REGULATORY LINKAGE AND COORDINATED EXPRESSION OF THE *CHKB* AND *CPT1B* GENES: IMPLICATIONS FOR AN ELEVATED METABOLIC CAPACITY AND THE ORIGIN OF BROWN ADIPOSE TISSUE

BY

BHAVIN V. PATEL

JULY, 2023

DIRECTOR OF DISSERTATION: BRIAN M. SHEWCHUK, PHD.

MAJOR DEPARTMENT: BIOCHEMISTRY AND MOLECULAR BIOLOGY

The emergence of endothermy was an important event in the evolution of vertebrates. Specifically, the acquisition of BAT was a key achievement for the radiation and diversification of eutherian (placental) mammals. However, the molecular mechanism behind the development of this thermogenic tissue has not been thoroughly characterized. The theory of the UCP1 gene being the key impetus for BAT formation has been challenged considering the UCP1 gene has been found in several ray-finned fish species that diverged early in the timescale of vertebrate evolution. Additionally, the UCP1 gene has been found in several non-placental mammals from the marsupial and monotreme clades, however finding BAT depots in these species has proven unsuccessful. Mitochondrial fatty acid oxidation (FAO) contributes to the proton motive force that drives ATP synthesis in many mammalian tissues. In eutherian mammals, brown adipose tissue (BAT) can also dissipate this proton gradient through uncoupling protein 1 (UCP1) to generate heat. An essential step in FAO is the transport of cytoplasmic long-chain acyl-coenzyme A (acyl-CoA) into the mitochondrial matrix, which requires the action of carnitine palmitoyltransferase 1B (CPT1B) in striated muscle and BAT. In eutherians, the *CPT1B* gene is closely linked to the *choline kinase beta* (*CHKB*) gene, which is transcribed from the same DNA strand and terminates just upstream of *CPT1B*. CHKB is a rate-limiting enzyme in the synthesis

of phosphatidylcholine (PC), a predominant mitochondrial membrane phospholipid, suggesting that the coordinated expression of *CHKB* and *CPT1B* may cooperatively enhance mitochondrial FAO. The present findings show that transcription of the eutherian *CHKB* and *CPT1B* genes is linked within a unitary epigenetic domain targeted to the *CHKB* gene, and that that this regulatory linkage appears to have resulted from an intergenic deletion in eutherians that significantly altered the distribution of *CHKB* and *CPT1B* expression. Informed by the timing of this event relative to the emergence of BAT, the phylogeny of *CHKB-CPT1B* synteny, and the insufficiency of UCP1 to account for eutherian BAT, these data support a mechanism for the emergence of BAT based on the acquisition of a novel capacity for adipocyte FAO in a background of extant UCP1. Additionally, this dissertation shows that Crispr-Cas9 mediated deletion of the *Chkb* 5' region, where a variety of chromatin modifications and transcriptional initiation is localized, has a drastic perturbance on *Cpt1b* expression indicating that the co-expression and regulation of both genes is facilitated by the one *Chkb* 5' CpG island promoter. The condensation of the intergenic region between *CHKB* and *CPT1B* through vertebrate evolution coincided with elevated basal metabolic rate and body temperature as well as the acquisition of BAT in placental mammals.

Regulatory Linkage and Coordinated Expression of the *CHKB* and *CPT1B* genes: Implications for an Elevated Metabolic Capacity and the Origin of Brown Adipose Tissue

A Dissertation

Presented to the Faculty of the Department of Biochemistry and Molecular Biology

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Biochemistry and Molecular Biology

By

Bhavin Patel

JULY, 2023

Regulatory Linkage and Coordinated Expression of the *CHKB* and *CPT1B* genes:

Implications for an Elevated Metabolic Capacity and the Origin of Brown Adipose Tissue

By

Bhavin V. Patel

APPROVED BY:

DIRECTOR OF DISSERTATION: _____

Brian M. Shewchuk, PhD.

COMMITTEE MEMBER: _____

Kyle D. Mansfield, PhD.

COMMITTEE MEMBER: _____

Brett D. Keiper PhD.

COMMITTEE MEMBER: _____

Espen E. Spangenburg, PhD.

CHAIR OF THE DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY:

_____

John Cavanagh, PhD.

DEAN OF THE GRADUATE SCHOOL:

_____

Kathleen Cox, PhD.

