTΦSΦΦDGTΦTΦKΦΦΦKK ΦΦΦSSΦKΦGGΦΦTΦΦΦΦΦTΦΦGΦΦKSG QSDSDQGGΦGΦEΦΦPENKKRΦHR
	**KLF8 (1-275aa)** ΦΦDΦDKΦΦNNΦEΦQΦNSEGGSΦQΦΦKQ ΦTΦSΦRNRDPPEΦEYRSNΦTSPTΦΦDΦNP ΦENPΦΦΦNDΦKΦEPPEEΦΦΦSDΦSΦPQΦ EPΦDΦSΦHKPKΦPΦQPΦSΦΦQΦPΦRPPK PQSSPQTΦΦΦSTSTSDΦSTSΦNΦPTΦΦTPG SΦΦTSSQSTGSQQΦΦHΦΦHTΦPSΦSΦPNK ΦGGΦKTΦPΦΦΦQSΦPΦΦYTTΦPΦDGGPΦ ΦΦTΦPΦΦGGDGKNΦGSΦKΦDPTSΦSPΦE ΦPSDSEESTΦESGSSΦΦQSΦQGΦQQEPΦΦ ΦΦQΦQGEESΦDΦKRRRΦHQ
	**KLF12 (1-318aaa)** ΦNΦHΦKRKTΦKNΦNTΦENRΦΦΦΦDGΦPΦ ΦRΦKTEΦΦESEQGSPNΦHNYPDΦEΦΦPΦΦ ΦNNΦKGEPPEDSΦSΦDHΦQTQTEPΦDΦSΦ NKΦRTSPTΦΦSSSPΦSΦTΦSΦSSPSSTSTS SSSSSRΦΦSSPTΦΦTSΦSSΦSSSSTΦΦTPG PΦΦΦSΦSGΦGGQQΦΦHΦΦHPΦPPSSPΦN ΦQSNKΦSHΦHRΦPΦΦΦQSΦPΦΦYTΦΦRS PGNΦNNTΦΦΦPΦΦEDGRGHGKΦQΦDPRG ΦSPRQSKSDSDDDDΦPNΦTΦDSΦNETGST ΦΦSΦΦRΦΦQEΦHPSPΦSRΦRGNRΦNNQK ΦPΦSΦSPΦSΦESTRRQRRSESPDSRKRRΦ HR



KLF15 (1-322aa)

ΦΦDHΦΦPΦDENΦSSPKΦPΦGYΦGDRΦΦG
RRΦYHΦΦPSPΦSEDDSDΦSSPΦSΦSSPDS
QΦΦΦSΦYGGGΦGTESQDSΦΦDΦΦΦSQΦT
ΦGSGGGSGSSΦGΦSSGPΦΦΦGPΦRRΦΦΦ
PΦKGEHΦΦΦPEΦPΦGDPDDΦPRPΦQPTΦE
EΦEEΦΦEENΦEPGΦKEΦPEGNSKDΦDΦΦS
QΦSΦGPHKSHΦHPGSSGRERΦSPPPGGΦS
ΦGGΦQGPGGGPTPDGPΦPΦΦΦQΦQPΦPΦ
KQESGTGPΦSPGQΦPENΦKΦΦQΦΦΦNΦQ
GQTΦΦΦΦPQΦΦPSSNΦNΦPSKΦΦRΦΦPΦ
PΦΦΦKPΦGSGPΦGPGPΦGΦΦΦGQKΦPKN
PΦΦEΦΦKΦHK



KLF17 (1-284aa)

ΦYGRPQΦEΦEQEΦGEΦSRΦQΦΦHQΦΦQD
NENSΦPΦΦNΦSSSSGSSGΦHTSΦNQGΦPS
ΦQHΦPHSΦEΦΦGSPΦΦSΦEΦPGQNΦNEG
GPQΦSΦPΦPERGΦSYΦPQΦTΦTPSRΦΦYΦ
QRΦSPPQQEΦTΦΦSGPQΦΦPΦGEPNΦPR
ΦΦRPΦGGNΦRΦPPNGΦPΦSΦSTGΦPΦΦS
HTGNPPΦPYPGΦSTΦPSDETΦΦGPTΦPSTE
ΦQΦΦΦPSΦΦQΦΦPPQDΦHDΦGΦPPΦESQ
SΦΦΦΦGSQDSΦΦSQPDSQEGPΦΦPEQPG
PΦPQTΦEKNSRPQEGTGRRGSSEΦRPYΦ

Supplementary Table 2: KLF11 overexpression regulates the insulin response pathway. Affymetrix gene expression data were examined for significant enrichment of genes clustered into a range of insulin response biological, molecular, and cellular functions using the Partek GO ANOVA function. The effects of the A347S variant are also examined as well as the significance of the disruption between the wild type to the variant. The top 100 results are included here.

GO Description	p-value(KLF11 vs. EV)	Fold-Change(KLF11 vs. EV)	p-value(V347 vs. EV)	Fold-Change(V347 vs. EV)
mesenchymal cell development	5.01E-10	-1.39052	7.46E-09	-1.26432
G1/S transition of mitotic cell cycle	1.28E-09	-1.20204	6.01E-09	-1.16354
dendrite development	1.53E-09	-1.29662	1.19E-08	-1.22204
embryonic morphogenesis	7.24E-09	-1.32189	2.16E-08	-1.27511
protein K48-linked deubiquitination	8.21E-09	-1.40555	2.44E-08	-1.34517
rRNA transcription	1.05E-08	1.52551	3.84E-08	1.4308
L-glutamate transport	1.14E-08	1.43933	3.71E-09	1.52134
regulation of protein localization	1.55E-08	-1.21315	1.50E-07	-1.15567
intracellular protein kinase cascade	1.95E-08	-1.26037	2.25E-07	-1.18449
protein deubiquitination	2.19E-08	-1.46812	7.02E-08	-1.39243
spindle organization	2.39E-08	-1.42709	3.62E-08	-1.40124
response to aluminum ion	2.71E-08	-1.35497	5.41E-07	-1.23004
SCF-dependent proteasomal ubiquitin-dependent protein catabolic process	2.72E-08	-1.32973	2.29E-07	-1.24247
positive regulation of microtubule polymerization	3.34E-08	-1.44954	3.26E-07	-1.31973
renal system process	3.54E-08	-1.4023	1.55E-07	-1.32336
oxaloacetate metabolic process	3.59E-08	-1.37133	4.43E-06	-1.18461
ear development	3.62E-08	-1.49717	3.14E-07	-1.35814
response to external stimulus	3.89E-08	-1.44702	1.57E-07	-1.36236
positive regulation of insulin secretion involved in cellular response to glucose stimulus	3.89E-08	-1.24569	6.41E-08	-1.22893
regulation of cell-matrix adhesion	4.14E-08	-1.44108	2.22E-07	-1.34292
homocysteine metabolic process	4.31E-08	-1.31413	1.85E-06	-1.18332
endoplasmic reticulum calcium ion homeostasis	4.61E-08	-1.38695	3.04E-07	-1.293

striated muscle cell differentiation	4.72E-08	-1.33691	3.31E-07	-1.25389
regulation of circadian rhythm	4.73E-08	-1.3389	6.47E-07	-1.23201
calcium ion-dependent exocytosis	4.91E-08	-1.50221	9.29E-07	-1.32168
nerve growth factor receptor signaling pathway	5.35E-08	-1.39293	5.78E-07	-1.27657
regulation of smoothened signaling pathway	5.40E-08	1.26274	1.15E-07	1.23592
positive regulation of protein serine/threonine kinase activity	5.63E-08	-1.40625	4.72E-07	-1.29638
protein localization to synapse	6.14E-08	-1.33427	5.09E-07	-1.24588
lipid biosynthetic process	6.55E-08	-1.53319	4.42E-07	-1.39724
regulation of viral genome replication	7.24E-08	-1.32982	3.21E-07	-1.26552
protein dephosphorylation	7.60E-08	-1.21527	1.37E-06	-1.14368
mitotic spindle organization	8.65E-08	-1.28237	4.43E-07	-1.22348
inositol phosphate dephosphorylation	9.54E-08	-1.14343	3.14E-07	-1.12194
response to inorganic substance	1.00E-07	-1.14719	6.71E-06	-1.08273
cellular response to glucose starvation	1.04E-07	-1.25547	4.27E-06	-1.15076
negative regulation of neuron death	1.04E-07	-1.13357	2.17E-06	-1.08836
fatty acid beta-oxidation	1.12E-07	-1.12029	5.91E-06	-1.07016
B cell proliferation	1.13E-07	-1.23073	9.01E-07	-1.17225
positive regulation of protein ubiquitination involved in ubiquitin-dependent protein catabolic process	1.18E-07	-1.11053	6.05E-08	-1.12089
positive regulation of exocytosis	1.21E-07	-1.40459	2.39E-06	-1.25974
positive regulation of nitric-oxide synthase biosynthetic process	1.24E-07	-1.36115	9.06E-07	-1.26957
autophagic vacuole assembly	1.24E-07	1.24965	8.09E-07	1.19137
regulation of exocytosis	1.46E-07	-1.52926	1.67E-06	-1.36335
positive regulation of insulin secretion	1.48E-07	-1.29859	1.16E-06	-1.22196
mRNA polyadenylation	1.49E-07	1.49056	1.73E-06	1.33739
response to folic acid	1.53E-07	-1.20648	3.60E-07	-1.18313
regulation of developmental	1.55E-07	-1.33828	5.09E-07	-1.28414

pigmentation				
decidualization	1.56E-07	1.18291	2.73E-07	1.16919
response to hydrogen peroxide	1.58E-07	-1.12074	8.50E-06	-1.07007
regulation of cell morphogenesis	1.63E-07	-1.3678	9.23E-07	-1.28469
phosphorylation	1.68E-07	-1.20536	4.23E-06	-1.13061
circadian regulation of gene expression	1.70E-07	-1.39288	2.67E-06	-1.26096
T cell homeostasis	1.79E-07	-1.22832	1.65E-06	-1.1668
myoblast fusion	1.79E-07	-1.43636	1.54E-06	-1.3153
peptidyl-threonine phosphorylation	1.84E-07	-1.26313	7.73E-07	-1.21421
positive regulation of protein ubiquitination	1.97E-07	-1.14452	1.30E-06	-1.11154
regulation of alternative mRNA splicing, via spliceosome	2.10E-07	-1.11256	2.53E-07	-1.10971
intrinsic apoptotic signaling pathway in response to oxidative stress	2.22E-07	-1.2897	2.86E-06	-1.20024
neuron-neuron synaptic transmission	2.24E-07	-1.18996	1.40E-06	-1.14705
V(D)J recombination	2.26E-07	-1.22659	4.77E-07	-1.20379
oocyte development	2.34E-07	-1.61139	1.14E-06	-1.47473
glycogen biosynthetic process	2.34E-07	-1.34121	4.62E-06	-1.22021
positive regulation of DNA repair	2.39E-07	1.14253	1.42E-06	1.11156
histone H4-K5 acetylation	2.42E-07	-1.1583	1.04E-07	-1.17786
histone H4-K8 acetylation	2.42E-07	-1.1583	1.04E-07	-1.17786
regulation of protein stability	2.50E-07	-1.15563	2.58E-07	-1.15496
antigen processing and presentation of exogenous peptide antigen via MHC class II	2.51E-07	1.24185	5.36E-06	1.15638
fatty acid metabolic process	2.60E-07	-1.16833	1.76E-05	-1.09347
B cell lineage commitment	2.69E-07	-1.32839	8.32E-07	-1.27809
growth	2.77E-07	-1.16397	1.10E-05	-1.09826
positive regulation of erythrocyte differentiation	2.87E-07	-1.26914	6.29E-06	-1.17264
positive regulation of interleukin-8 production	2.90E-07	1.2234	8.83E-07	1.19078
cell projection	3.04E-07	1.15778	0.000447915	1.05562

organization				
regulation of I-kappaB kinase/NF-kappaB cascade	3.06E-07	1.19668	1.78E-06	1.1536
regulation of apoptotic process	3.10E-07	-1.10279	3.36E-06	-1.07439
glucose metabolic process	3.18E-07	-1.25251	1.98E-06	-1.19421
negative regulation of ossification	3.49E-07	-1.21558	1.17E-05	-1.1309
oocyte maturation	3.99E-07	-1.3428	5.51E-06	-1.23259
dopamine metabolic process	3.99E-07	-1.30377	1.27E-05	-1.18324
epidermal growth factor receptor signaling pathway	4.03E-07	-1.14529	2.30E-05	-1.08273
positive regulation of peptidyl-serine phosphorylation	4.09E-07	-1.20445	1.93E-06	-1.16424
oxidative phosphorylation	4.36E-07	-1.33135	3.98E-06	-1.23913
dopamine receptor signaling pathway	4.79E-07	-1.3151	2.62E-06	-1.24546
positive regulation of GTPase activity	4.84E-07	-1.25783	4.18E-06	-1.18895
regulation of ARF GTPase activity	4.84E-07	-1.3397	5.01E-06	-1.24035
positive regulation of neuroblast proliferation	4.91E-07	-1.23322	5.73E-06	-1.16412
filopodium assembly	4.92E-07	-1.34275	3.79E-06	-1.25323
negative regulation of peptidyl-serine phosphorylation	5.11E-07	1.42863	2.75E-05	1.23402
response to virus	5.19E-07	1.07692	1.24E-06	1.06839
regulation of protein kinase A signaling cascade	5.36E-07	-1.23193	4.13E-06	-1.17316
response to sucrose stimulus	5.57E-07	1.25002	5.85E-06	1.17815
cellular respiration	5.59E-07	1.42991	1.84E-05	1.25273
response to organic nitrogen	5.73E-07	-1.09883	0.000133857	-1.0462
negative regulation of protein complex assembly	5.76E-07	-1.27837	6.61E-06	-1.1952
chromatin modification	5.99E-07	-1.07438	5.10E-07	-1.076
axonogenesis	6.02E-07	-1.21985	4.69E-06	-1.16403
Golgi organization	6.23E-07	-1.13535	1.35E-06	-1.12161
cholesterol biosynthetic process	6.26E-07	-1.73814	2.52E-05	-1.40319

CHAPTER 3:

Single and Combinatorial Chromatin Coupling Events Underlies the Function of Transcript Factor Krüppel-like Factor 11 in the Regulation of Gene Networks

Ezequiel Calvo[1], Adrienne Grzenda[2], Gwen Lomberk[2], Angela Mathison[2], Juan Iovanna[3], and Raul Urrutia[2,4,5]

[1] Molecular Endocrinology and Oncology Research Center, CHUL Research Center, Quebec, Canada; [2] Laboratory of Epigenetics and Chromatin Dynamics, Mayo Clinic, Rochester, MN 55905, USA; [3] INSERM U.624, Stress Cellulaire, 163 Avenue de Luminy, Case 915, Parc Scientifique et Technologique de Luminy, 13288 Marseille Cedex 9, France; [4] Translational Epigenomics Program, Center for Individualized Medicine (CIM), Mayo Clinic, Rochester, MN 55905, USA; [5] Departments of Medicine, Physiology and Biochemistry, Mayo Clinic, Rochester, MN 55905, USA

Manuscript in preparation.

SUMMARY

Krüppel-like factors (KLFs) are a group of master regulators of gene expression conserved from flies to human. However, scant information is available on either the mechanisms or functional impact of the coupling of KLF proteins to chromatin remodeling machines, a deterministic step in transcriptional regulation. In the current study, we have used genome-wide analyses of chromatin immunoprecipitation (ChIP-on-Chip) and Affymetrix-based expression profiling to gain insight into how the KLF11, a human transcription factor involved in tumor suppression and metabolic diseases, works by coupling to three co-factor groups: the Sin3-histone deacetylase system, WD40-domain containing proteins, and HP1-histone methyltransferase system. Our results reveal that KLF11 regulates distinct gene networks involved in metabolism and growth by using single or combinatorial coupling events. This study, the first of its type for any KLF protein, reveals that interactions with multiple chromatin systems are required for the full gene regulatory function of these proteins.

INTRODUCTION

The traditional strategy for studying the role of transcription factors in gene regulation relies on either small or large-scale expression analysis following overexpression, somatic knockdown, or germ line deletion. While these approaches permit the identification of gene networks regulated by transcription factors, little information is gathered about the chromatin coupling events that ultimately drive gene expression. This gap in knowledge is unfortunate as transcription factors are often modular proteins armed with multiple sites for potential interactions with chromatin cofactors. Thus, knocking out a transcription factor will, in theory, disrupt all chromatin-coupling events. For this reason, it is important to dissect whether define transcriptional regulators use single or combinatorial coupling events to regulate different gene expression networks.

Such knowledge is particularly important for proteins in which dysfunction contributes to disease since impairment in each of the coupling events can have potentially differential impact, serving as a cause of variation in symptoms as well as progression. Moreover, with the development of novel chromatin-centric pharmacology, different drugs can partially inactivate certain gene networks regulated by a transcription factor while leaving others intact. Therefore, dissecting individual and combinatorial chromatin coupling events has both biological and medical relevance.

Consequently, the current study focuses on studying KLF11 as a model system for addressing these important questions. KLF11 is a well characterized human disease causing gene, which couples to several chromatin partners (156). Alterations in KLF11, originally discovered by its role in growth regulation, causes juvenile (MODY7) and neonatal diabetes (Ins-331 mutation) diabetes (118,157,158). KLF11 is an inducible gene, responsive to a large variety of growth regulatory and metabolic stimuli, and functions in the nucleus to regulate gene expression by coupling to distinct chromatin partners. Upon stimuli, KLF11 binds to promoters containing the consensus CCCCGC/CCCCAC

sequences via its three C-terminal zinc finger domains and, through well-characterized protein-protein interaction modules in its N-terminus, differentially recruits chromatin partners such as the Sin3-HDAC complex (159), WD40 proteins (160), and the HP1-HMT system (161). Thus, through the use of these domains, KLF11 can translate environmental signals into distinct programs of gene expressions, which remain being defined in detail.

In the current study, we have employed a combination of genome-wide ChIP-on-Chip and expression profiles, to reconstruct both direct and indirect effects of KLF11 on the regulation of different gene networks. Furthermore, we used specific site-directed mutants known to disrupt specifically each of the chromatin coupling domains to generate expression profiles as a genome-wide reporter for the identification of the subset of genes regulated by KLF11 in response to binding to each of its chromatin cofactors. The results of these experiments demonstrate that a single chromatin machine can either regulate KLF11-regulated genes while others require interactions with more than one system. Accordingly, we provide analyses of the distinct gene networks regulated by this transcription factor and its partners. This type of comprehensive genome-wide analysis has never been performed for any member of the KLF family and, when applied to KLF11, demonstrates that these proteins behaves a modular with the ability of using single chromatin proteins or a combination of them to achieve their gene regulatory function.

RESULTS

Distinct gene expression networks are regulated by the differential coupling of KLF11 to individual chromatin partners.

Recent studies support a model whereby KLF11 functions by binding to GC-rich sites within promoters of different gene networks involved in the regulation of metabolism and cell growth (118,157,158,161-166). This data is in agreement with the biological role of this transcription factors in cancer and diabetes. However, how binding of KLF11 to these promoters regulates gene expression remains to be fully understood. Recent data

demonstrate that the KLF11 protein behaves as a scaffold for recruiting different chromatin cofactors via distinct structural motifs (Figure 1A). The Sin3 Interacting Domain (SID) between amino acids 22-40 enables coupling of KLF11 to the Sin3/HDAC system (91,165-172). Introduction of proline residues at amino acids 29-30 interrupts the association between KLF11 and the Sin3 scaffold protein, which includes interactions with HDAC1/2. The region between 281 and 373 amino acids is a proline-rich domain that couples with a variety of WD40 proteins, including G-protein coupled receptors (173). The A347S mutation observed in MODY7 (neonatal diabetes) falls within this domain and has been demonstrated to decouple KLF11 from novel transcription factor Gβ2. The zinc finger domain has been demonstrated to interact with histone acetyltransferases p300 and CBP (158,174,175). Finally, the region between 483 and 487 dictates interaction between KLF11 and HP1α (161). Deletion of this C-terminal portion decouples HP1α from histone methyltransferases SUV39 and G9a, leading to defects in metabolism and tumor suppression. Currently, it is unknown whether these systems work in isolation or in a cooperative manner to regulate gene expression.

To address this question, we first preformed genome-wide promoter binding analysis of wild type KLF11. 404 identifiable gene promoters were bound by KLF11. Gene ontological analysis of these targets reveals association to a large number of metabolism processes, including many already characterized as regulated by KLF11, including insulin regulation (118,158,166) and Akt/TOR signaling (165) (Figure 1B). Next, we performed genome-wide expression profiles using specific mutants for these sites which have been shown to specifically uncouple KLF11 from each of these chromatin remodeling proteins, namely EAPP (Sin3/HDAC), A347S (WD40 proteins), and Δ486 (HP1α/HMT). Of the 404 gene targets identified by chromatin immunoprecipitation, 19% (n=75) display alteration in the presence of the wild type KLF11 ($p<0.05$). Clustering analysis reveals that the expression pattern of these gene targets is disrupted in the presence of one or all of the described mutants. Thus, we conclude that KLF11 mediated gene expression may be disrupted by decoupling of the transcription factor

from its chromatin co-factors in a combinatorial or singular fashion.

Distinct gene networks are regulated by the combinatorial coupling of KLF11 to chromatin partners.

Next, we investigated which gene targets are uniquely regulated by each of these pathways. Statistical analyses of each expression profiles show that, indeed, when compared with the empty vector control condition, each mutation induces changes in gene expression which are either specific to each chromatin system or similar to the wild type. Thus, comparison of the genes significantly affected by each condition (p-value <0.05 and fold change ± 1.5) shows that 97% of transcripts that are modulated by the wild type KLF11 are also regulated by any of the three mutations, with the A347S and Δ486 mutations sharing the largest number of genes with the wild type (Figure 2A). In fact, of the 801 genes modulated by the Δ486 mutation, 93.9% are also modulated by the A347S mutation. Finally, the three mutants share with KLF11 only 11 different genes, which likely require an intact coupling of this transcription factor to all the chromatin proteins involved in its function (Figure 2B). We observe that the A347S mutation shows the largest number of significantly regulated transcripts (n=708) upon the decoupling of KLF11 from WD40 proteins (Figure 3A). By contrast, the subsets of genes uniquely regulated by the Δ486 or EAPP mutants are much smaller, 44 and 21 genes, respectively (Figures 3B-C).

The Δ486 mutation uniquely modulates only 44 genes (5.5%) of the 801 genes significantly modulated under expression of the mutant. Almost a half of all Δ486 modulated genes (n=390) are common to the A347S mutant (Figure 3D) and the other half are common to both A347S and wild type KLF11 (Figure 3E). Interestingly, the Δ486 and the A347S mutants have, in general, a similar direction of modulation, although with varying degrees of signal intensity. On the other hand, the A347S and Δ486 modulated genes are almost always completely reverted by the EAPP mutation, in which the signals are constantly proximal to empty vectors values.

Under the same conditions, the A347S mutant, which associates to the development of human juvenile diabetes, induces changes in 1521 genes. 46.5 % of the modulated genes (708/1521) are specific to this mutation, while approximately half of the genes (48%) are common to the Δ486 mutant and 43.3% are common to both A347S and wild type KLF11 (352/423, Figure 3F). Finally, 7% of the genes (n=50) are common to wild type KLF11, and only 15 common to the EAPP mutant. The EAPP mutation, which disrupts coupling to the Sin3-HDAC complex, reverses the expression of most genes induced by the wild type KLF11 or by the other two mutations.

Identification of KLF11-regulated genes for which expression is independent of known chromatin coupling events.

Subsequently, we examined genes for which expression might be regulated by KLF11 in a manner that does not involved the chromatin remodeling machines examined above. A small percentage of genes are still significantly modulated uniquely by wild type KLF11 (Figure 3G). Ontological classification demonstrate that these 14 genes are associated with diabetes (ALMS1P; RAET1L; GYS1; RBM47; EFR3B), cancer (RPL27A; SPAM1; PTCH1; ZNF277; USP35; CDC86), or metabolism (C3orf15; AOC2; GOT1). Taken together, these data indicate that the vast majority of KLF11's biological processes are dependent on its interaction with chromatin regulators although a repertoire of chromatin-independent functions is expected.

Pathway reconstruction of combinatorial KLF11-regulated chromatin pathways by ontological approach.

Gene Ontology (GO) enrichment analysis of the genes modulated by wild type KLF11 and each of the three mutations show various common biological processes (Figure 4A). Indeed, the wild type shares 60% (34/57) of the biological process with at least one of the mutants. One of the biological processes, lipid cellular metabolic process, is enriched in wild type KLF11 and all three mutants (Supplemental Table 1). Among the most highly enriched biological processes, 14 terms are common to wild type and the A347S and

Δ486 mutations. Variations of the intensity of signal, the number of genes implicated, and the score of GO enrichment for these 14 biological processes are shown in Supplemental Table 1. Interestingly, wild type KLF11 and mutants conserve several biological processes related to anabolism or catabolism of lipids, amino acids, and glycoconjugates. Other common groups include redox processes of protein homodimerization, regulation of the activity of protein kinases, and response to the estrogen.

Two of the wild type KLF11 biological processes, the G2/M transition during DNA damage checkpoint and cilium assembly are also conserved in the A347S mutation. There are 16 biological processes exclusively conserved between wild type KLF11 and the Δ486 mutant. These include cellular metabolic processes such as aldehyde and nitrogen compound metabolism, isocitrate and 2-oxoglutarate metabolism, glycosaminoglycan catabolic processes, and fatty acid beta-oxidation.

Interestingly, the A347S and Δ486 mutation share more than a quarter of biological processes (n=39) that are absent in wild type KLF11 and the EAPP mutant. Between these 39 GO terms, we found enrichment of terms associated to the monitoring of the DNA transcription, RNA polymerase I transcription promoter regulation, and epigenetic control of gene expression such as histone H3 acetylation, chromatin remodeling, protein acetylation and deacetylation, and regulation of phosphorylation. Wild type KLF11 and EAPP share only two GO-terms in common: branched chain amino acid family catabolic processes and negative regulation of TOR and signaling cascade.

Overall, 22 of 57 (38.6%) of the biological processes enriched in the wild type KLF11 are not present any of the three mutants. Among the 22 most enriched biological process found only in the wild type KLF11, we observed several metabolic processes, such as glucose, carbohydrate, and very long-chain fatty acid metabolism, oxaloacetate metabolism, fatty acid homeostasis, and oligosaccharide metabolism. Top processes associated with wild type KLF11 expression are shown in Figure 4B and Supplemental Table 2. Other specific mechanisms of wild type KLF11 lost by mutation are two

mechanisms associated with gene expression (transcription from the RNA polymerase III promoter and mRNA capping) and two biological processes associated with the integrity of the DNA (the DNA repair and the response to DNA damage stimulus). Finally three important biological processes (e.g., the mitotic cell cycle, the negative regulation of insulin-like growth factor receptor signaling pathway, and the negative regulation of epithelial cell proliferation) are also lost by the mutants. From these data we conclude that the majority of KLF11's function, particularly in regulating gene expression and metabolism, are dictated by the combinatorial association between the transcription factor and a variety of chromatin coupling systems. The high degree in over suggests a form of "epigenetic redundancy" to ensure the proper regulation of these gene targets and processes.

Pathway reconstruction of singular KLF11-regulated chromatin pathways by ontological approach.

Of the mutants, A347S displays the highest number of unique biological process that are not enriched in the wild type (n=93). Top processes associated with A347S mutant expression are shown in Figure 4C and Supplemental Table 2. Interestingly, almost 30% of these biological processes are associated with the structure and maintenance of DNA or telomeres (Supplemental Table 3, List 1), processes related to phosphorylation and dephosphorylation, deacetylation, methylation, and demethylation (Supplemental Table 3, List 2), and processes of transcriptional regulation by poly-II and mRNA transport (Supplemental Table 3, List 3).

In turn, the Δ486 mutant exclusively enriches 45 biological processes that are mediated by the wild type KLF11 or in other mutants. Top processes associated with Δ486 expression are shown in Figure 4C and Supplemental Table 2. Unlike the A347S mutant, mutation Δ486 is particularly enriched in processes associated with metabolic and biosynthetic processes, protein export-import, process and energy protein modifications associated processes (Supplemental Table 4, List 1). Other groups of processes specifically enriched in this mutant are associated with proliferation and epithelial cell

survival (Supplemental Table 4, List 2).

The EAPP mutant did not display robust gene expression changes within the experiment. However, utilizing a more permissive cutoff yields a list of 53 biological categories, although each only contains a single gene. We find enrichment of biological processes involving nitric oxide transport, catabolic processes, developmental morphogenesis functions, among others. Top processes associated with EAPP expression are shown in Figure 4E and Supplemental Table 2. From these data, we are able to identify singular KLF11 chromatin coupling events dictating a number of critical processes, particularly in management of cell cycle control and DNA replication, ones that inherently require a more specific and more finely regulated degree of control to execute their programs under strict spatial and temporal constraints.

Identification of canonical signaling pathways mediated by KLF11 and mutants.

Subsequently, we wished to further assess the relatedness the gene targets significantly modulated by KLF11 and its mutants utilizing a semantic-based approach and the Ingenuity Global Canonical Pathways (GPC) algorithm. The input dataset of significantly regulated gene targets was compared against the IPA canonical pathways, which are curated from published literature. The significance of the associations between the data set and a given canonical pathway is determined by the ratio of the number of genes mapping to the pathway divided by the total number of pathway genes. A p-value is calculated using Fischer's Exact Test determining the probability that the association between the data set and the pathway occurs by chance alone. GPC analysis of genes modulated by the wild type KLF11 and the three mutants shows a variety of significantly enriched pathways (p-value <0.05). Indeed, the wild type KLF11 shares about 71% (30/42) of its pathways with at least one of the mutants (Figure 5A). The most important intersection of common pathways is observed between wild type KLF11 and mutants A347S and Δ486. Indeed, 90% (38/42) of these pathways are also enriched in at least one of these two mutants. Only two pathways are shared between the wild and the EAPP mutant.

Nevertheless, all genes implicated in these pathways are modulated differently by wild type KLF11 and each one of the mutants. The expression of these genes significantly regulated by wild type KLF11 and the A347S and Δ486 mutants are almost always modulated in the same sense and clustered together (Figure 5B). In contrast, the deletion of the EAPP site produces a complete reversion of the signals, and the levels of expression of EAPP's genes became close to the empty vector level, as evidence by its clustering with empty vector apart from the wild type KLF11 and other mutants.

Twelve of the 42 significantly enriched pathways are conserved only in the wild type KLF11. These normal functions of the wild type KLF11 are potentially lost by the three studied mutations. The most enriched of these pathways includes the nucleotide excision repair pathway and the alpha-adrenergic signaling pathways. Interestingly, of all the remaining pathways, 80% (9/11) of them are metabolic biosynthetic pathways.

Mutant A347S possesses the most exclusive pathways. Indeed, 28 pathways are enriched in the A347S mutant compared to wild type KLF11 and other mutants (Figure 5C). Interestingly, over 30% (9/28) of these pathways are signaling pathways, including, EGF, PDGR, p53, RANK and reelin signaling, and glucocorticoid receptor signaling. Furthermore, pathways involved in diabetes and pancreatic adenocarcinoma are also present. These pathways may explain the massive number of genes differentially modulated by this mutation (n=1521 genes) and suggests that this mutant could induce new potential unexpected and significant physiopathological changes.

86% of the enriched pathways in the Δ486 mutant are shared with the wild type KLF11 or the other mutants. Only four new pathways are specifically enriched by the Δ486 mutation (Figure 5D). Interestingly, the pathways are implicated in the degradation of biogenic amines (an important field associated to physiopathology of depression and metabolism of drug of abuse). Finally, the EAPP mutant influences only six pathways exclusively (Figure 5E), also primarily related to the degradation of biogenic amines.

Taken together, we conclude that KLF11 chromatin-coupled gene targets delineate into specific, well ordered pathways, consistent with KLF11's position as a master transcriptional regulator that mediates large cascades of other transcription factors and regulators. The effects of KLF11 chromatin coupling are therefore not incidental to its function but a primary means by which to execute gene activation or repression.

Reconstruction of downstream biological and disease networks mediated by KLF11 and mutants.

We next explored the degree of connectivity between KLF11 and mutant regulated genes and their pathobiological associations. Networks of significantly regulated genes were algorithmically generated based on their connectivity and assigned a score that encapsulates the relevance of the generated network is to the original list of focus genes. All edges are supported by at least one literature reference of direct physical, transcriptional and enzymatic interactions or from canonical information stored in the Ingenuity Pathways Knowledge Base. A right-tailed Fisher's Exact Test was used to calculate the p-value for networks and a threshold of $p<0.05$ used to determine significance. A functional analysis of a network then determined the biological functions and/or diseases that are most significant to the genes contained within the network (Figure 6).

Of the identified processes, only three were common to wild type KLF11 and all three mutants: cell death and survival, hereditary disorders, and molecular transport mechanisms. Wild type KLF11 possessed the most unique networks (n=10) and included a number of characterized biological and disease networks, including cellular growth and proliferation (176), endocrine disorders, reproductive system development and function (177), and biliary hyperplasia, among others. The A347S mutant uniquely associates to 6 networks, including ones linked to auditory disease, cellular compromise, and organismal development. The Δ486 mutant uniquely associated to 8 networks. Among the Δ486 mediated networks are significant associations to gastrointestinal disease, connective tissue development, and immune presentation and response. Together, these data

recapitulate the repertoire of known KLF11-mediated functions and uncover a number of new potential functional and disease associates for future studies.

Identification of upstream regulators of KLF11-chromatin coupled pathways.

Finally, we wished to identify the upstream transcriptional regulators using Ingenuity's Upstream Regulator (UR) analysis. The UR analytic is based on prior knowledge of the expected effects of transcriptional regulators and their target genes stored in the Ingenuity Public Knowledge Base. The algorithm examines how many known targets of each transcriptional regulator are present in the list of significantly regulated targets and compares direction of expression to that expected from previously published data. For each potential regulator, an overlap p-value using Fisher's Exact Test and an activation z-score was computed, the latter of which determines the activation state. A threshold p-value of <0.01 and activation score of +/- 2 was considered significant for our purposes. The end result was the identification of upstream regulators of KLF11 or its mutants that permits the generation of plausible signaling cascades mediated by these upstream regulators through KLF11 chromatin coupling.

The results of the UR analysis are presented in Supplemental Table 5. Wild type KLF11 displayed two significant upstream inhibitory regulators, PPARG and HNF1A, both of which have previously been implicated in the regulation of the transcription factor or as a co-regulator (Figure 7A) (178). The A347S mutant, consistent with the gene level, pathway, and network analyses is the most disruptive, possessing 9 activating upstream regulators and 23 inhibitory upstream regulators. Interestingly, EZH2, a histone methyltransferases, is identified as potential upstream activator. Previous reports have identified antagonistic relationship between KLF10 and the functions of EZH2/Polycomb Repressive Complex (175). This data hints to other important chromatin couplings to KLF11 that remain to be completely elucidated. Figure 7B shows the genes affected in the presence of the mutant when 486 is inhibited by SREBF1 or JUN and ATF4. For the 486 mutant, 6 upstream activators were identified and 15 upstream inhibitors, including SREBF2 and HIF1A (Figure 7C). Although the EAPP did not display any significant

upstream regulators as defined by our criteria, a number of other KLF proteins (KLF3, 7, 13, 1) were significant by p-value, suggesting that inter-association of KLF proteins is also critical to their function. Although a number of the regulators are common to the wild type KLF11 and the A347S and Δ486 mutants, indicating that mutation does not necessarily impair upstream regulation or co-regulation, the activation or inhibition of these proteins regulates a distinct subset of endpoint genes. The absence of wild type KLF11 targets and acquisition of new, inappropriate targets in the presence of mutation places KLF11 chromatin coupling events at the center of the transcription factor's function or dysfunction.

In summary, here we have identified the precise interplay between KLF11 and three of its chromatin coupling systems, namely Sin3/HDAC, HP1/HMTs, and WD40 proteins, on a genome scale. The functional repertoire of KLF11 appears to be dictated by both combinatorial effects between these systems and singular chromatin coupling events. Redundancy, particularly in critical biological processes such as metabolism and organismal development, occurs between systems to ensure preservation of critical functions. On the other hand, singular chromatin coupling events exist to ensure precision in processes requiring strict spatial and temporal parameters to avoid inappropriate signaling.

DISCUSSION

Until recently, the relationship between chromatin and transcription factors appeared largely passive. However, the evidence that overexpression of four transcription factors—Oct4, Sox2, m-Myc, and KLF4—is sufficient to reprogram fully differentiated fibroblasts into pluripotent ES-like cells suggests an active role for transcription factors in the epigenetic regulation of chromatin state (8). KLF proteins, in particular, have demonstrated a multitude of feed-forward effects on epigenetic regulation of gene expression. Here we propose KLF11 as a model for understanding the critical role of sequence-specific factors in epigenetic regulation of a multitude of pathobiological processes.

KLF proteins deliver epigenetic information to gene promoters through three primary mechanisms: (i) sensing and translating environmental stimuli into a program of gene expression, (ii) sequence-specific targeting of chromatin remodeling complexes to gene promoters, and (iii) transactivation of other transcription factors to assist in the regulation of large networks of interdependent genes. Transient regulation of gene expression occurs through association with histone acetyltransferases (HATs) and histone deacetyltransferases (HDACs), while long-term gene silencing is enacted by interactions with histone methyltransferases (HMTs) and DNA methyltransferases (DNMTs). No single classification scheme may accurately characterize the function of KLF11 in its entirety. However, the classification of family members by virtue of its co-factors provides a framework to understand the functional differences in the manner in which it engages chromatin to activate or repress transcription in a dynamic and reversible manner.

The association between KLF11 and chromatin coupling is well established. Initial biochemical characterization KLF11 revealed that its N-terminus possesses a domain adopts an α-helical conformation and that mutations within this domain significantly disrupted its binding to Sin3 and the subsequent recruitment of HDACs (32,34).

Subsequently, KLF11 was found to associate with HP1α, one of the epigenetic "gatekeepers" of gene silencing through its extreme C-terminus (64-66). HP1 proteins repress gene expression by binding to H3K9me marks and interacting to H3K9 HMTs, such as G9a or SUV39H1, which methylate this residue on adjacent nucleosomes (67,68). Deletion of the C-terminus leads to deregulation of tumor suppression functions mediated by KLF11. Finally, KLF11 has also been demonstrated to associate with WD40, WWI, WWII, and SH3-domain containing proteins through a proline-rich domain. Interaction of KLF11 and WD40 protein Gβ2, for example, is disrupted in the presence of the A347S mutation, the variant associated with development of Maturity Onset Diabetes of the Young 7 (MODY7). The effect of decoupling of KLF11 to chromatin co-factors, however, has only been studied on small subsets of genes related to the system of focus.