ACKNOWLEDGEMENTS

I would like to start off by thanking my parents Harsha and Vinod, my sister Jayna, and my girlfriend Rudra for their unconditional unrelenting love and support throughout this journey. My parents have always allowed my sister and me to pursue our dreams no matter the cost, even if they had to work twice as hard. They tried their hardest to never make us feel like we couldn't do something, and I dedicate this PhD to them. My girlfriend Rudra has always been my support system and she made it ever so apparent by her willingness to always be in my corner as I traversed this long strenuous six-year trek. Every milestone I hit; she has been there to celebrate my achievements even when I didn't want to. Maintaining a long-distance relationship has been difficult, but her constant efforts in driving to Greenville from Durham even after consecutive grueling 12-hour shifts dealing with the emotionally taxing job of taking care of kids with cancer have not gone unnoticed and I am forever grateful to and in awe of her. Rudra, you are the greatest nurse I know and have been an angel for all those children that needed someone like you during their fight. I am so excited to witness your new journey as a Nurse Practitioner and I have no doubt you will be great at that too. You are my rock and the love of my life I cannot wait to start our life together.

I also want to thank my mentor and doctoral advisor, Dr. Brian Shewchuk, for welcoming a bright-eyed, naïve student into your lab in November of 2017 and trusting me with this project. Under your tutelage, I have seen myself grow not only as a scientist, but as a man. Your approachability and willingness to let me come and bother you with questions, suggestions, and ask for general life advice has meant more to me than you know. Thank you for your patience, I know I haven't been perfect and there have been times when I've needed an extra push, but you've been so understanding and willing to guide me to where I needed to be at the end of all of this. You've been a guiding hand every step of the way, but you've also given me the

autonomy to figure things out on my own and develop as an independent thinker. Thank you for priming me for a successful scientific career by being a fantastic mentor. I would also like to give a special thank you to Dr. Fanrong Yao for teaching me the experiments when I first came into the lab and constantly being there to help with anything I needed these past six years. You have always been there to support all my ideas in the lab, no matter how crazy they seemed. Lastly, I would like to thank my dissertation committee and every professor and collaborator that has been willing to provide their expertise throughout my training. Every meeting or conversation has helped shape the success of my project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION AND BACKGROUND

Obesity

Obesity is a rapidly growing global public health concern, affecting millions of people worldwide. The prevalence of obesity has been increasing steadily over the past few decades and is now considered a global epidemic. According to the World Health Organization, in 2016, more than 2 billion adults, 18 years and older, were classified as being overweight, or having a BMI of 25 kg/m^2 or higher. Of these, over 650 million individuals or 13% of the world's adult population were classified as obese, which is defined as having a body mass index (BMI) of 30 kg/m^2 or higher [3]. Sadly, it is projected that by 2030 around 50% of the world's population will be classified as overweight or obese. As prevalence continues to grow at this alarming rate, obesity is no longer merely a concern but rather becoming an epidemic. Thus, research to understand, treat, and even reverse the symptoms of obesity and its associated disorders is ongoing.

Obesity is associated with several comorbidities or coexisting conditions including type 2 diabetes (T2D), cancers, hypertension, non-alcoholic fatty liver disease, and inflammatory bowel disease. Type 2 diabetes, for example, is a rapidly growing concern that is closely associated with obesity from excess caloric intake and subsequent expanding lipid stores that result in elevated plasma free fatty acids, insulin resistance, and elevated fasting blood glucose levels. According to the International Diabetes Federation (IDF), over 500 million people are diagnosed with T2D worldwide as of 2021 [4]. This number is projected to rise to over 700 million people by 2045 globally [4]. Obesity also leads to an increased risk for cardiovascular disease through hypertension, high cholesterol, and high triglycerides. Furthermore, obesity is associated with certain types of cancer, including breast cancer, colon cancer, and liver cancer. Obesity is a complex and complicated disease as there is no single factor that contributes to its onset;

rather, there are many contributing factors which make it difficult to treat. The rise in obesity is largely attributed to changes in lifestyle and diet. The modern diet is often high in processed foods, sugar, and saturated fats, which compounded with reduced energy expenditure due to a rise in sedentary lifestyles results in an excess caloric intake and storage of lipids in the form of triglycerides. The increase in stored lipids can lead to expansion of adipose tissue and contribute to an increase of visceral fat, which has negative effects on vital abdominal organs, and even cellular lipotoxicity. A prolonged obesogenic phenotype has also been shown to result in metabolic inflexibility through epigenetic reprogramming which results in reduced expression of various metabolic genes.

Taking all of this into consideration, it is vital to discover the important obesity-associated genes and the changes in their epigenetic patterns during the onset and progression of obesity and the metabolic syndrome. My dissertation research involves characterizing a pair of transcriptionally linked loci that each play an important role in mitochondrial biogenesis and fatty acid oxidation. Additionally, this project will also propose a possible link between the arrangement of the *chkb-cpt1b* gene locus and the emergence of a specialized thermogenic tissue in Eutherian mammals: Brown Adipose Tissue (BAT). In the context of obesity, BAT has a recognized therapeutic potential and is an intense focus of research due to its high oxidative capacity and ability to burn fat quite literally.

CPT1B is a major rate-limiting enzyme for mitochondrial FAO

Carnitine palmitoyltransferase 1 (CPT1) is a crucial enzyme that catalyzes the rate-limiting step of mitochondrial fatty acid oxidation (β-oxidation). In total, there are three isoforms of the CPT1 enzyme that are each encoded by their respective genes. CPT1A, although ubiquitously expressed, is the predominant isoform found in the liver and kidneys. CPT1B is also widely expressed, however is more tissue-specific and serves as the predominant isoform

in striated muscle such as skeletal muscle and heart tissue, as well as in testes and brown adipose tissue (BAT). The final isoform, CPT1C, is the major isoform found in the brain. The CPT1 enzyme is required to transport long chain fatty acids (LCFAs) from the cytoplasm of the cell into the mitochondrial matrix where it can go through subsequent rounds of β-oxidation. Each cycle of β-oxidation produces an NADH and $FADH_2$ molecule as well as the 2-carbon acetyl-CoA molecule, which gets further oxidized in the citric acid cycle producing even more electron carrier molecules and 1 energy molecule of GTP. The increased quantity of electron carriers results in an increased flux of electrons through the electron transport chain of the mitochondria culminating in the production of ATP energy molecules via oxidative phosphorylation.

CPT1 is a transmembrane protein located in the outer mitochondrial membrane and is responsible for the removal of the coenzyme-A group from activated long chain fatty acyl-CoAs and the replacement with a free carnitine molecule, forming fatty acyl-carnitines (Figure 1). This switch is integral for β-oxidation to proceed because, unlike small-chain and medium-chain fatty acids, LCFAs are unable to traverse the double membrane of the mitochondria without the carnitine/acylcarnitine shuttle system. LCFAs are defined as straight chain fatty acids comprised of 13 to 21 carbon atoms and of note, the bulk of fatty acids comprising animal tissues and animal diets have 16 and 18 carbon chain lengths. Palmitic acid is a 16-carbon saturated fatty acid that makes up a large percentage of our saturated fatty acid intake. Therefore, our lab utilizes palmitic acid for our lipid incubation studies in cultured cells. The fatty acylcarnitines are then transported into the matrix of the mitochondria by a translocase protein in exchange for free carnitine. Once in the mitochondrion, the fatty acylcarnitine is subjected to the reverse reaction mediated by carnitine palmitoyl transferase 2 (CPT2) fixed in the inner mitochondrial membrane which removes the carnitine group and adds back the Coenzyme-A [5, 6]

CPT1C, the brain isoform, shows no detectable palmitoyltransferase activity which could be explained by the fact that neuronal cells largely rely on glucose oxidation to meet their

bioenergetic demands. Nevertheless, the expression of a tissue specific CPT1 enzyme in the brain that has high sequence similarity to the canonical CPT1 enzymes yet shows little to no definitive acyltransferase activity or enhanced oxidation of fatty acids is confounding. The enzymatic role of CPT1C and its function in neurochemistry remains elusive. Thus far, only defects in CPT1A have been clinically characterized [6-9]. Patients with a complete loss of *CPT1B* have not been identified indicating the importance of the gene for survival.

While a homozygous knockout of *Cpt1b* in mice proves to be embryonically lethal, a heterozygous knockout of *Cpt1b*, reducing CPT1B protein levels by 50%, results in mice being predisposed to insulin resistance, lipotoxicity, and fatal hypothermia [10-12]. A common single nucleotide polymorphism (SNP) was found in the *CPT1B* gene in an Italian cohort population who were classified as severely obese (BMI >30 kg/m^2) and undergoing bariatric surgery [13]. This suggest that a defect in the *CPT1B* gene can lead to an obese phenotype as well as liver steatosis [13].