The investigation presented here is the first to attempt to dissect the roles of KLF11 and chromatin coupling on a genome-wide level to ascertain the combinatorial or unique effects of each type of previously identified interacting chromatin system. Utilizing a single cell type with overexpression of wild type KLF11 and the three previously described mutants, A347S, Δv486, and EAPP, we were able to examine the interplay between coupling of the transcription factor to the WD40, HP1/HMT, and Sin3/HDAC systems, respectively. By introducing saturating amounts of each mutant into pancreatic epithelial cells, we were able to tease apart the relative and overlapping contributions of each chromatin coupling mechanism to the regulation of pathobiological gene networks. Our results reveal that in the pancreatic epithelial cell utilized as our model, a large number of the affected genes, networks, and signaling cascades are modulated by the coupling of KLF11 to its variety of chromatin systems. The number of genes identified as regulated independently of these systems represented <1% of the total genes mediated in the presence of wild type or mutant proteins.

CONCLUSION

More than fifteen years ago, our research team hypothesized that the discovery and study of KLF proteins and their chromatin cofactors would assist in unraveling complex human diseases. In the intervening years, our understanding of the interplay between KLF proteins and epigenetic machinery in transient and long-term gene regulation has grown exponentially. With knowledge inferred from the predictive power of rationally derived computational models, we propose a new paradigm for KLF mediation of gene networks through the translation of input from cellular milieu into epigenetic information to effect changes in chromatin structure. KLF proteins currently provide the best model for understanding the interactions between mechanisms of "hard inheritance" (environmental and genetic variation) and "soft inheritance" (epigenetic variation) in underscoring the phenotypic variability observed in complex disease mechanisms. Perhaps most importantly, these new pathways open up infinite possibilities for targeting new molecules and processes in therapeutic intervention and disease management.

MATERIALS AND METHODS

Cell culture - Panc1 cell lines were obtained from the American Type Culture Collection (ATCC, Rockville, MD). Cells were cultured as described previously (118,119).

Constructs - Standard molecular biology techniques were used to clone full length KLF11, KLF11-A347S, KLF11-EAPP, and KLF11-486 into pcDNA3.1/His (Invitrogen, Carlsbad, CA). All constructs were verified by sequencing at the Mayo Clinic Molecular Biology Core Facility. Epitope-tagged (6XHis-Xpress™) KLF11, KLF11-A347S, KLF11-EAPP, KLF11Δ486 variants as well as empty vector (Ad5CMV) were generated as recombinant adenovirus in collaboration with the Gene Transfer Vector Core at the University of Iowa.

Genome-wide expression profiles of KLF11 and mutants - Panc1 epithelial cells were plated at a density of $1x10^6$ cells/100mm dish and transduced with empty vector, KLF11, KLF11-A347S, KLF11-Δ486, or KLF11-EAPP at an MOI of 150. RNA was prepared as

previously described from pooled biological triplicates (179). Global gene expression profiling was carried out in technical duplicate at the Microarrays Facility of the Research Center of Laval University CRCHUL utilizing the Affymetrix Human Gene 1.0 ST arrays (28,869 well-annotated genes and 764,885 distinct probes). Intensity files were generated by Affymetrix GCS 3000 7G and the Gene-Chip Operating Software (Affymetrix, Santa Clara, CA). A subset of genes was validated by qPCR as previously described (Supplemental Figure 1) (124,129)

Genome-wide promoter binding profile of KLF11 - Panc1 epithelial cells were transfected with full-length His-tagged KLF11. ChIP was performed as previously described (70,116,129,131) using an antibody against the His-Tag (OMNI D8; Santa Cruz Biotechnology) to detect recombinant expression of KLF11. Non-specific IgG antibody was utilized as a negative control. Binding activity was derived using the NimbleGen human promoter hybridization system (Madison, WI). Peaks were detected by searching for >4 probes where signals were above the specified cutoff values (90% to 15%) using a 500bp sliding window along 5kb upstream of the transcriptional start site in human promoters. Each peak was assigned a score that is the log2 ratio of the fourth highest probe in each peak. If multiple peaks are present, the peak nearest the TSS is reported. Ratio data was then randomized 20 times to evaluate the false discovery rate (FDR). Only peaks with FDR scores <0.2 were deemed high confidence binding sites and reported.

Data analysis - Data analysis, background subtraction and intensity normalization was performed using Robust Multiarray Analysis (RMA) (180). Genes that were differentially expressed along with false discovery rate (FDR) were estimated from t test (>0.005) and corrected using Bayes approach (181,182). A threshold of +/- 1.5 log2 fold change with a p-value with FDR of less than 0.05 (without FDR for EAPP mutant) was used to determine significantly regulated targets. Data analysis, hierarchical clustering, and ontology were performed with the OneChanelGUI to extend affylmGUI graphical interface capabilities and Partek Genomics Suite, version 6.6 (Partek Inc., St. Louis, MO)

with ANOVA and GO ontological analysis (183). Selected probes and their fold changes were loaded into Ingenuity Pathways Analysis Software (Ingenuity® Systems, www.ingenuity.com) for annotation, redundancy checks, canonical pathway, biological network, and upstream regulator analysis using default parameters.

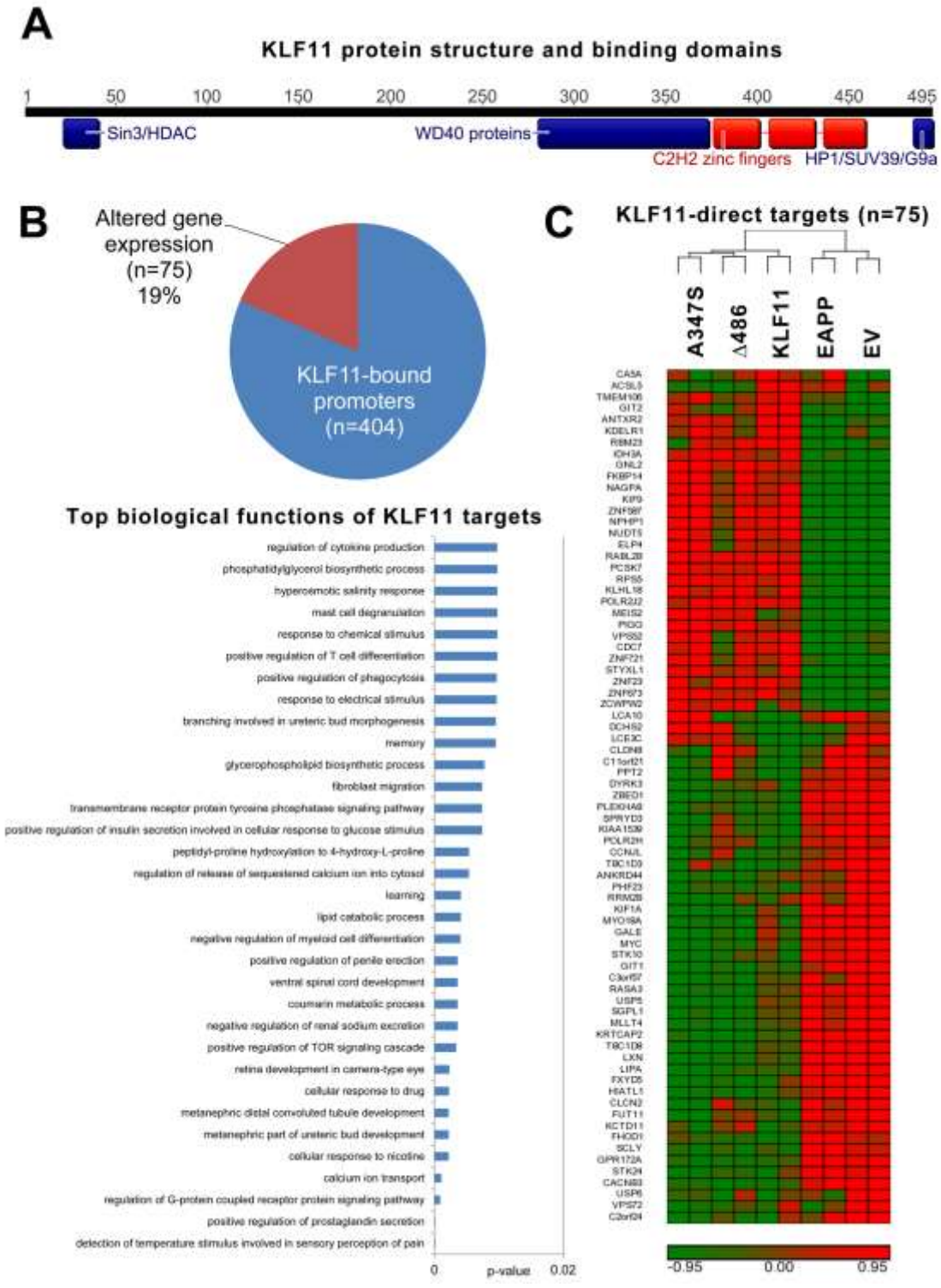


Figure 1: KLF11-mediated gene expression is disrupted in the presence of chromatin coupling mutations. (A) The KLF11 protein is 495 amino acids and contains a highly conserved C_2H_2 zinc finger

domain at its C-terminus. Known chromatin coupling domains explored in this study are highlighted. (B) Panc1 epithelial cells were transfected with wild type KLF11 and chromatin immunoprecipitation performed and hybridized to a whole genome promoter array. 404 genes were found bound by KLF11 with 2000 base pairs upstream and 500 base pairs downstream of transcriptional start sites. Gene ontological analysis reveals enrichment of genes in a number of known KLF11-associated biological processes, including immune response, TOR signaling, and insulin sensitivity. For whole genome analysis, mutants were designed against three of the previously characterized chromatin coupling domains. The EAPP mutation in the N-terminus of the protein decouples the transcription factor from the Sin3/histone deacetylase system. The A347S mutation in the proline rich domain decouples KLF11 from WD40 containing proteins. Finally, the deletion mutation starting at amino acid 486 disconnects KLF11 from the HP1/histone methyltransferase system. Panc1 epithelial cells were transduced with empty vector, wild type KLF11 or the A347S, Δ486, or EAPP mutants. Whole genome transcriptional profiling was performing using the Affymetrix Human Gene 1.0 ST array system. The criteria for significant regulation over empty vector was set at a threshold of +/- 1.5 log2 fold change and a p-value with false discovery rate of less than 0.05. For the EAPP mutant, the p-value did not include false discovery rate thresholding due to the limited experimental effects of this mutant. (C) 75 genes were significantly regulated by overexpression of KLF11 ($p<0.05$) that are directly bound by KLF11 as determined by chromatin Immunoprecipitation. Examination of the effects of the overexpression of the three chromatin decoupling mutants reveals that expression is frequently altered in the presence of one or more of these variants.

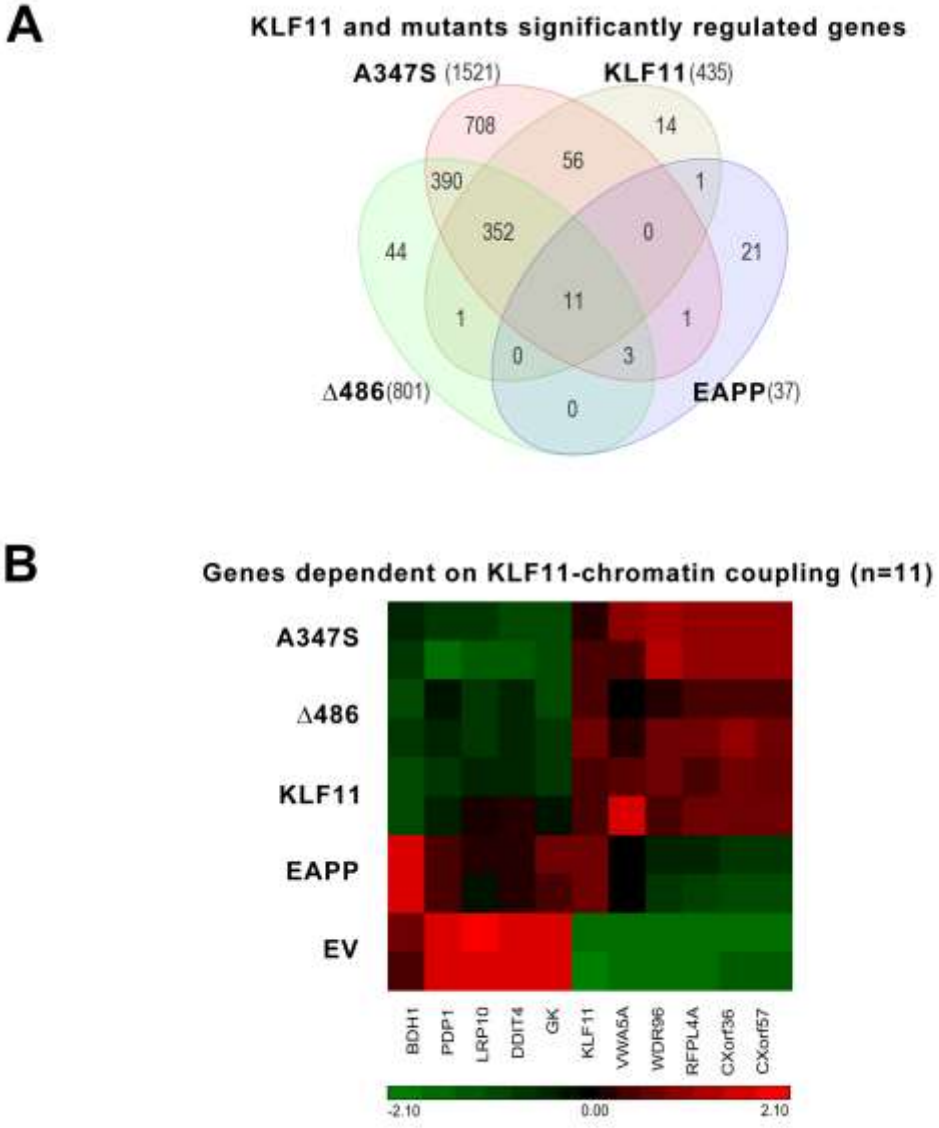


Figure 2: KLF11 differentially couples to distinct chromatin co-factor systems. (A) Venn diagram of the overlap of genes significantly regulated by wild type KLF11 and its mutants. Genes that occur at the overlap of KLF11 and all three of its chromatin-coupling mutants were deemed as independent of the effects of chromatin de-coupling and therefore inherent to the protein itself. (B) Heatmap of the 11 KLF11 regulated gene that are dependent on interaction with all three chromatin coupling systems.

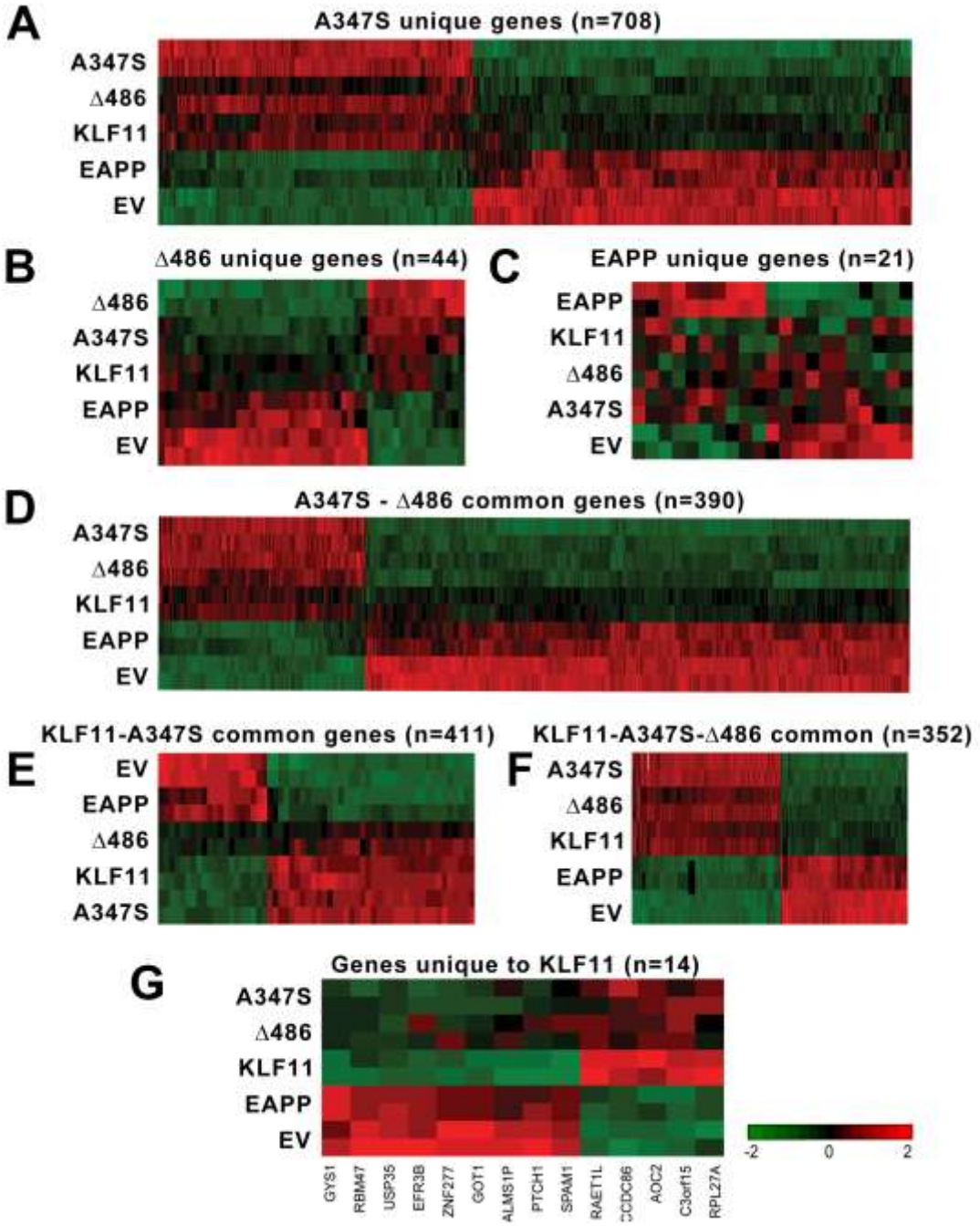


Figure 3: KLF11 regulates gene expression through combinatorial and singular chromatin coupling events. (A) The A347S mutant, which decouples KLF11 from WD40 proteins, displays the most genome wide effects, uniquely affecting the expression of 708 genes, whereas the (B) Δ486 deletion mutation (decouples KLF11 from HP1/histone methyltransferases) and the (C) EAPP mutation (decouples KLF11 from binding the Sin3 scaffold protein and subsequently histone deacetylases) uniquely regulate only 44 and 21 genes, respectively. (D) Examination of the overlap between genes significantly regulated by the A347S mutant reveals that approximately 50% are regulated in a similar fashion to the Δ486 mutant,

although with varying degrees of intensity. The EAPP mutant, however, displays near complete reversal of these targets. (E) Wild type KLF11 and the A347S mutant share 411 targets apart from either the EAPP or Δ486 mutants, although (F) 352 targets exists that are shared between all three systems. These data indicate that KLF11 chromatin coupling occurs in a largely combinatorial fashion. (G) Only 14 genes were identified that are uniquely regulated by KLF11, independent of chromatin coupling to the transcription factor. Ontological analysis of these genes reveals roles in cancer, cellular proliferation, and metabolism.

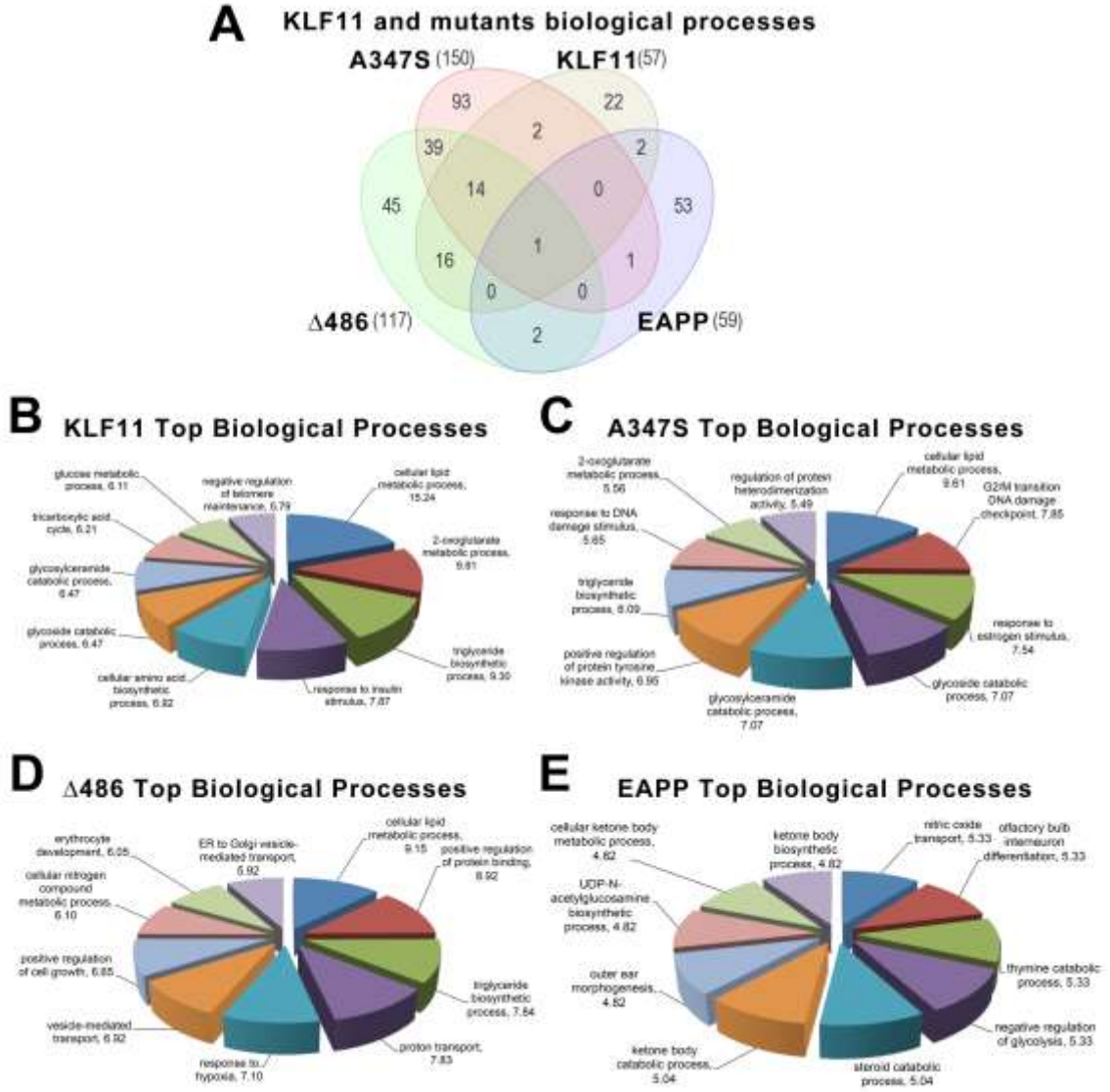


Figure 4: Biological regulated by KLF11 and its chromatin binding partners. Genes significantly regulated by wild type KLF11 and its mutants compared to empty vector were analyzed for enrichment of biological processes by an ontological approach. A threshold off 3 genes and a p-value of less than 0.05 by Fisher's Exact Test were required to be considered significantly regulated by the transcription factor and its mutants. (A) Venn diagram of biological processes demonstrates that the A347S mutant causes the largest number of genome-wide effects with significant alterations in 150 processes, 93 of which are unique to the mutant. Only one biological process, lipid cellular metabolism, is dependent on the combinatorial effects of all three chromatin coupling systems. The top biological processes were computed for each of the proteins compared to empty vector. The pie charts for (B) Wild type KLF11, (C) A347S mutant, (D) Δ486 mutant, and (E) EAPP mutant display the top ten scoring biological processes for each with enrichment percentage displayed. These results indicate that the effects of KLF11 coupling to chromatin co-factors impact on a large number of biological functions.

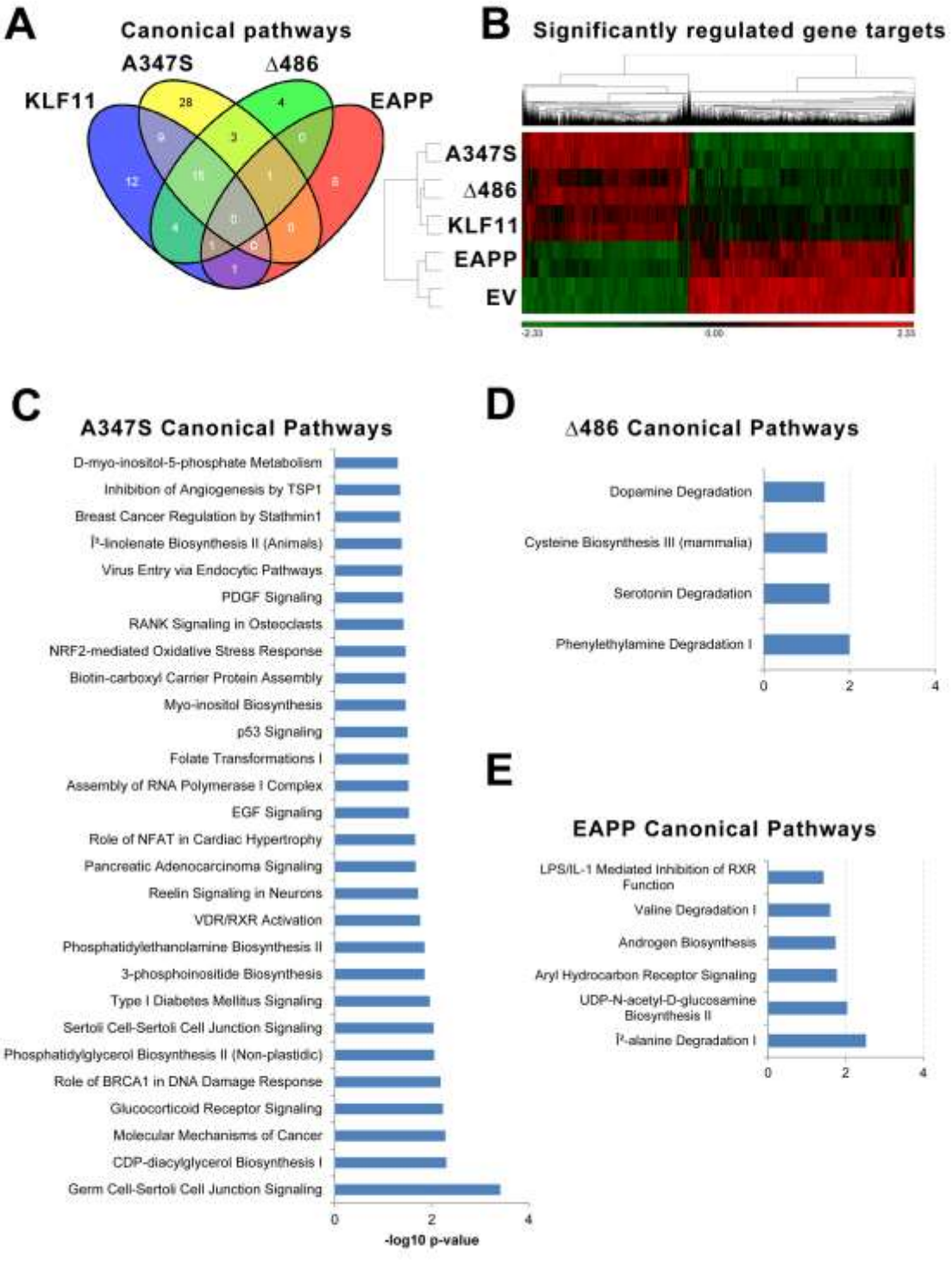


Figure 5: Canonical signaling pathways mediated by KLF11 and mutants. Using the Ingenuity Global Canonical Pathways algorithm, significantly modulated signaling pathways were identified for wild type KLF11 and its mutants. A p-value of less than 0.05 as determined by Fisher's Exact Test was used as

criteria for significant association of focus genes to predetermined pathways curated from published literature. (A) Venn diagram of significant canonical pathways regulated by wild type KLF11 and mutants reveals that the effects of the EAPP mutation are less significant in this experiment compared to the A347S and 486 mutants. These data are supported by the clustering of all significantly regulated targets as shown in (B) which demonstrated that the A347S and Δ486 mutants cluster closely with wild type KLF11 while the EAPP mutant clusters with the empty vector control, indicating complete reversal of the effects of the other proteins. (C) Among the 28 top scoring unique canonical pathways mediated by the 374 mutant are signaling pathways, including, EGF, PDGR, p53, RANK and reelin signaling, and glucocorticoid receptor signaling. The Δ486 and EAPP mutants mediate 4 and 6 pathways respectively, both of which are centered on processes involved on the degradation of biogenic amines. These data support KLF11 as a master transcriptional regulator that mediates large, interconnected signaling cascades.

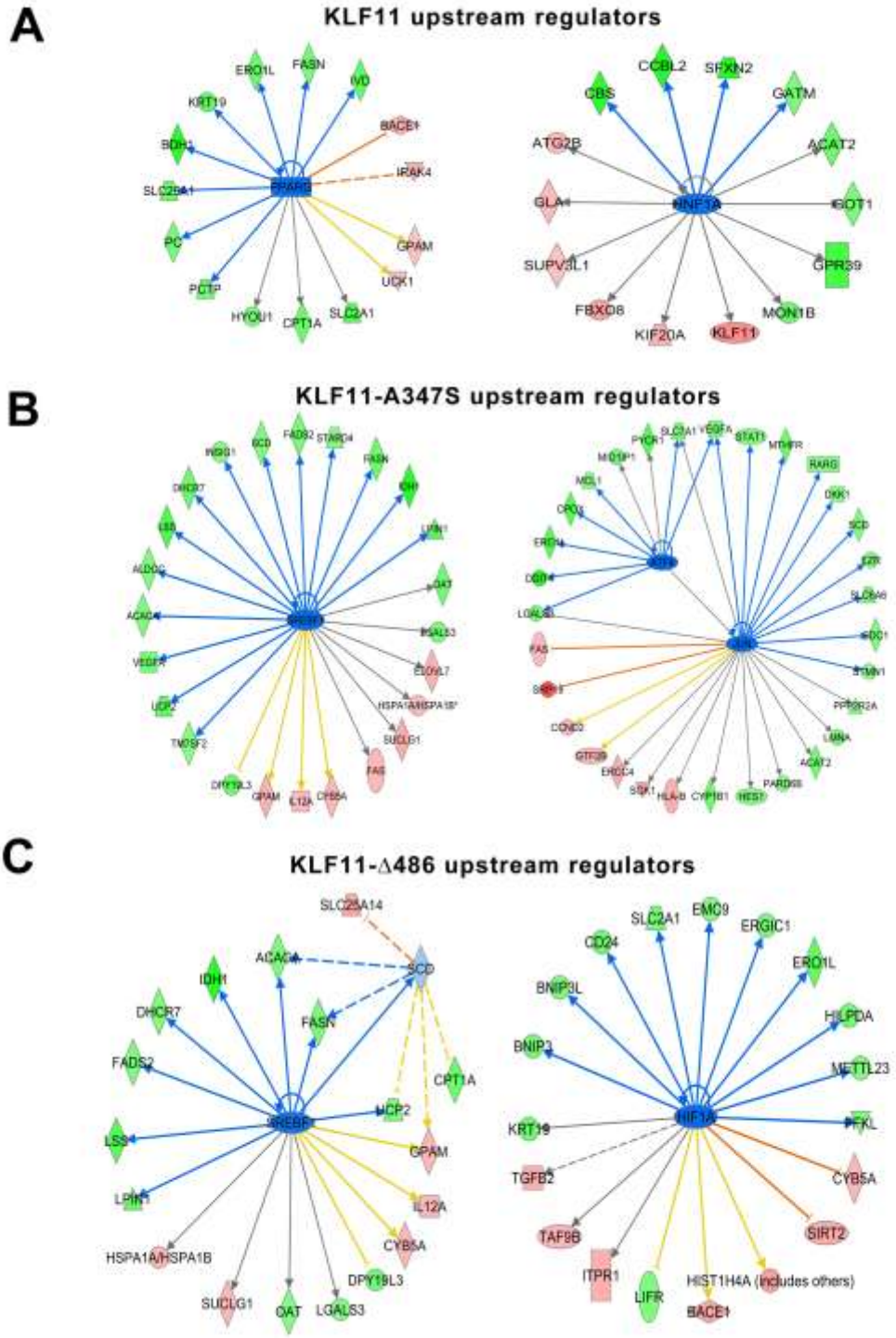


Figure 6: Biological disease networks mediated by KLF11 and its mutants. Using Ingenuity's network analysis tool, significantly regulated genes by KLF11 and its mutants were examined for connectivity. Networks were scored by relevance and a p-value assigned by Fisher's Exact Test to discriminate networks

generated by chance alone. A threshold of $p<0.05$ was employed to determine significant networks. Networks were then assessed for association to known biological or disease processes. Wild type KLF11 and its mutants share three processes: death and survival, hereditary disorders, and molecular transport. Of the identified processes, 10 are unique to wild type KLF11 and 6 and 8 are unique to the A347S and Δ486 mutants, respectively. A number of known KLF11-mediated biological and diseases processes were revealed by this analysis, providing an internal control, including endocrine disorders, gastrointestinal disorders, and cellular growth and proliferation processes. However, the analysis also generated a number of novel biological and disease processes that remain to be experimentally validated for KLF11.

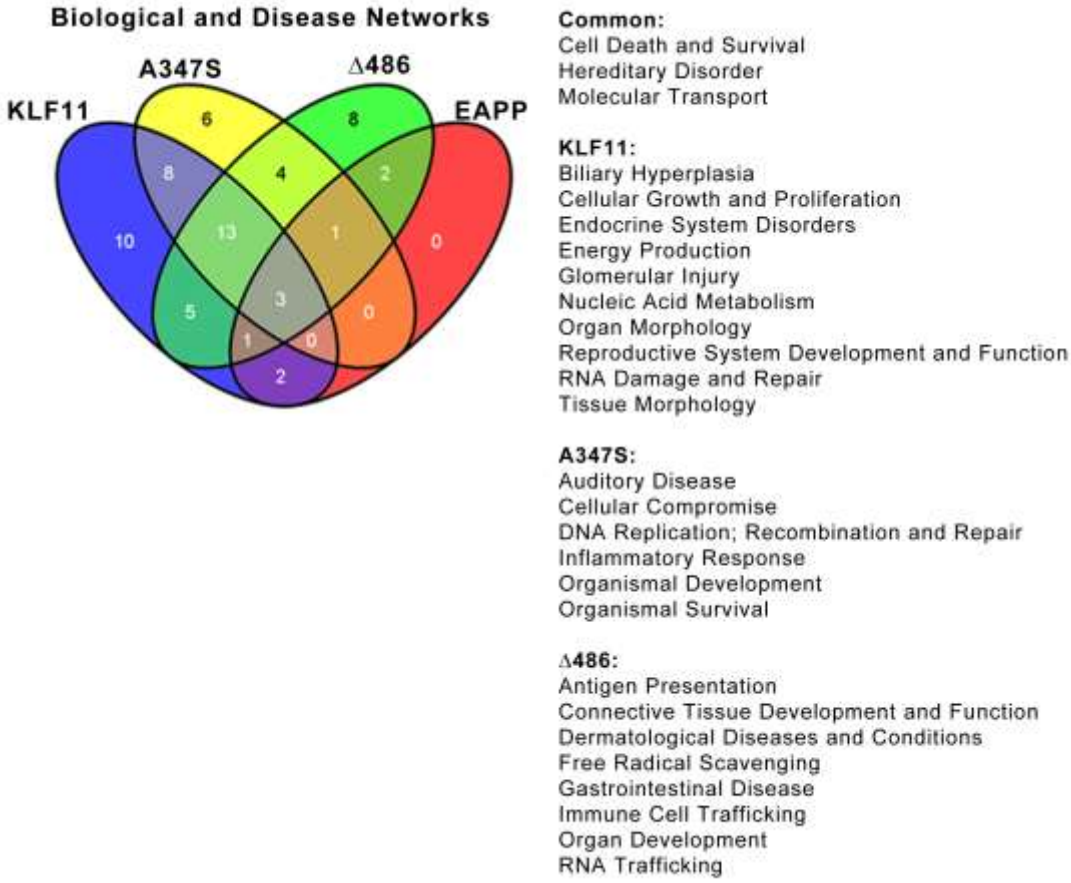


Figure 7: Upstream regulators of KLF11 and its mutants. Gene targets significantly regulated by KLF11 and its mutants were analyzed by Ingenuity's Upstream Regulator Analytic that compares the experimentally derive activation or inhibition of focus molecules to relationships between upstream regulators and target molecules known from published data. For each potential upstream regulator and its targets, a p-value for the degree of overlap between that of the protein under study and a z-score for the activation status are calculated. For our study, a threshold of $p<0.05$ and a minimum z-score of +/- 2 were employed. The full results are published in Supplemental Table 5 and sample signaling cascades presented for wild type KLF11 (A), A347S (B), and Δ486 (C). No cascades were generated that proved significant for the EAPP mutant although a more permissive criteria of a p-value of less than 0.05 did generate a number of single gene associations listed in Supplemental Table 5, including a variety of other KLF proteins, hinting at co-regulation within the KLF family. Many of the upstream regulators, such as PPARG, are previously identified co-regulators of KLF11. While the upstream regulators frequently overlap between wild type and mutant conditions, the end gene targets are unique to each condition. These results suggest that the deregulation of KLF11 from a large number of gene networks is completely dependent on the effective coupling of the transcription factor to its chromatin co-factor system.

Supplemental Table 1: Overlapping biological function mediated by KLF11 and mutants.