Considering the important role CPT1B plays in lipid metabolism in skeletal muscle, it's no surprise that it serves as a hot topic for research and a promising target for potential therapeutics to combat obesity. In fact, *CPT1B* seems to be differentially expressed between healthy lean individuals and obese individuals with primary myotubes cultured from skeletal muscle biopsies of obese women showing a blunted expression of the gene in response to palmitate treatment compared to their lean counterparts [14, 15]. Perhaps more strikingly, the blunted expression of *CPT1B* in response to lipid treatment persisted in primary myocytes that were cultured for an extended period of time under identical conditions. The researchers discovered that the cells from obese individuals had been epigenetically re-programmed. In other words, changes in CpG methylation, histone acetylation, and transcription factor binding were observed between the two experimental groups suggesting that the *CPT1B* gene is at least partly regulated by epigenetic modifications in skeletal muscle [14, 15].

Further investigation of the promoter region of the *CPT1B* gene revealed the presence of the 3' end of the choline kinase beta (*CHKB*) gene less than 300 bp upstream in humans and less than 500 bp upstream in mice (http://genome.ucsc.edu). Both genes seemingly exist on an island in chromosome 15 of the mouse genome, with the closest genes being 6.1 kb downstream of the 3' end of *Cpt1b* and 24 kb upstream of *Chkb*. The juxtaposition of the genes brings up the obvious question of whether there is transcriptional or regulatory interplay between the genes.

CHKB catalyzes the first committed step of phosphatidylcholine biosynthesis.

The choline kinase enzyme is a cytoplasmic protein that catalyzes the phosphorylation of free choline to phosphocholine, marking the first committed step in the CDP-choline pathway for phosphatidylcholine biosynthesis. Phosphatidylcholine (PC) is one of the most abundant and integral phospholipids present in biological membranes, including both inner and outer mitochondrial membranes (Figure 1). PC comprises approximately 50% of total phospholipid mass in mammalian cells. There are two separate genes, *CHKA* and *CHKB*, that have evolved from duplication and divergence of a common ancestral gene and encode three isoforms of choline kinase: CHKA1, CHKA2 and CHKB [16, 17]. CHKA 1 and 2 are products of alternative splicing of *CHKA* transcripts while CHKB is a product of the *CHKB* gene. The length of the *Chka* gene is 40 kb whereas *Chkb* is only 3.5 kb in length [17].

Interestingly, both genes contain 11 exons each and much of the difference in genomic length is contributed by significantly longer introns in *CHKA*. Choline kinase proteins function as either homo- or heterodimers [16, 17]. *CHKA* is ubiquitously expressed and accounts for a large percentage of enzymatic activity, with highest activity levels being observed in liver and testis [16, 17]. *CHKB* is also widely expressed in most tissues, with the highest activity levels being detected in heart, skeletal muscle, and brown adipose tissue [16-18]

Wittenberg and Kornberg published the first characterization of choline kinase in brewer's yeast in 1953 [19] followed by several decades of minimal progress until 1984, when Ishidate *et al.* were able to purify the choline kinase protein to homogeneity from rat kidneys [20]. Following this, several groups worked to clone and characterize different genes and cDNAs of choline kinase from rodents, mammals, yeast, and plants [21-27]. In 2000, Aoyama *et al.* used rapid amplification of cDNA ends (RACE) of mouse and rat *CHKA* and *CHKB* transcripts to compare the sequences to the newly available human genome sequences reported from the human genome project [17]. Promoter reporter assays identified a proximal promoter region of the *Chkb* gene in murine Hepa-1 cells located within 110 bp upstream of the transcription start site. This region contained two CCAAT boxes with inverse orientation, a CREB binding motif, as well as an SP1 binding site. There also was the presence of a TATA-like binding element however a functional TATA box was not present [16, 17]. All these findings suggest *CHKB* behaves like a housekeeping gene.