		Number of genes				Enrichment Score			
GO ID	Common to KLF11, A347S, Δ486, and EAPP	KLF11	Δ486	A347S	EAPP	KLF11	Δ486	A347S	EAPP
44255	cellular lipid metabolic process	15	16	10	2	15.24	9.15	9.61	4.01

		Number of genes				Enrichment Score			
GO ID	Common to KLF11 and A347S and Δ486 mutants	KLF11	Δ486	A347S	EAPP	KLF11	Δ486	A347S	EAPP
6103	2-oxoglutarate metabolic process	5	4	3	1	9.81	4.89	5.56	3.40
19432	triglyceride biosynthetic process	6	7	4	1	9.30	7.84	6.09	3.50
8652	cellular amino acid biosynthetic process	5	5	3	1	6.92	4.29	3.93	3.27
16139	glycoside catabolic process	2	2	2	1	6.47	5.25	7.07	3.95
46477	glycosylceramide catabolic process	2	2	2	1	6.47	5.25	7.07	3.73
43497	regulation of protein heterodimerization activity	2	2	2	1	4.91	3.72	5.49	3.50
35338	long-chain fatty-acyl-CoA biosynthetic process	3	4	2	1	4.89	5.09	3.30	4.48
55114	oxidation-reduction process	11	18	9	1	4.55	5.38	4.50	3.31
43496	regulation of protein homodimerization activity	2	2	2	1	4.32	3.16	4.89	3.50
51607	defense response to virus	4	5	3	1	4.10	3.36	3.31	3.31
16192	vesicle-mediated transport	9	18	8	1	3.58	6.92	4.24	3.95
6081	cellular aldehyde metabolic process	2	3	2	1	3.52	4.37	4.08	4.64
61098	positive regulation of protein tyrosine kinase activity	2	3	3	1	3.52	4.37	6.95	3.80
43627	response to estrogen stimulus	4	8	6	1	3.03	5.60	7.54	3.50

Supplemental Table 2: Top biological processes mediated by KLF11 and mutants compared to empty vector.

KLF11	EAPP	Δ486	A347S
cellular lipid metabolic process	nitric oxide transport	cellular lipid metabolic process	cellular lipid metabolic process
2-oxoglutarate metabolic process	olfactory bulb interneuron differentiation	positive regulation of protein binding	G2/M transition DNA damage checkpoint
triglyceride biosynthetic process	thymine catabolic process	triglyceride biosynthetic process	response to estrogen stimulus
response to insulin stimulus	negative regulation of glycolysis	proton transport	glycoside catabolic process
cellular amino acid biosynthetic process	steroid catabolic process	response to hypoxia	glycosylceramide catabolic process
glycoside catabolic process	ketone body catabolic process	vesicle-mediated transport	positive regulation of protein tyrosine kinase activity
glycosylceramide catabolic process	outer ear morphogenesis	positive regulation of cell growth	triglyceride biosynthetic process
tricarboxylic acid cycle	UDP-N-acetylglucosamine biosynthetic process	cellular nitrogen compound metabolic process	response to DNA damage stimulus
glucose metabolic process	cellular ketone body metabolic process	erythrocyte development	2-oxoglutarate metabolic process
negative regulation of telomere maintenance	ketone body biosynthetic process	ER to Golgi vesicle-mediated transport	regulation of protein heterodimerization activity
fatty acid beta-oxidation	response to nutrient	tricarboxylic acid cycle	response to cholesterol
response to heat	response to ethanol	embryonic digit morphogenesis	response to ionizing radiation
negative regulation of IGFR signaling pathway	glycerol-3-phosphate metabolic process	response to estrogen stimulus	positive regulation of protein binding
cellular response to stimulus	mesenchymal to epithelial transition	polyamine metabolic process	positive regulation of DNA repair
biosynthetic process	valine metabolic process	oxidation-reduction process	proton transport
positive regulation of cell growth	regulation of neural precursor cell proliferation	chromatin remodeling	induction of apoptosis by intracellular signals
cellular nitrogen compound metabolic process	forelimb morphogenesis	post-embryonic development	regulation of transcription, DNA-dependent
isocitrate metabolic process	hindlimb morphogenesis	glycoside catabolic process	regulation of protein homodimerization activity
regulation of protein heterodimerization activity	fructose 6-phosphate metabolic process	glycosylceramide catabolic process	response to folic acid
L-methionine salvage from methylthioadenosine	oxygen transport	myeloid leukocyte differentiation	oxidation-reduction process

Supplemental Table 3: Biological processes unique to the decoupling of KLF11 from WD40 proteins (A347S mutant)

List 1	List 2	List 3
chromatin modification chromatin silencing at rDNA DNA replication DNA-dependent DNA replication initiation negative regulation of centrosome duplication nucleotide-excision repair positive regulation of DNA repair telomere maintenance telomere maintenance via telomere lengthening telomere maintenance via telomere shortening protection from non-homologous end joining at telomere protein localization to chromosome, telomeric region	negative regulation of protein autophosphorylation protein dephosphorylation histone H3 deacetylation methylation histone H3-K4 demethylation	transcription elongation from RNA polymerase II promoter transcription initiation from RNA polymerase II promoter regulation of transcription from RNA polymerase II promoter in response to oxidative stress positive regulation of transcription from RNA polymerase I promoter regulation of transcription from RNA polymerase I promoter termination of RNA polymerase I transcription mRNA transport

Supplemental Table 4: Biological processes unique to the decoupling of KLF11 from HP1a/HMT (Δ486 mutant).

List 1	List 2
beta-amyloid metabolic process positive regulation of triglyceride biosynthetic process L-serine metabolic process porphyrin metabolic process heme biosynthetic process sterol biosynthetic process polyamine biosynthetic process dopamine biosynthetic process folic acid metabolic process response to cholesterol response to glucose stimulus aromatic amino acid family metabolic process response to iron ion electron transport chain regulation of proteasomal protein catabolic process protein deubiquitination protein modification process positive regulation of protein export from nucleus protein import into nucleus, translocation positive regulation of protein import into nucleus, translocation protein autophosphorylation vesicle docking involved in exocytosis actin cytoskeleton reorganization protein tetramerization	positive regulation of epidermal growth factor receptor signaling pathway positive regulation of epithelial to mesenchymal transition inner cell mass cell proliferation positive regulation of epithelial cell migration cellular response to estradiol stimulus negative regulation of survival gene product expression negative regulation of growth induction of apoptosis by intracellular signals cell death

Supplemental Table 5: Upstream regulators of KLF11 and mutants.

KLF11 URs	State	Z-score	p-value	Target molecules in dataset
PPARG	Inhibited	-2,323	9,74E-03	BACE1,BDH1,CPT1A,ERO1L,FASN,GPAM,HYOU1,IRAK4,IVD,KRT19,PC,PCTP,SLC25A1,SLC2A1,UCK1
HNF1A	Inhibited	-2,000	8,90E-02	ACAT2,ATG2B,CBS,CCBL2,FBXO8,GATM,GLA,GOT1,GPR39,KIF20A,KLF11,MON1B,SFXN2,SUPV3L1

A347S URs	State	Z-score	p-value	Target molecules in dataset	Mechanistic Network
PPARG	Inhibited	-3,585	2,82E-04	ACAA1,ACACA,ANGPTL4,ATP6V1D,BACE1,BCL6,BDH1,CA2,CAT,CCPG1,CPT1A,ERO1L,F11R,FASN,GPAM,HES1,HYOU1,IL12A,INSIG1,IRAK4,IVD,JUN,KLF6,KRT19,LNPEP,NDUFA5,ODC1,PC,PCTP,PDHB,PMM1,SAT1,SCD,SDC1,SLC25A1,SLC25A20,SLC2A1,SORBS1,SREBF1,TGFBR1,TJP1,TKT,UCK1,UCP2,VEGFA	
RXRA	Inhibited	-3,386	1,55E-01	ACACA,ACSL3,ARL4C,CAT,CLMN,CPT1A,FADS2,FASN,HIF1A,HSD17B4,ILK,INSIG1,MID1IP1,PC,RARG,SAT1,SCD,SDC1,SORBS1,SREBF1,TGFB2,VEGFA	
STAT4	Inhibited	-3,323	1,69E-02	ACSS1,ADSS,ALDOC,ARFGAP3,ATF4,ATP7A,EPOR,ERO1L,ERRFI1,FZD7,GRTP1,HILPDA,HSPA1A/HSPA1B,ING2,LPIN1,LRRFIP1,MBOAT2,P4HA2,PRDX6,RNF128,SAT1,SERPINB1,SETD5,STAT1,VEGFA	FOXO4,HIF1A,TP53
HIF1A	Inhibited	-3,257	3,56E-03	ALDOC,ANGPTL4,BACE1,BNIP3,BNIP3L,CD24,CHKA,CYB5A,EMC9,EPAS1,EPOR,ERGIC1,ERO1L,HIF1A,HILPDA,HIST1H4A (includes others),HIST2H2AC,ITPR1,JUN,JUP,KDM3A,KRT19,LIFR,MCL1,P4HA1,P4HA2,PFKL,QKI,SDC4,SIRT2,SLC25A37,SLC29A1,SLC2A1,TAF9B,TGFB2,VEGFA	HIF1A,PPARG,SREBF1,SREBF2,TP53
PPARGC1A	Inhibited	-2,930	1,00E00	ACACA,CALM1 (includes others),CAT,CPT1A,FASN,LPIN1,SCD,SLC25A20,SREBF1,UCP2	PPARG,SCAP,SREBF1,SREBF2,TP53
SREBF2	Inhibited	-2,875	2,18E-03	ACACA,ALDOC,CYB5A,DHCR7,FADS2,FASN,IDH1,INSIG1,LSS,SCD,SREBF1,STARD4,TM7SF2	HIF1A,PPARG,SREBF1,SREBF2,TP53
ERG	Inhibited	-2,828	3,34E-01	ADD1,ARHGAP17,CDC42BPB,DBN1,ILK,PIM1,PTPN4,RAB2A,RASA2	FOXO4,HIF1A,HIF3A,TP53

ATF4	Inhibited	-2,736	1,74E-01	ATF4,CPOX,DDIT4,ERO1L,JUN,LGALS3,MCL1,MID1IP1,PYCR1,SLC7A1,VEGFA	
JUN	Inhibited	-2,660	1,00E00	ACAT2,CCND2,CYP1B1,DKK1,ERCC4,EZR,FAS,GTF2B,HES1,HLA-B,JUN,LGALS3,LMNA,MTHFR,PARD6B,PPP2R2A,RARG,SCD,SDC1,SGK1,SLC6A6,SLC7A1,STAT1,STMN1,VEGFA	
NFE2L2	Inhibited	-2,587	1,00E00	ABCC4,AKR1A1,ATF4,BNIP3,C5,CAT,CCRN4L,DCTN3,DHCR7,EPAS1,EPB41,ESD,IFNGR2,LMNA,NQO2,SAT1,SEC23A,SLC1A4,SLC2A1,SREBF1,SYT1,TBRG1,TCN2,UGDH,USP14,VEGFA	FOXO4,HIF1A,PTEN,TP53
MITF	Inhibited	-2,540	3,49E-01	ASAH1,CCNG2,CHKA,GM2A,HIF1A,ITGA3,IVNS1ABP,LGALS3,PSEN2,SDC1,SLC19A2,SORT1,TFAP2A	ERBB2,FOXO4,HIF1A,PPARG,PTEN,TP53,TSC2
SREBF1	Inhibited	-2,530	1,79E-03	ACACA,ALDOC,CYB5A,DHCR7,DPY19L3,ELOVL7,FADS2,FAS,FASN,GPAM,HSPA1A/HSPA1B,IDH1,IL12A,INSIG1,LGALS3,LPIN1,LSS,SCD,SREBF1,STARD4,SUCLG1,TM7SF2,UCP2,VEGFA	
NR1H3	Inhibited	-2,453	1,28E-01	ACACA,ACSL3,ARL4C,FASN,GPAM,MID1IP1,SCD,SREBF1,VEGFA	
ARNT	Inhibited	-2,425	1,63E-01	BNIP3,CCND2,CYP1B1,ERO1L,HIF1A,KIF20A,MYO1C,SLC2A1,VEGFA	
GLI1	Inhibited	-2,268	5,23E-02	ASPM,CCND2,CD24,CMBL,DKK1,EZR,IMPA2,INSIG1,JUP,KRT19,LMNA,MRPS6,NQO2,PIM1,PLOD1,PPAP2C,RPS6KA1,VEGFA,ZC3HAV1L	
NR1H2	Inhibited	-2,236	1,54E-01	ACACA,FASN,HSD17B4,LRP8,SCD,SREBF1,VEGFA	
RORA	Inhibited	-2,236	3,59E-01	ADIPOR1,ARNTL,CCRN4L,CYP2C18,ELOVL7,FASN,GSTM2,HSD17B7,SCD,SLC16A10,SLC2A13,SREBF1,UCP2	
MLXIPL	Inhibited	-2,195	2,52E-02	ACACA,CPT1A,FASN,MID1IP1,SCD	
SP1	Inhibited	-2,125	1,31E-01	ACSS1,ALDH3A2,BACE1,BNIP3L,CAT,CBS,CCND2,CD99,CDC25B,CYP1B1,EPOR,EZR,FAS,FASN,HDAC1,HIF1A,HIST1H4A (includes others),HSD17B7,HSPA1A/HSPA1B,IL12A,ITGA2,JUN,KRT19,MAT2B,MCL1,MSH6,OGG1,PIM1,RBL1,RECK,SETDB1,SGK1,SLC2A1,SLC39A8,SLC7A1,SMAD3,SREBF1,S	

				TAT1,TGFB2,TGFBR1,UGDH,VEGFA	
PPARA	Inhibited	-2,009	2,70E-02	ACAA1,ACACA,ALDH3A2,ANGPTL4,C1S,C5,CAT,CHKA,CPT1A,CYP2C18,DHCR7,F11R,FADS2,FASN,GPAM,GPD2,HLA-E,HSD17B4,INSIG1,KIF20A,LIFR,LSS,OGG1,PC,PCTP,PPM1D,PRDX6,QPCT,RAD51B,RETSAT,SAT1,SCD,SLC25A20,SLC27A2,SOCS2,SREBF1,TJP1,UCP2,VEGFA	
E2f	Inhibited	-2,000	1,13E-01	EPAS1,HDAC1,HIF1A,HIST1H3A (includes others),HIST1H4A (includes others),ITGA6,MFAP1,MIR17HG,RBBP5,RBL1,RECQL,TRMT13	
PXR ligand-PXR-Retinoic acid-RXRÎ±	Inhibited	-2,000	2,87E-01	ALDH3A2,CAT,GSTM2,SCD	
PPRC1	Inhibited	-2,000	4,40E-01	ATF4,DDIT4,ERRFI1,NAMPT	
GFI1	Activated	2,646	3,38E-01	ATF1,ETS2,JUN,KAT2B,RIPK1,SMAD3,STAT1,TNFRSF1A	
STAT1	Activated	2,538	1,00E00	CCND2,FAS,HIF1A,IFI27,IL12A,JUN,PIM1,SAMHD1,SLFN5,SMAD2,SMAD3,STAT1,USP18	
PIAS1	Activated	2,449	1,40E-02	ACACA,FASN,MCL1,SCD,SREBF1,STAT1	
HOXD10	Activated	2,449	3,43E-01	EZR,HBEGF,ITGA3,NCS1,TJP1,USP14	
SPDEF	Activated	2,449	3,89E-01	HIF1A,ITGA3,ITGA6,SDC1,SMAD2,SMAD3	
HIC1	Activated	2,345	2,14E-02	AHNAK2,CA2,CCDC176,ITPR1,LRP8,PLEC,SIRT1,SNAPC1,TNS3,WDR6	
EZH2	Activated	2,141	1,00E00	ANXA6,CCND2,CLDN10,CYB5R2,CYP1B1,DDT,DKK1,EZR,FUCA1,SIRT1,TBX3,WTAP	
VHL	Activated	2,137	1,00E00	ATF4,BNIP3,CLDN4,EPAS1,HIF1A,LMNA,SLC2A1,VEGFA	
N-cor	Activated	2,000	1,85E-01	ACACA,FASN,SCD,SREBF1,USP18	

Δ486 URs	**State**	**Z-score**	**p-value**	**Target molecules in dataset**	**Mechanistic Network**
RXRA	Inhibited	-3,104	4,65E-02	ACACA,ARL4C,CAT,CLMN,CPT1A,FADS2,FASN,HIF1A,HSD17B4,MID1IP1,SAT1,SCD,SORBS1,SREBF1,TGFB2	FOXO3, FOXO4, Insulin,PPARG,SREBF1,

					SREBF2,TP53,arachidonic acid
PPARGC1A	Inhibited	-2,774	1,34E-01	ACACA,CALM1 (includes others),CAT,CPT1A,FASN,LPIN1,SCD,SREBF1,UCP2	
PPARG	Inhibited	-2,652	7,65E-05	ACACA,BACE1,BCL6,BDH1,CA2,CAT,CCPG1,CPT1A,ERO1L,FASN,GPAM,HYOU1,IL12A,IRAK4,IVD,KRT19,LNPEP,ODC1,PCTP,PDHB,PMM1,SAT1,SCD,SLC25A1,SLC2A1,SORBS1,SREBF1,TGFBR1,UCK1,UCP2	FOXO3,PPARG,SREBF1,SREBF2,TP53
NFE2L2	Inhibited	-2,578	1,64E-01	ABCC4,AKR1A1,BNIP3,C5,CAT,DCTN3,DHCR7,EPB41,ESD,IFNGR2,NQO2,SAT1,SEC23A,SLC1A4,SLC2A1,SREBF1,UGDH,USP14	
GLI1	Inhibited	-2,574	5,22E-01	CD24,IMPA2,KRT19,NQO2,PIM1,PLOD1,ZC3HAV1L	
MITF	Inhibited	-2,414	1,00E00	GM2A,HIF1A,IVNS1ABP,LGALS3,PSEN2,SORT1	
ATF4	Inhibited	-2,200	1,33E-01	CPOX,DDIT4,ERO1L,LGALS3,MID1IP1,PYCR1,SLC7A1	
MLXIPL	Inhibited	-2,195	1,80E-03	ACACA,CPT1A,FASN,MID1IP1,SCD	
TP73	Inhibited	-2,138	5,02E-01	ARNTL,CCNG1,CLMN,DLG1,FASN,IDH2,SAT1,UBE2D1	
SREBF2	Inhibited	-2,128	2,01E-03	ACACA,CYB5A,DHCR7,FADS2,FASN,IDH1,LSS,SCD,SREBF1	
NR1H3	Inhibited	-2,105	3,55E-02	ACACA,ARL4C,FASN,GPAM,MID1IP1,SCD,SREBF1	
HIF1A	Inhibited	-2,068	3,30E-02	BACE1,BNIP3,BNIP3L,CD24,CYB5A,EMC9,ERGIC1,ERO1L,HIF1A,HILPDA,HIST1H4A (includes others),ITPR1,KRT19,LIFR,PFKL,SIRT2,SLC2A1,TAF9B,TGFB2	
NR1H2	Inhibited	-2,000	3,00E-02	ACACA,FASN,HSD17B4,LRP8,SCD,SREBF1	
ERG	Inhibited	-2,000	3,68E-01	ADD1,CDC42BPB,PIM1,PTPN4,RAB2A	
RORA	Inhibited	-2,000	1,00E00	ARNTL,FASN,SCD,SLC2A13,SREBF1,UCP2	
HIC1	Activated	2,433	4,37E-02	CA2,CCDC176,ITPR1,LRP8,TNS3,WDR6	
HOXA10	Activated	2,236	1,00E00	ATF6B,IDH2,KLF10,SAT1,SCD	
STAT1	Activated	2,219	1,00E00	HIF1A,IFI27,IL12A,PIM1,SAMHD1,SLFN5,SMAD2	
PIAS1	Activated	2,000	2,22E-02	ACACA,FASN,SCD,SREBF1	
N-cor	Activated	2,000	7,71E-02	ACACA,FASN,SCD,SREBF1	
GFI1	Activated	2,000	2,73E-01	ATF1,ETS2,KAT2B,RIPK1,TNFRSF1A	

EAPP URs	p-value	Target molecules in dataset
ZNF496	1,53E-03	HBB
FOXN2	1,53E-03	HBB
HLTF	3,05E-03	HBB
DLX4	4,57E-03	HBB
TIAL1	4,57E-03	HBB
KLF7	6,09E-03	HBB
ATF4	9,55E-03	DDIT4,STC2
KLF3	1,06E-02	HBB
HNRNPD	1,22E-02	HBB
MAFF	1,37E-02	HBB
BRCA1	1,48E-02	DDIT4,HBB
CEBPG	1,52E-02	HBB
CBX2	1,67E-02	HBB
MTA2	1,67E-02	HBB
KLF13	1,82E-02	HBB
NSD1	1,97E-02	HBB
CTBP2	2,27E-02	HBB
ZFPM1	2,27E-02	HBB
GLI1	2,32E-02	GREM2,SALL1
COMMD3-BMI1	2,57E-02	HBB
MAFK	2,57E-02	HBB
SIX1	2,71E-02	SALL1
NFE2	2,71E-02	HBB
BACH1	2,71E-02	HBB
CARM1	3,16E-02	STC2
DRAP1	3,60E-02	HIST1H2BH/HIST1H2BO
KLF1	4,04E-02	HBB
GTF2B	4,04E-02	HIST1H2BH/HIST1H2BO
YBX1	4,19E-02	HBB
KAT2B	4,34E-02	HBB
HOXC8	4,34E-02	GREM2

CONCLUSION

These studies on KLF in the delivery of epigenetic information reveal a novel role for the nucleosome as a cellular nanomachine that translates KLF inputs into epigenetic outputs to regulate networks of related gene targets. We hypothesize that these mechanisms have evolved in eukaryotes as a means for translating environmental stimuli into the regulation of gene networks, thereby serving as the basis for the epigenetics of environment-gene interactions. While much of the evidence for this paradigm emerges from the field of metabolism, it is tempting to speculate—given the diverse range of KLF functions—that these principles also apply to the regulation of gene networks in general (e.g. cell cycle, differentiation).

In this paradigm (Figure 1), environmental signals are channeled from the membrane to the nucleus through signaling transduction cascades that impinge upon KLF transcription factors, activating or repressing their activity through posttranslational modification. Activated KLF proteins function as master transcription factors at the apex of a cascade of gene promoters (nodes) that form an interrelated gene network affecting a specific function (e.g. glucose metabolism, differentiation). The character of the KLF factor, the context of the promoter sequence, and the interaction with co-repressors or co-activators determine the type of epigenetic information delivered to a gene promoter by a particular factor. Antagonism between HATs and HDACs serves as a switch for the transient epigenetic status of the promoter, permitting rapid activation or repression of a gene target, respectively. In repressed targets, KLF/HDACs may recruit long-term silencing machinery, HMTs and DNMTs, to make the repressed status of a gene permanent. Among the targets transactivated by KLF proteins are additional transcriptional factors (regulation nodes). Increases or decreases in the intracellular concentration of these transcription factors act as a feedback to modulate the output of the pathway.

KLF conducts a veritable symphony of gene expression through diverse gene orchestras to underscore the entire gamut of biological functions. The old paradigm of transcriptional factors turning genes “on” or “off” by simply binding to a single site consensus sequence is woefully outdated. Transcription is regulated through intricate interactions between the transcription factor, its promoter, and its associated co-factors interacting in close proximity to fine-tune the intensity of activating or repressive signals from the environment or within the cellular milieu.

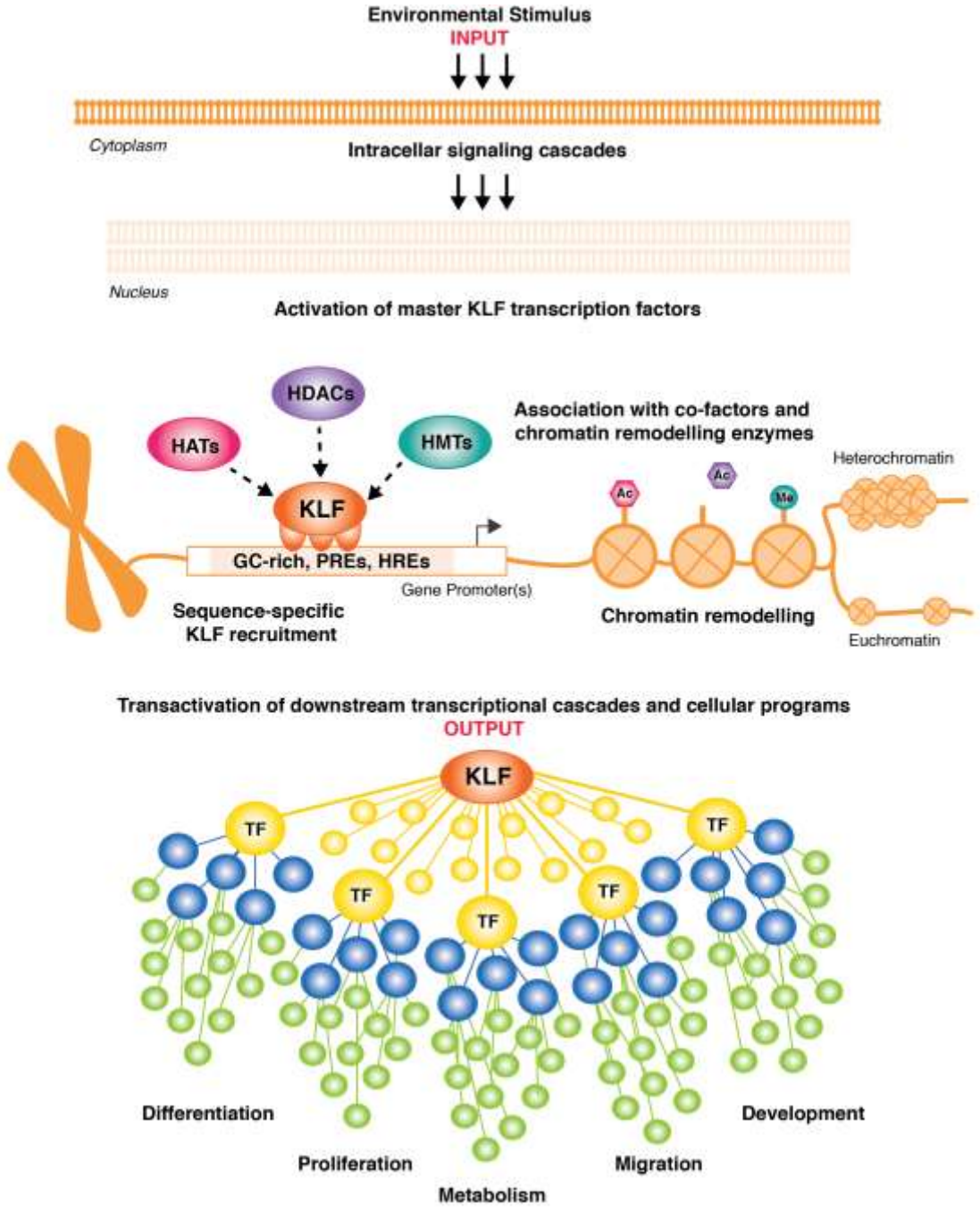


FIGURE 1. KLF are conductors of diverse gene orchestras. KLF is the central mediator in a diverse range of biological functions, from development to metabolism. Environmental stimuli from the external environment are transduced through intercellular signaling cascades leading to the sequence-specific recruitment of KLF molecules to gene promoters, which, thanks to their coupling with a variety of epigenetic molecules that determine their activating or repressing quality, are responsible for remodeling chromatin and turning genes "on" and "off." As other transcription factors are frequently the targets of KLF-mediated activation or repression, the molecule is poised to regulate large gene signaling cascades.

BIBLIOGRAPHY

1. Ptak, C., and Petronis, A. (2008) Epigenetics and complex disease: from etiology to new therapeutics. *Annu Rev Pharmacol Toxicol* **48**, 257-276
2. Zhan, S., Merlin, C., Boore, J. L., and Reppert, S. M. The monarch butterfly genome yields insights into long-distance migration. *Cell* **147**, 1171-1185
3. Rakyan, V. K., Down, T. A., Balding, D. J., and Beck, S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet* **12**, 529-541
4. Daxinger, L., and Whitelaw, E. Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat Rev Genet*
5. Feinberg, A. P., and Irizarry, R. A. Evolution in health and medicine Sackler colloquium: Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc Natl Acad Sci U S A* **107 Suppl 1**, 1757-1764
6. Jirtle, R. L., and Skinner, M. K. (2007) Environmental epigenomics and disease susceptibility. *Nat Rev Genet* **8**, 253-262
7. Richards, E. J. (2006) Inherited epigenetic variation--revisiting soft inheritance. *Nat Rev Genet* **7**, 395-401
8. Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., and Jaenisch, R. (2007) In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318-324
9. Dynan, W. S., and Tjian, R. (1983) The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* **35**, 79-87
10. Kadonaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. (1987) Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51**, 1079-1090
11. Tadepally, H. D., Burger, G., and Aubry, M. (2008) Evolution of C2H2-zinc finger genes and subfamilies in mammals: species-specific duplication and loss of clusters, genes and effector domains. *BMC Evol Biol* **8**, 176
12. Schuh, R., Aicher, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schroder, C., Kemler, R., and et al. (1986) A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a Drosophila segmentation gene. *Cell* **47**, 1025-1032
13. Imataka, H., Sogawa, K., Yasumoto, K., Kikuchi, Y., Sasano, K., Kobayashi, A., Hayami, M., and Fujii-Kuriyama, Y. (1992) Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat P-4501A1 gene. *EMBO J* **11**, 3663-3671
14. Hagen, G., Muller, S., Beato, M., and Suske, G. (1992) Cloning by recognition site screening of two novel GT box binding proteins: a family of Sp1 related genes. *Nucleic Acids Res* **20**, 5519-5525

15. Kingsley, C., and Winoto, A. (1992) Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. *Mol Cell Biol* **12**, 4251-4261
16. Miller, I. J., and Bieker, J. J. (1993) A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. *Mol Cell Biol* **13**, 2776-2786
17. Cook, T., Gebelein, B., Mesa, K., Mladek, A., and Urrutia, R. (1998) Molecular cloning and characterization of TIEG2 reveals a new subfamily of transforming growth factor-beta-inducible Sp1-like zinc finger-encoding genes involved in the regulation of cell growth. *J Biol Chem* **273**, 25929-25936
18. Tachibana, I., Imoto, M., Adjei, P. N., Gores, G. J., Subramaniam, M., Spelsberg, T. C., and Urrutia, R. (1997) Overexpression of the TGFbeta-regulated zinc finger encoding gene, TIEG, induces apoptosis in pancreatic epithelial cells. *J Clin Invest* **99**, 2365-2374
19. Lomberk, G., and Urrutia, R. (2005) The family feud: turning off Sp1 by Sp1-like KLF proteins. *Biochem J* **392**, 1-11
20. Bieker, J. J. (2001) Kruppel-like factors: three fingers in many pies. *J Biol Chem* **276**, 34355-34358
21. Kaczynski, J., Cook, T., and Urrutia, R. (2003) Sp1- and Kruppel-like transcription factors. *Genome Biol* **4**, 206
22. Gibson, T. J., Postma, J. P., Brown, R. S., and Argos, P. (1988) A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc finger') common to many eukaryotic transcriptional regulatory proteins. *Protein Eng* **2**, 209-218
23. Klevit, R. E. (1991) Recognition of DNA by Cys2,His2 zinc fingers. *Science* **253**, 1367, 1393
24. Ruppert, J. M., Kinzler, K. W., Wong, A. J., Bigner, S. H., Kao, F. T., Law, M. L., Seuanez, H. N., O'Brien, S. J., and Vogelstein, B. (1988) The GLI-Kruppel family of human genes. *Mol Cell Biol* **8**, 3104-3113
25. Elrod-Erickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions. *Structure* **4**, 1171-1180
26. Kim, C. A., and Berg, J. M. (1996) A 2.2 A resolution crystal structure of a designed zinc finger protein bound to DNA. *Nat Struct Biol* **3**, 940-945
27. Crossley, M., Whitelaw, E., Perkins, A., Williams, G., Fujiwara, Y., and Orkin, S. H. (1996) Isolation and characterization of the cDNA encoding BKLF/TEF-2, a major CACCC-box-binding protein in erythroid cells and selected other cells. *Mol Cell Biol* **16**, 1695-1705
28. Matsumoto, N., Laub, F., Aldabe, R., Zhang, W., Ramirez, F., Yoshida, T., and Terada, M. (1998) Cloning the cDNA for a new human zinc finger protein defines a group of closely related Kruppel-like transcription factors. *J Biol Chem* **273**, 28229-28237
29. Shields, J. M., and Yang, V. W. (1998) Identification of the DNA sequence that interacts with the gut-enriched Kruppel-like factor. *Nucleic Acids Res* **26**, 796-802

30. Pandya, K., and Townes, T. M. (2002) Basic residues within the Kruppel zinc finger DNA binding domains are the critical nuclear localization determinants of EKLF/KLF-1. *J Biol Chem* **277**, 16304-16312
31. Shields, J. M., and Yang, V. W. (1997) Two potent nuclear localization signals in the gut-enriched Kruppel-like factor define a subfamily of closely related Kruppel proteins. *J Biol Chem* **272**, 18504-18507
32. Cook, T., Gebelein, B., Belal, M., Mesa, K., and Urrutia, R. (1999) Three conserved transcriptional repressor domains are a defining feature of the TIEG subfamily of Sp1-like zinc finger proteins. *J Biol Chem* **274**, 29500-29504
33. Ryu, S., Zhou, S., Ladurner, A. G., and Tjian, R. (1999) The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* **397**, 446-450
34. Zhang, J. S., Moncrieffe, M. C., Kaczynski, J., Ellenrieder, V., Prendergast, F. G., and Urrutia, R. (2001) A conserved alpha-helical motif mediates the interaction of Sp1-like transcriptional repressors with the corepressor mSin3A. *Mol Cell Biol* **21**, 5041-5049
35. Brey, C. W., Nelder, M. P., Hailemariam, T., Gaugler, R., and Hashmi, S. (2009) Kruppel-like family of transcription factors: an emerging new frontier in fat biology. *Int J Biol Sci* **5**, 622-636
36. Munoz-Descalzo, S., Terol, J., and Paricio, N. (2005) Cabut, a C2H2 zinc finger transcription factor, is required during Drosophila dorsal closure downstream of JNK signaling. *Dev Biol* **287**, 168-179
37. Rodriguez, I. Drosophila TIEG is a modulator of different signalling pathways involved in wing patterning and cell proliferation. *PLoS One* **6**, e18418
38. Munoz-Descalzo, S., Belacortu, Y., and Paricio, N. (2007) Identification and analysis of cabut orthologs in invertebrates and vertebrates. *Dev Genes Evol* **217**, 289-298
39. Guertin, D. A., Guntur, K. V., Bell, G. W., Thoreen, C. C., and Sabatini, D. M. (2006) Functional genomics identifies TOR-regulated genes that control growth and division. *Curr Biol* **16**, 958-970
40. Bulow, M. H., Aebersold, R., Pankratz, M. J., and Junger, M. A. The Drosophila FoxA ortholog Fork head regulates growth and gene expression downstream of Target of rapamycin. *PLoS One* **5**, e15171
41. Buck, A., Buchholz, M., Wagner, M., Adler, G., Gress, T., and Ellenrieder, V. (2006) The tumor suppressor KLF11 mediates a novel mechanism in transforming growth factor beta-induced growth inhibition that is inactivated in pancreatic cancer. *Mol Cancer Res* **4**, 861-872
42. Seetharam, A., Bai, Y., and Stuart, G. W. A survey of well conserved families of C2H2 zinc-finger genes in Daphnia. *BMC Genomics* **11**, 276
43. De Graeve, F., Smaldone, S., Laub, F., Mlodzik, M., Bhat, M., and Ramirez, F. (2003) Identification of the Drosophila progenitor of mammalian Kruppel-like factors 6 and 7 and a determinant of fly development. *Gene* **314**, 55-62