CHK activity has also been implicated in fibroblast proliferation; studies using NIH3T3 cells have shown that stimulation of choline kinase expression with certain growth factors and a subsequent increase in phosphocholine was necessary for cell proliferation. Higher levels of the choline kinase protein and phosphocholine, the direct product of choline kinase, are also detected in cancer cell lines and therefore may contribute to carcinogenesis. Chka, unlike Chkb, has been demonstrated to be an indispensable enzyme for early embryonic development in mice [28]. Thus, only the complete knockout of *Chka* in mice proves to be embryonically lethal [29], while a knockout of the *Chkb* gene in mice or a defect in the human *CHKB* gene leads to a peculiar form of congenital muscular dystrophy [28, 30-33].

Congenital muscular dystrophies (CMD) are a group of muscular dystrophies that present early at birth and progress aggressively. These disorders are commonly degenerative and often affect voluntary muscles. Generally, most CMD cases reported thus far are a result of an autosomal recessive mutation in extracellular matrix protein encoding genes. To date, there

are approximately 30 different types of CMD discovered, each based on the respective gene or protein malfunction (Muscular Dystrophy Association).

Interestingly, in 1998, Nishino *et al.* reported a peculiar new form of CMD in four unrelated Japanese patients in their adolescence. All four patients exhibited extreme cognitive impairment along with signs of neonatal hypotonia, delayed motor development, increased muscle wasting, and dilated cardiomyopathy [34]. Perhaps more interesting, these patients displayed abnormally enlarged mitochondria, or megamitochondria, located at the periphery of skeletal muscle fibers with the center being devoid of mitochondria [34]. This genetic disorder displayed recessive autosomal inheritance, as none of the patients' parents or siblings were afflicted with CMD. The most intriguing and prominent phenotype of CMD was the presence of permanently localized peripheral megamitochondria, therefore mitochondrial DNA was analyzed for possible mutations. However, no duplication or deletions were found in the mitochondrial DNA and instead the most likely cause for the dystrophic phenotype was a mutation in a gene found in nuclear DNA. However, the association of this phenotype with the *CHKB* gene wouldn't be made until after the development of a transgenic mouse line presenting a similar muscular dystrophy phenotype.

In 2006 Dr. Gregory Cox at the Jackson Laboratory identified a spontaneous recessive mutation in mice that coincided with the onset of a CMD that had a rostral to caudal gradient of severity in muscular atrophy as well as neonatal forelimb bone deformity. The hind limb muscles of these mice were atrophied more severely than the forelimb muscles, and as a result, these mice were aptly named as *RMD* to signify this rostrocaudal muscular dystrophy phenotype [31]. It was discovered through positional cloning that these mice carry a spontaneous 1600 bp intragenic deletion of *Chkb* on chromosome 15 that encompassed exon 3 through intron 9 resulting in truncated mRNA transcripts. Consequently, CHKB protein and enzyme activity levels are also undetectable in all skeletal muscle tissue samples from RMD mice compared to wild type littermates [31]. While forelimb muscles remained largely unaffected, RMD mutant

mice showed severe progression of atrophy in hindlimb muscles culminating in the loss of motor activity by two to three months of age. Histological analysis of quadriceps muscles showed the presence of enlarged megamitochondria at the periphery of the skeletal muscle fibers. PC levels were reduced in forelimb and more so in the hindlimb skeletal muscle tissue, although when we compare absolute levels against WT there is only a slight reduction in skeletal muscle tissues of RMD mice [31]. In fact, only a 31% reduction in PC levels was reported in *Chkb*$^{-/-}$ RMD mouse hindlimb muscles compared to wildtype hindlimb muscles while CHKB enzyme levels were undetectable. It is widely accepted that although choline kinase is necessary for PC biosynthesis, CK activity is generally present in excess and under these circumstances the level of CK activity does not influence the rate of PC synthesis [35]. Although CHKB is ubiquitously expressed, loss of enzyme activity results in a PC reduction in only skeletal muscle tissue showing a tissue specific defect. Interestingly, *Chka* expression is downregulated as mice age into adulthood compared to young mice, so it seems that CHKB is the dominant choline kinase isoform in adult skeletal muscle tissue. However, it has also been reported that upregulating *Chka* expression in affected RMD skeletal muscle can reverse the disease phenotype by compensating for PC synthesis [33].

Following the characterization of a *Chkb* gene knockout in the newly developed RMD mice, patients with CMD and megamitochondria were sequenced and found to contain mutations in the *CHKB* gene that lead to loss of function. The disease was reclassified as congenital muscular dystrophy with megaconical myopathy (MDCMC) and is currently the only muscular dystrophy caused by a defect in phospholipid biosynthesis [30, 36-44].