44. Oates, A. C., Pratt, S. J., Vail, B., Yan, Y., Ho, R. K., Johnson, S. L., Postlethwait, J. H., and Zon, L. I. (2001) The zebrafish klf gene family. *Blood* **98**, 1792-1801
45. Vo, N., and Goodman, R. H. (2001) CREB-binding protein and p300 in transcriptional regulation. *J Biol Chem* **276**, 13505-13508
46. Yang, X. J. (2004) Lysine acetylation and the bromodomain: a new partnership for signaling. *Bioessays* **26**, 1076-1087
47. Chen, H., Lin, R. J., Xie, W., Wilpitz, D., and Evans, R. M. (1999) Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase. *Cell* **98**, 675-686
48. Gu, J., Rubin, R. M., and Yuan, Z. M. (2001) A sequence element of p53 that determines its susceptibility to viral oncoprotein-targeted degradation. *Oncogene* **20**, 3519-3527
49. Turner, J., and Crossley, M. (1998) Cloning and characterization of mCtBP2, a co-repressor that associates with basic Kruppel-like factor and other mammalian transcriptional regulators. *EMBO J* **17**, 5129-5140
50. Kaczynski, J., Zhang, J. S., Ellenrieder, V., Conley, A., Duenes, T., Kester, H., van Der Burg, B., and Urrutia, R. (2001) The Sp1-like protein BTEB3 inhibits transcription via the basic transcription element box by interacting with mSin3A and HDAC-1 co-repressors and competing with Sp1. *J Biol Chem* **276**, 36749-36756
51. Kaczynski, J. A., Conley, A. A., Fernandez Zapico, M., Delgado, S. M., Zhang, J. S., and Urrutia, R. (2002) Functional analysis of basic transcription element (BTE)-binding protein (BTEB) 3 and BTEB4, a novel Sp1-like protein, reveals a subfamily of transcriptional repressors for the BTE site of the cytochrome P4501A1 gene promoter. *Biochem J* **366**, 873-882
52. Schuierer, M., Hilger-Eversheim, K., Dobner, T., Bosserhoff, A. K., Moser, M., Turner, J., Crossley, M., and Buettner, R. (2001) Induction of AP-2alpha expression by adenoviral infection involves inactivation of the AP-2rep transcriptional corepressor CtBP1. *J Biol Chem* **276**, 27944-27949
53. van Vliet, J., Turner, J., and Crossley, M. (2000) Human Kruppel-like factor 8: a CACCC-box binding protein that associates with CtBP and represses transcription. *Nucleic Acids Res* **28**, 1955-1962
54. Koipally, J., and Georgopoulos, K. (2000) Ikaros interactions with CtBP reveal a repression mechanism that is independent of histone deacetylase activity. *J Biol Chem* **275**, 19594-19602
55. Sundqvist, A., Sollerbrant, K., and Svensson, C. (1998) The carboxy-terminal region of adenovirus E1A activates transcription through targeting of a C-terminal binding protein-histone deacetylase complex. *FEBS Lett* **429**, 183-188
56. Meloni, A. R., Smith, E. J., and Nevins, J. R. (1999) A mechanism for Rb/p130-mediated transcription repression involving recruitment of the CtBP corepressor. *Proc Natl Acad Sci U S A* **96**, 9574-9579
57. Phippen, T. M., Sweigart, A. L., Moniwa, M., Krumm, A., Davie, J. R., and Parkhurst, S. M. (2000) Drosophila C-terminal binding protein functions as a

context-dependent transcriptional co-factor and interferes with both mad and groucho transcriptional repression. *J Biol Chem* **275**, 37628-37637
58. Sewalt, R. G., Gunster, M. J., van der Vlag, J., Satijn, D. P., and Otte, A. P. (1999) C-Terminal binding protein is a transcriptional repressor that interacts with a specific class of vertebrate Polycomb proteins. *Mol Cell Biol* **19**, 777-787
59. Turner, J., Nicholas, H., Bishop, D., Matthews, J. M., and Crossley, M. (2003) The LIM protein FHL3 binds basic Kruppel-like factor/Kruppel-like factor 3 and its co-repressor C-terminal-binding protein 2. *J Biol Chem* **278**, 12786-12795
60. Daftary, G. S., Lomberk, G. A., Buttar, N. S., Allen, T. W., Grzenda, A., Zhang, J., Zheng, Y., Mathison, A. J., Gada, R. P., Iovanna, J. L., Billadeau, D. D., Prendergast, F. G., and Urrutia, R. A Detailed Structural-Functional Analysis of Kruppel-Like Factor 16 (Klf16) Reveals Novel Mechanisms for Silencing Sp/Klf-Sites Involved in Metabolism and Endocrinology. *J Biol Chem*
61. Grzenda, A., Lomberk, G., Zhang, J. S., and Urrutia, R. (2009) Sin3: master scaffold and transcriptional corepressor. *Biochimica et biophysica acta* **1789**, 443-450
62. Chen, X., and Bieker, J. J. (2001) Unanticipated repression function linked to erythroid Kruppel-like factor. *Mol Cell Biol* **21**, 3118-3125
63. Truty, M. J., Lomberk, G., Fernandez-Zapico, M. E., and Urrutia, R. (2009) Silencing of the transforming growth factor-beta (TGFbeta) receptor II by Kruppel-like factor 14 underscores the importance of a negative feedback mechanism in TGFbeta signaling. *J Biol Chem* **284**, 6291-6300
64. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-124
65. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-120
66. Aasland, R., and Stewart, A. F. (1995) The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. *Nucleic Acids Res* **23**, 3168-3173
67. Chin, H. G., Esteve, P. O., Pradhan, M., Benner, J., Patnaik, D., Carey, M. F., and Pradhan, S. (2007) Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. *Nucleic Acids Res* **35**, 7313-7323
68. Fuks, F., Hurd, P. J., Deplus, R., and Kouzarides, T. (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* **31**, 2305-2312
69. Stephens, G. E., Slawson, E. E., Craig, C. A., and Elgin, S. C. (2005) Interaction of heterochromatin protein 2 with HP1 defines a novel HP1-binding domain. *Biochemistry* **44**, 13394-13403
70. Lomberk, G., Bensi, D., Fernandez-Zapico, M. E., and Urrutia, R. (2006) Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat Cell Biol* **8**, 407-415

71. Huang, Y., Myers, M. P., and Xu, R. M. (2006) Crystal structure of the HP1-EMSY complex reveals an unusual mode of HP1 binding. *Structure* **14**, 703-712
72. Lomberk, G., Mathison, A. J., Grzenda, A., Seo, S., Demars, C. J., Rizvi, S., Bonilla-Velez, J., Calvo, E., Fernandez-Zapico, M. E., Iovanna, J., Buttar, N. S., and Urrutia, R. Sequence-specific recruitment of heterochromatin protein 1 via interaction with Kruppel-like factor 11, a human transcription factor involved in tumor suppression and metabolic diseases. *J Biol Chem*
73. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039-1043
74. Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185-196
75. Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998) SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci* **54**, 80-93
76. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* **16**, 2893-2905
77. Vincenz, C., and Kerppola, T. K. (2008) Different polycomb group CBX family proteins associate with distinct regions of chromatin using nonhomologous protein sequences. *Proc Natl Acad Sci U S A* **105**, 16572-16577
78. Tonini, T., D'Andrilli, G., Fucito, A., Gaspa, L., and Bagella, L. (2008) Importance of Ezh2 polycomb protein in tumorigenesis process interfering with the pathway of growth suppressive key elements. *J Cell Physiol* **214**, 295-300
79. Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B. E., Bergman, Y., Simon, I., and Cedar, H. (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nat Genet* **39**, 232-236
80. Kondo, Y., Shen, L., Cheng, A. S., Ahmed, S., Boumber, Y., Charo, C., Yamochi, T., Urano, T., Furukawa, K., Kwabi-Addo, B., Gold, D. L., Sekido, Y., Huang, T. H., and Issa, J. P. (2008) Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nat Genet* **40**, 741-750
81. Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y., and Fuks, F. (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871-874
82. Cheng, A. S., Lau, S. S., Chen, Y., Kondo, Y., Li, M. S., Feng, H., Ching, A. K., Cheung, K. F., Wong, H. K., Tong, J. H., Jin, H., Choy, K. W., Yu, J., To, K. F., Wong, N., Huang, T. H., and Sung, J. J. EZH2-mediated concordant repression of Wnt antagonists promotes beta-catenin-dependent hepatocarcinogenesis. *Cancer Res* **71**, 4028-4039

83. Dellino, G. I., Schwartz, Y. B., Farkas, G., McCabe, D., Elgin, S. C., and Pirrotta, V. (2004) Polycomb silencing blocks transcription initiation. *Mol Cell* **13**, 887-893
84. Francis, N. J., Saurin, A. J., Shao, Z., and Kingston, R. E. (2001) Reconstitution of a functional core polycomb repressive complex. *Mol Cell* **8**, 545-556
85. Brown, J. L., and Kassis, J. A. Spps, a Drosophila Sp1/KLF family member, binds to PREs and is required for PRE activity late in development. *Development* **137**, 2597-2602
86. Xiong, Y., Khanna, S., Grzenda, A. L., Sarmento, O. F., Svingen, P. A., Lomberk, G. A., Urrutia, R. A., and Faubion, W. A., Jr. (2012) Polycomb antagonizes p300/CREB-binding protein-associated factor to silence FOXP3 in a Kruppel-like factor-dependent manner. *The Journal of biological chemistry* **287**, 34372-34385
87. Song, A., Chen, Y. F., Thamatrakoln, K., Storm, T. A., and Krensky, A. M. (1999) RFLAT-1: a new zinc finger transcription factor that activates RANTES gene expression in T lymphocytes. *Immunity* **10**, 93-103
88. Scohy, S., Gabant, P., Van Reeth, T., Hertveldt, V., Dreze, P. L., Van Vooren, P., Riviere, M., Szpirer, J., and Szpirer, C. (2000) Identification of KLF13 and KLF14 (SP6), novel members of the SP/XKLF transcription factor family. *Genomics* **70**, 93-101
89. Martin, K. M., Cooper, W. N., Metcalfe, J. C., and Kemp, P. R. (2000) Mouse BTEB3, a new member of the basic transcription element binding protein (BTEB) family, activates expression from GC-rich minimal promoter regions. *Biochem J* **345 Pt 3**, 529-533
90. McConnell, B. B., and Yang, V. W. Mammalian Kruppel-like factors in health and diseases. *Physiol Rev* **90**, 1337-1381
91. Ellenrieder, V., Buck, A., Harth, A., Jungert, K., Buchholz, M., Adler, G., Urrutia, R., and Gress, T. M. (2004) KLF11 mediates a critical mechanism in TGF-beta signaling that is inactivated by Erk-MAPK in pancreatic cancer cells. *Gastroenterology* **127**, 607-620
92. Buttar, N. S., DeMars, C. J., Lomberk, G., Rizvi, S., Bonilla-Velez, J., Achra, S., Rashtak, S., Wang, K. K., Fernandez-Zapico, M. E., and Urrutia, R. Distinct role of Kruppel-like factor 11 in the regulation of prostaglandin E2 biosynthesis. *J Biol Chem* **285**, 11433-11444
93. Ellenrieder, V., Zhang, J. S., Kaczynski, J., and Urrutia, R. (2002) Signaling disrupts mSin3A binding to the Mad1-like Sin3-interacting domain of TIEG2, an Sp1-like repressor. *EMBO J* **21**, 2451-2460
94. Casares, F., and Sanchez-Herrero, E. (1995) Regulation of the infraabdominal regions of the bithorax complex of Drosophila by gap genes. *Development* **121**, 1855-1866
95. Brown, J. L., Grau, D. J., DeVido, S. K., and Kassis, J. A. (2005) An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the Drosophila engrailed gene. *Nucleic Acids Res* **33**, 5181-5189
96. Venuprasad, K., Huang, H., Harada, Y., Elly, C., Subramaniam, M., Spelsberg, T., Su, J., and Liu, Y. C. (2008) The E3 ubiquitin ligase Itch regulates expression

of transcription factor Foxp3 and airway inflammation by enhancing the function of transcription factor TIEG1. *Nat Immunol* **9**, 245-253
97. Seo, S., Lomberk, G., Mathison, A., Buttar, N., Podratz, J., Calvo, E., Iovanna, J., Brimijoin, S., Windebank, A., and Urrutia, R. (2012) Kruppel-like factor 11 differentially couples to histone acetyltransferase and histone methyltransferase chromatin remodeling pathways to transcriptionally regulate dopamine D2 receptor in neuronal cells. *The Journal of biological chemistry* **287**, 12723-12735
98. Singh, M., and Maitra, A. (2007) Precursor lesions of pancreatic cancer: molecular pathology and clinical implications. *Pancreatology* **7**, 9-19
99. Ellenrieder, V. (2008) TGFbeta regulated gene expression by Smads and Sp1/KLF-like transcription factors in cancer. *Anticancer Res* **28**, 1531-1539
100. Fernandez-Zapico, M. E., Mladek, A., Ellenrieder, V., Folch-Puy, E., Miller, L., and Urrutia, R. (2003) An mSin3A interaction domain links the transcriptional activity of KLF11 with its role in growth regulation. *EMBO J* **22**, 4748-4758
101. Kawai-Kowase, K., Kurabayashi, M., Hoshino, Y., Ohyama, Y., and Nagai, R. (1999) Transcriptional activation of the zinc finger transcription factor BTEB2 gene by Egr-1 through mitogen-activated protein kinase pathways in vascular smooth muscle cells. *Circ Res* **85**, 787-795
102. Nandan, M. O., Yoon, H. S., Zhao, W., Ouko, L. A., Chanchevalap, S., and Yang, V. W. (2004) Kruppel-like factor 5 mediates the transforming activity of oncogenic H-Ras. *Oncogene* **23**, 3404-3413
103. Chanchevalap, S., Nandan, M. O., McConnell, B. B., Charrier, L., Merlin, D., Katz, J. P., and Yang, V. W. (2006) Kruppel-like factor 5 is an important mediator for lipopolysaccharide-induced proinflammatory response in intestinal epithelial cells. *Nucleic Acids Res* **34**, 1216-1223
104. Pei, H., Yao, Y., Yang, Y., Liao, K., and Wu, J. R. Kruppel-like factor KLF9 regulates PPARgamma transactivation at the middle stage of adipogenesis. *Cell Death Differ* **18**, 315-327
105. Lee, H., Kang, R., Kim, Y. S., Chung, S. I., and Yoon, Y. Platycodin D inhibits adipogenesis of 3T3-L1 cells by modulating Kruppel-like factor 2 and peroxisome proliferator-activated receptor gamma. *Phytother Res* **24 Suppl 2**, S161-167
106. Lomberk, G., and Urrutia, R. (2005) The family feud: turning off Sp1 by Sp1-like KLF proteins. *Biochem. J.* **392**, 1-11
107. Munoz-Descalzo, S., Terol, J., and Paricio, N. (2005) Cabut, a C2H2 zinc finger transcription factor, is required during Drosophila dorsal closure downstream of JNK signaling. *Developmental Biology* **287**, 168-179
108. Bonnefond, A., Lomberk, G., Buttar, N., Busiah, K., Vaillant, E., Lobbens, S., Yengo, L., Dechaume, A., Mignot, B., Simon, A., Scharfmann, R., Neve, B., Tanyolac, S., Hodoglugil, U., Pattou, F., Cave, H., Iovanna, J., Stein, R., Polak, M., Vaxillaire, M., Froguel, P., and Urrutia, R. (2011) Disruption of a novel KLF-p300-regulated pathway for insulin biosynthesis revealed by studies of the c.-331 INS mutation found in neonatal diabetes mellitus. *J. Biol. Chem.* **286**, 28414-28424

109. Neve, B., Fernandez-Zapico, M., Ashkenazi-Katalan, V., Dina, C., Hamid, Y., Joly, E., Vaillant, E., Benmezroua, Y., Durand, E., Bakaher, N., Delannoy, V., Vaxillaire, M., Cook, T., Dallinga-Thie, G., Jansen, H., Charles, M., Clement, K., Galan, P., Hercberg, S., Helbecque, N., Charpentier, G., Prentki, M., Hansen, T., Pedersen, O., Urrutia, R., Melloul, D., and Froguel, P. (2005) Role of transcription factor KLF11 and its diabetes-associated gene variants in pancreatic beta cell function. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4807-4812
110. Buttar, N. S., DeMars, C. J., Lomberk, G., Rizvi, S., Bonilla-Velez, J., Achra, S., Rashtak, S., Wang, K. K., Fernandez-Zapico, M. E., and Urrutia, R. (2010) Distinct Role of Kruppel-like Factor 11 in the Regulation of Prostaglandin E2 Biosynthesis. *Journal of Biological Chemistry* **285**, 11433-11444
111. Cao, S., Fernandez-Zapico, M. E., Jin, D., Puri, V., Cook, T. A., Lerman, L. O., Zhu, X.-Y., Urrutia, R., and Shah, V. (2005) KLF11-mediated Repression Antagonizes Sp1/Sterol-responsive Element-binding protein-induced Transcriptional Activation of Caveolin-1 in Response to Cholesterol Signaling. *Journal of Biological Chemistry* **280**, 1901-1910
112. Grunewald, M., Johnson, S., Lu, D., Wang, Z., Lomberk, G., Albert, P. R., Stockmeier, C. A., Meyer, J. H., Urrutia, R., Miczek, K. A., Austin, M. C., Wang, J., Paul, I. A., Woolverton, W. L., Seo, S., Sittman, D. B., and Ou, X.-M. (2012) Mechanistic Role for a Novel Glucocorticoid-KLF11 (TIEG2) Protein Pathway in Stress-induced Monoamine Oxidase A Expression. *Journal of Biological Chemistry* **287**, 24195-24206
113. Ou, X.-M., Chen, K., and Shih, J. C. (2004) Dual Functions of Transcription Factors, Transforming Growth Factor-Œ≤-inducible Early Gene (TIEG)2 and Sp3, Are Mediated by CACCC Element and Sp1 Sites of Human Monoamine Oxidase (MAO) B Gene. *Journal of Biological Chemistry* **279**, 21021-21028
114. Yamamoto, K.-i., Sakaguchi, M., Medina, R. J., Niida, A., Sakaguchi, Y., Miyazaki, M., Kataoka, K., and Huh, N.-h. (2010) Transcriptional regulation of a brown adipocyte-specific gene, UCP1, by KLF11 and KLF15. *Biochemical and Biophysical Research Communications* **400**, 175-180
115. Fernandez-Zapico, M. E., van Velkinburgh, J. C., Gutierrez-Aguilar, R., Neve, B., Froguel, P., Urrutia, R., and Stein, R. (2009) MODY7 Gene, KLF11, Is a Novel p300-dependent Regulator of Pdx-1 (MODY4) Transcription in Pancreatic Islet beta Cells. *J. Biol. Chem.* **284**, 36482-36490
116. Lomberk, G., Mathison, A. J., Grzenda, A., Seo, S., DeMars, C. J., Rizvi, S., Bonilla-Velez, J., Calvo, E., Fernandez-Zapico, M. E., Iovanna, J., Buttar, N. S., and Urrutia, R. (2012) Sequence-specific Recruitment of Heterochromatin Protein 1 via Interaction with Kruppel-like Factor 11, a Human Transcription Factor Involved in Tumor Suppression and Metabolic Diseases. *Journal of Biological Chemistry* **287**, 13026-13039
117. Cook, T., Gebelein, B., Belal, M., Mesa, K., and Urrutia, R. (1999) Three conserved transcriptional repressor domains are a defining feature of the TIEG subfamily of Sp1-like zinc finger proteins. *J. Biol. Chem.* **274**, 29500-29504

118. Mathison, A., Grzenda, A., Lomberk, G., Velez, G., Buttar, N., Tietz, P., Hendrickson, H., Liebl, A., Xiong, Y. Y., Gores, G., Fernandez-Zapico, M., Larusso, N. F., Faubion, W., Shah, V. H., and Urrutia, R. (2013) Role for Kruppel-like transcription factor 11 in mesenchymal cell function and fibrosis. *PloS one* **8**, e75311
119. Gebelein, B., Fernandez-Zapico, M., Imoto, M., and Urrutia, R. (1998) KRAB-independent suppression of neoplastic cell growth by the novel zinc finger transcription factor KS1. *J. Clin. Invest.* **102**, 1911-1919
120. Cheng, Z.-J., Harikumar, K. G., Holicky, E. L., and Miller, L. J. (2003) Heterodimerization of Type A and B Cholecystokinin Receptors Enhance Signaling and Promote Cell Growth. *Journal of Biological Chemistry* **278**, 52972-52979
121. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132
122. Vila, J. A., Baldoni, H. A., Ripoll, D. R., Ghosh, A., and Scheraga, H. A. (2004) Polyproline II helix conformation in a proline-rich environment: a theoretical study. *Biophys J* **86**, 731-742
123. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797
124. Mathison, A., Liebl, A., Bharucha, J., Mukhopadhyay, D., Lomberk, G., Shah, V., and Urrutia, R. (2010) Pancreatic stellate cell models for transcriptional studies of desmoplasia-associated genes. *Pancreatology.* **10**, 505-516
125. Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264
126. Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001) Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* **125**, 279-284
127. Smyth, G. K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3
128. Wettenhall, J. M., Simpson, K. M., Satterley, K., and Smyth, G. K. (2006) affylmGUI: a graphical user interface for linear modeling of single channel microarray data. *Bioinformatics* **22**, 897-899
129. Seo, S., Lomberk, G., Mathison, A., Buttar, N., Podratz, J., Calvo, E., Iovanna, J., Brimijoin, S., Windebank, A., and Urrutia, R. (2012) Kruppel-like Factor 11 Differentially Couples to Histone Acetyltransferase and Histone Methyltransferase Chromatin Remodeling Pathways to Transcriptionally Regulate Dopamine D2 Receptor in Neuronal Cells. *Journal of Biological Chemistry* **287**, 12723-12735
130. Hodel, A. E., Harreman, M. T., Pulliam, K. F., Harben, M. E., Holmes, J. S., Hodel, M. R., Berland, K. M., and Corbett, A. H. (2006) Nuclear Localization Signal Receptor Affinity Correlates with in Vivo Localization in Saccharomyces cerevisiae. *Journal of Biological Chemistry* **281**, 23545-23556

131. Fernandez-Zapico, M. E., Lomberk, G. A., Tsuji, S., DeMars, C. J., Bardsley, M. R., Lin, Y. H., Almada, L. L., Han, J. J., Mukhopadhyay, D., Ordog, T., Buttar, N. S., and Urrutia, R. (2011) A functional family-wide screening of SP/KLF proteins identifies a subset of suppressors of KRAS-mediated cell growth. *Biochem J* **435**, 529-537
132. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *The FASEB Journal* **14**, 231-241
133. Sudol, M., Chen, H. I., Bougeret, C., Einbond, A., and Bork, P. (1995) Characterization of a novel protein-binding module ,Äî the WW domain. *FEBS Letters* **369**, 67-71
134. Freund, C., Schmalz, H. G., Sticht, J., and Kühne, R. (2008) Proline-Rich Sequence Recognition Domains (PRD): Ligands, Function and Inhibition
Protein-Protein Interactions as New Drug Targets. (Klussmann, E., and Scott, J. eds.), Springer Berlin Heidelberg. pp 407-429
135. Venuprasad, K., Huang, H., Harada, Y., Elly, C., Subramaniam, M., Spelsberg, T., Su, J., and Liu, Y.-C. (2008) The E3 ubiquitin ligase Itch regulates expression of transcription factor Foxp3 and airway inflammation by enhancing the function of transcription factor TIEG1. *Nat Immunol* **9**, 245-253
136. Conkright, M. D., Wani, M. A., and Lingrel, J. B. (2001) Lung Kruppel-like Factor Contains an Autoinhibitory Domain That Regulates Its Transcriptional Activation by Binding WWP1, an E3 Ubiquitin Ligase. *Journal of Biological Chemistry* **276**, 29299-29306
137. Chen, C., Sun, X., Guo, P., Dong, X.-Y., Sethi, P., Cheng, X., Zhou, J., Ling, J., Simons, J. W., Lingrel, J. B., and Dong, J.-T. (2005) Human Kruppel-like Factor 5 Is a Target of the E3 Ubiquitin Ligase WWP1 for Proteolysis in Epithelial Cells. *Journal of Biological Chemistry* **280**, 41553-41561
138. Zhi, X., Zhao, D., Zhou, Z., Liu, R., and Chen, C. (2012) YAP Promotes Breast Cell Proliferation and Survival Partially through Stabilizing the KLF5 Transcription Factor. *The American Journal of Pathology* **180**, 2452-2461
139. Daftary, G. S., Lomberk, G. A., Buttar, N. S., Allen, T. W., Grzenda, A., Zhang, J., Zheng, Y., Mathison, A. J., Gada, R. P., Calvo, E., Iovanna, J. L., Billadeau, D. D., Prendergast, F. G., and Urrutia, R. (2012) Detailed Structural-Functional Analysis of the Kruppel-like Factor 16 (KLF16) Transcription Factor Reveals Novel Mechanisms for Silencing Sp/KLF Sites Involved in Metabolism and Endocrinology. *Journal of Biological Chemistry* **287**, 7010-7025
140. Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E., and Sigler, P. B. (1996) Crystal structure of a GA protein [beta][gamma]dimer at 2.1 A resolution. *Nature* **379**, 369-374
141. Weinstein, L. S., Chen, M., Xie, T., and Liu, J. (2006) Genetic diseases associated with heterotrimeric G proteins. *Trends in pharmacological sciences* **27**, 260-266
142. Smrcka, A., DM, L., and Dessal, A. (2008) G protein betagamma subunits as targets for small molecule therapeutic development. *Comb Chem High Throughput Screen.* **11**, 382-395.

143. Dupre, D. J., Robitaille, M., Rebois, R. V., and Hebert, T. E. (2009) The Role of Gbetagamma Subunits in the Organization, Assembly, and Function of GPCR Signaling Complexes. *Annual Review of Pharmacology and Toxicology* **49**, 31-56
144. Ryden, M., Faulds, G., Hoffstedt, J., Wennlund, A., and Arner, P. (2002) Effect of the (C825T) Gbeta3 Polymorphism on Adrenoceptor-Mediated Lipolysis in Human Fat Cells. *Diabetes* **51**, 1601-1608
145. Kino, T., Tiulpakov, A., Ichijo, T., Chheng, L., Kozasa, T., and Chrousos, G. P. (2005) G protein beta interacts with the glucocorticoid receptor and suppresses its transcriptional activity in the nucleus. *The Journal of Cell Biology* **169**, 885-896
146. Robitaille, M., Gora, S., Wang, Y., Goupil, E., Petrin, D., Del Duca, D., Villeneuve, L. R., Allen, B. G., Laporte, S. A., Bernard, D. J., and Hebert, T. E. (2010) Gbetagamma is a negative regulator of AP-1 mediated transcription. *Cellular Signalling* **22**, 1254-1266
147. Tallman, J. F., Smith, C. C., and Henneberry, R. C. (1977) Induction of functional beta-adrenergic receptors in HeLa cells. *Proceedings of the National Academy of Sciences* **74**, 873-877
148. Al-Nakkash, L., Simmons, N., Lingard, J., and Argent, B. (1996) Adenylate cyclase activity in human pancreatic adenocarcinoma cell lines. *International Journal of Gastrointestinal Cancer* **19**, 39-47
149. Wank, S. A., Pisegna, J. R., and De Weerth, A. (1994) Cholecystokinin Receptor Family. *Annals of the New York Academy of Sciences* **713**, 49-66
150. Civelek, M., and Lusis, A. J. (2011) Conducting the metabolic syndrome orchestra. *Nat Genet* **43**, 506-508
151. Voight, B. F., Scott, L. J., Steinthorsdottir, V., Morris, A. P., Dina, C., Welch, R. P., Zeggini, E., Huth, C., Aulchenko, Y. S., Thorleifsson, G., McCulloch, L. J., Ferreira, T., Grallert, H., Amin, N., Wu, G., Willer, C. J., Raychaudhuri, S., McCarroll, S. A., Langenberg, C., Hofmann, O. M., Dupuis, J., Qi, L., Segre, A. V., van Hoek, M., Navarro, P., Ardlie, K., Balkau, B., Benediktsson, R., Bennett, A. J., Blagieva, R., Boerwinkle, E., Bonnycastle, L. L., Bostrom, K. B., Bravenboer, B., Bumpstead, S., Burtt, N. P., Charpentier, G., Chines, P. S., Cornelis, M., Couper, D. J., Crawford, G., Doney, A. S. F., Elliott, K. S., Elliott, A. L., Erdos, M. R., Fox, C. S., Franklin, C. S., Ganser, M., Gieger, C., Grarup, N., Green, T., Griffin, S., Groves, C. J., Guiducci, C., Hadjadj, S., Hassanali, N., Herder, C., Isomaa, B., Jackson, A. U., Johnson, P. R. V., Jorgensen, T., Kao, W. H. L., Klopp, N., Kong, A., Kraft, P., Kuusisto, J., Lauritzen, T., Li, M., Lieverse, A., Lindgren, C. M., Lyssenko, V., Marre, M., Meitinger, T., Midthjell, K., Morken, M. A., Narisu, N., Nilsson, P., Owen, K. R., Payne, F., Perry, J. R. B., Petersen, A.-K., Platou, C., Proenca, C., Prokopenko, I., Rathmann, W., Rayner, N. W., Robertson, N. R., Rocheleau, G., Roden, M., Sampson, M. J., Saxena, R., Shields, B. M., Shrader, P., Sigurdsson, G., Sparso, T., Strassburger, K., Stringham, H. M., Sun, Q., Swift, A. J., Thorand, B., Tichet, J., Tuomi, T., van Dam, R. M., van Haeften, T. W., van Herpt, T., van Vliet-Ostaptchouk, J. V., Walters, G. B., Weedon, M. N., Wijmenga, C., Witteman, J., Bergman, R. N., Cauchi, S., Collins, F. S., Gloyn, A. L., Gyllensten, U., Hansen, T., Hide, W. A.,

Hitman, G. A., Hofman, A., Hunter, D. J., Hveem, K., Laakso, M., Mohlke, K. L., Morris, A. D., Palmer, C. N. A., Pramstaller, P. P., Rudan, I., Sijbrands, E., Stein, L. D., Tuomilehto, J., Uitterlinden, A., Walker, M., Wareham, N. J., Watanabe, R. M., Abecasis, G. R., Boehm, B. O., Campbell, H., Daly, M. J., Hattersley, A. T., Hu, F. B., Meigs, J. B., Pankow, J. S., Pedersen, O., Wichmann, H. E., Barroso, I., Florez, J. C., Frayling, T. M., Groop, L., Sladek, R., Thorsteinsdottir, U., Wilson, J. F., Illig, T., Froguel, P., van Duijn, C. M., Stefansson, K., Altshuler, D., Boehnke, M., and McCarthy, M. I. (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* **42**, 579-589

152. Small, K., Hedman, A., Grundberg, E., Nica, A., Thorleifsson, G., Kong, A., Thorsteindottir, U., Shin, S., Richards, H., the GIANT Consortium, the MAGIC Investigators, the DIAGRAM Consortium, Soranzo, N., Ahmadi, K., Lindgren, C., Stefansson, K., Dermitzakis, E., Deloukas, P., Spector, T., McCarthy, M., and Consortium, f. t. M. (2011) Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nat Genet* **43**, 561-564
153. Migliori, V., Mapelli, M., and Guccione, E. (2012) On WD40 proteins: propelling our knowledge of transcriptional control? *Epigenetics.* **7**, 815-822. doi: 810.4161/epi.21140. Epub 22012 Jul 21119.
154. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**, 1025-1040
155. Del Vecchio, I., Zuccotti, A., Pisano, F., Canneva, F., Lenzken, S. C., Rousset, F., Corsini, E., Govoni, S., and Racchi, M. (2009) Functional mapping of the promoter region of the GNB2L1 human gene coding for RACK1 scaffold protein. *Gene* **430**, 17-29
156. Lomberk, G., and Urrutia, R. (2005) The family feud: turning off Sp1 by Sp1-like KLF proteins. *Biochem J* **392**, 1-11
157. Bonnefond, A., Lomberk, G., Buttar, N., Busiah, K., Vaillant, E., Lobbens, S., Yengo, L., Dechaume, A., Mignot, B., Simon, A., Scharfmann, R., Neve, B., Tanyolac, S., Hodoglugil, U., Pattou, F., Cave, H., Iovanna, J., Stein, R., Polak, M., Vaxillaire, M., Froguel, P., and Urrutia, R. (2011) Disruption of a novel Kruppel-like transcription factor p300-regulated pathway for insulin biosynthesis revealed by studies of the c.-331 INS mutation found in neonatal diabetes mellitus. *J Biol Chem* **286**, 28414-28424
158. Fernandez-Zapico, M. E., van Velkinburgh, J. C., Gutierrez-Aguilar, R., Neve, B., Froguel, P., Urrutia, R., and Stein, R. (2009) MODY7 gene, KLF11, is a novel p300-dependent regulator of Pdx-1 (MODY4) transcription in pancreatic islet beta cells. *J Biol Chem* **284**, 36482-36490
159. Grzenda, A., Lomberk, G., Zhang, J. S., and Urrutia, R. (2009) Sin3: master scaffold and transcriptional corepressor. *Biochim Biophys Acta* **1789**, 443-450
160. Grzenda, A., Lomberk, G., Svingen, P., Mathison, A., Calvo, E., Iovanna, J., Xiong, Y., Faubion, W., and Urrutia, R. (2013) Functional characterization of EZH2beta reveals the increased complexity of EZH2 isoforms involved in the regulation of mammalian gene expression. *Epigenetics Chromatin* **6**, 3

161. Lomberk, G., Mathison, A. J., Grzenda, A., Seo, S., DeMars, C. J., Rizvi, S., Bonilla-Velez, J., Calvo, E., Fernandez-Zapico, M. E., Iovanna, J., Buttar, N. S., and Urrutia, R. (2012) Sequence-specific recruitment of heterochromatin protein 1 via interaction with Kruppel-like factor 11, a human transcription factor involved in tumor suppression and metabolic diseases. *J Biol Chem* **287**, 13026-13039
162. Seo, S., Lomberk, G., Mathison, A., Buttar, N., Podratz, J., Calvo, E., Iovanna, J., Brimijoin, S., Windebank, A., and Urrutia, R. (2012) Kruppel-like factor 11 differentially couples to histone acetyltransferase and histone methyltransferase chromatin remodeling pathways to transcriptionally regulate dopamine D2 receptor in neuronal cells. *J Biol Chem* **287**, 12723-12735
163. Fan, Y., Guo, Y., Zhang, J., Subramaniam, M., Song, C. Z., Urrutia, R., and Chen, Y. E. (2012) Kruppel-like factor-11, a transcription factor involved in diabetes mellitus, suppresses endothelial cell activation via the nuclear factor-kappaB signaling pathway. *Arterioscler Thromb Vasc Biol* **32**, 2981-2988
164. Fernandez-Zapico, M. E., Lomberk, G. A., Tsuji, S., DeMars, C. J., Bardsley, M. R., Lin, Y. H., Almada, L. L., Han, J. J., Mukhopadhyay, D., Ordog, T., Buttar, N. S., and Urrutia, R. (2011) A functional family-wide screening of SP/KLF proteins identifies a subset of suppressors of KRAS-mediated cell growth. *Biochem J* **435**, 529-537
165. Buttar, N. S., DeMars, C. J., Lomberk, G., Rizvi, S., Bonilla-Velez, J., Achra, S., Rashtak, S., Wang, K. K., Fernandez-Zapico, M. E., and Urrutia, R. (2010) Distinct role of Kruppel-like factor 11 in the regulation of prostaglandin E2 biosynthesis. *J Biol Chem* **285**, 11433-11444
166. Neve, B., Fernandez-Zapico, M. E., Ashkenazi-Katalan, V., Dina, C., Hamid, Y. H., Joly, E., Vaillant, E., Benmezroua, Y., Durand, E., Bakaher, N., Delannoy, V., Vaxillaire, M., Cook, T., Dallinga-Thie, G. M., Jansen, H., Charles, M. A., Clement, K., Galan, P., Hercberg, S., Helbecque, N., Charpentier, G., Prentki, M., Hansen, T., Pedersen, O., Urrutia, R., Melloul, D., and Froguel, P. (2005) Role of transcription factor KLF11 and its diabetes-associated gene variants in pancreatic beta cell function. *Proc Natl Acad Sci U S A* **102**, 4807-4812
167. Pang, Y. P., Kumar, G. A., Zhang, J. S., and Urrutia, R. (2003) Differential binding of Sin3 interacting repressor domains to the PAH2 domain of Sin3A. *FEBS Lett* **548**, 108-112
168. Fernandez-Zapico, M. E., Mladek, A., Ellenrieder, V., Folch-Puy, E., Miller, L., and Urrutia, R. (2003) An mSin3A interaction domain links the transcriptional activity of KLF11 with its role in growth regulation. *Embo J* **22**, 4748-4758
169. Kaczynski, J. A., Conley, A. A., Fernandez Zapico, M., Delgado, S. M., Zhang, J. S., and Urrutia, R. (2002) Functional analysis of basic transcription element (BTE)-binding protein (BTEB) 3 and BTEB4, a novel Sp1-like protein, reveals a subfamily of transcriptional repressors for the BTE site of the cytochrome P4501A1 gene promoter. *Biochem J* **366**, 873-882

170. Ellenrieder, V., Zhang, J. S., Kaczynski, J., and Urrutia, R. (2002) Signaling disrupts mSin3A binding to the Mad1-like Sin3-interacting domain of TIEG2, an Sp1-like repressor. *Embo J* **21**, 2451-2460
171. Zhang, J. S., Moncrieffe, M. C., Kaczynski, J., Ellenrieder, V., Prendergast, F. G., and Urrutia, R. (2001) A conserved alpha-helical motif mediates the interaction of Sp1-like transcriptional repressors with the corepressor mSin3A. *Mol Cell Biol* **21**, 5041-5049
172. Kaczynski, J., Zhang, J. S., Ellenrieder, V., Conley, A., Duenes, T., Kester, H., van Der Burg, B., and Urrutia, R. (2001) The Sp1-like protein BTEB3 inhibits transcription via the basic transcription element box by interacting with mSin3A and HDAC-1 co-repressors and competing with Sp1. *J Biol Chem* **276**, 36749-36756
173. Lomberk, G., Grzenda, A., Mathison, A., Escande, C., Zhang, J. S., Calvo, E., Miller, L. J., Iovanna, J., Chini, E. N., Fernandez-Zapico, M. E., and Urrutia, R. (2013) Kruppel-like Factor 11 Regulates the Expression of Metabolic Genes via an Evolutionarily Conserved Protein-Interaction Domain Functionally Disrupted in Maturity Onset Diabetes of the Young. *J Biol Chem*
174. Mathison, A., Liebl, A., Bharucha, J., Mukhopadhyay, D., Lomberk, G., Shah, V., and Urrutia, R. (2010) Pancreatic stellate cell models for transcriptional studies of desmoplasia-associated genes. *Pancreatology* **10**, 505-516
175. Xiong, Y., Khanna, S., Grzenda, A. L., Sarmento, O. F., Svingen, P. A., Lomberk, G. A., Urrutia, R. A., and Faubion, W. A., Jr. (2012) Polycomb antagonizes p300/CREB-binding protein-associated factor to silence FOXP3 in a Kruppel-like factor-dependent manner. *J Biol Chem* **287**, 34372-34385
176. Xiao, T. Z., Bhatia, N., Urrutia, R., Lomberk, G. A., Simpson, A., and Longley, B. J. (2011) MAGE I transcription factors regulate KAP1 and KRAB domain zinc finger transcription factor mediated gene repression. *PLoS One* **6**, e23747
177. Daftary, G. S., Zheng, Y., Tabbaa, Z. M., Schoolmeester, J. K., Gada, R. P., Grzenda, A. L., Mathison, A. J., Keeney, G. L., Lomberk, G. A., and Urrutia, R. (2013) A novel role of the Sp/KLF transcription factor KLF11 in arresting progression of endometriosis. *PloS one* **8**, e60165
178. Yin, K. J., Fan, Y., Hamblin, M., Zhang, J., Zhu, T., Li, S., Hawse, J. R., Subramaniam, M., Song, C. Z., Urrutia, R., Lin, J. D., and Chen, Y. E. (2013) KLF11 mediates PPARgamma cerebrovascular protection in ischaemic stroke. *Brain* **136**, 1274-1287
179. Gebelein, B., and Urrutia, R. (2001) Sequence-specific transcriptional repression by KS1, a multiple-zinc-finger-Kruppel-associated box protein. *Mol Cell Biol* **21**, 928-939
180. Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264
181. Prasanth, S. G., Prasanth, K. V., Siddiqui, K., Spector, D. L., and Stillman, B. (2004) Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. *Embo J* **23**, 2651-2663

182. Kondo, Y., Shen, L., Ahmed, S., Boumber, Y., Sekido, Y., Haddad, B. R., and Issa, J. P. (2008) Downregulation of histone H3 lysine 9 methyltransferase G9a induces centrosome disruption and chromosome instability in cancer cells. *PLoS One* **3**, e2037
183. Wettenhall, J. M., Simpson, K. M., Satterley, K., and Smyth, G. K. (2006) affylmGUI: a graphical user interface for linear modeling of single channel microarray data. *Bioinformatics* **22**, 897-899