Conjoined genes: *Chkb*, *Cpt1b*, and the readthrough transcript

A detectable readthrough transcript containing exons from both genes is present that initiates at the *Chkb* 5’ region and ends at the 3’ end of *Cpt1b*. This RNA product is a candidate

for nonsense mediated decay and is unlikely to translate into a functional protein. Instead, this long noncoding RNA (lncRNA) product may play a regulatory role in maintaining an open chromatin domain that includes both genes. Interestingly, a few studies that have investigated *Chkb* or *Cpt1b* separately have mentioned the existence of this readthrough transcript and occasional co-transcription of these two genes. However, no one has studied the possible regulatory or transcriptional linkage of the *Chkb* and *Cpt1b* genes in further detail, and certainly not in the context of the evolution of BAT-based homeothermy in eutherians. Our theory is that the *Chkb* and *Cpt1b* are members of a class of genes known as “conjoined genes”. Conjoined genes in eukaryotes are predicted at several hundred loci and are generally present when two or more genes are in relative proximity to one another and are transcribed in the same direction [45]. The classical definition of a conjoined gene is a gene that produces transcript(s) by combining a part of one exon from each of the two distinct parent genes that exist on the same chromosome and are in the same orientation and translate independently into separate proteins [45-47]. Based on these stipulations, the *chkb-cpt1b* gene locus fits under the conjoined gene category and our findings will provide further evidence for this phenomenon.

Association of the *CHKB-CPT1B* locus with narcolepsy

A variant between the *CHKB* and *CPT1B* genes containing a single nucleotide polymorphism (SNP) is associated with a narcolepsy phenotype in East Asian populations [48, 49]. The SNP was located approximately 250 bp downstream of the *CHKB* stop codon and approximately 100 bp upstream of the *CPT1B* gene placing it within the intergenic region between both genes [48, 49]. Interestingly, QRTPCR results in patients with the risk allele showed lower expression of both *CHKB* and *CPT1B* genes [48, 49]. Several groups have implicated FAO and the carnitine shuttle system in sleep regulation thus making CPT1B expression a likely culprit for the narcolepsy phenotype. CHKB has an important role in CDP-

choline based PC synthesis and perturbations in this system may also lead to narcolepsy, albeit in a more complicated pathway. The important discovery is that a SNP within the intergenic region of the *CHKB* and *CPT1B* genes in these patients leads to reduced expression of both genes suggesting a regulatory and transcriptional connection.

Coordinated expression of *Chkb* and *Cpt1b* in mouse tissues

It has been previously reported that the expression of *Cpt1b* can be regulated by serum levels of long chain fatty acids that are taken up by the cells via PPARa activation and subsequent binding to the peroxisome proliferator response elements (PPRE) located within the *CPT1B* promoter region in humans. PPARs (Peroxisome Proliferator-Activated Receptors) and PGC1a (Peroxisome Proliferator-Activated Receptor Gamma Coactivator 1-alpha) are two important proteins that act as master regulators of mitochondrial function and biogenesis in the cell. PPARs are a family of nuclear receptors that are activated by binding of overlapping derivatives of fatty acids. Once activated, these nuclear receptors can bind to specific response elements within the proximal promoters of multiple genes involved in lipid metabolism, inflammation, and energy homeostasis to regulate their expression. There are three isoforms of PPARs: PPAR alpha, PPAR beta/delta, and PPAR gamma. Each isoform has distinct functions and target genes, but all of them play an important role in the regulation of mitochondrial function and biogenesis. PGC1a is a coactivator protein that binds to various transcription factors, including PPARs, to increase the transcriptional activity of target genes. PGC1a is highly expressed in tissues with high energy demands, such as skeletal muscle and the brain, and plays a critical role in the regulation of mitochondrial function and biogenesis. PGC1a activates genes involved in the production of new mitochondria, the import of proteins into the mitochondria, and the maintenance of the mitochondrial DNA. Both PPARs and PGC1a have been implicated in a wide range of physiological and pathological conditions.

We first wanted to determine if expression of *Cpt1b* and *Chkb* in cultured myotubes responded in parallel to long chain fatty acid treatment. Using the transformed C2C12 murine skeletal muscle cell line, we found that both genes respond in parallel to palmitate treatment. Expression of not only *Cpt1b,* but also *Chkb* increases in response to LCFA treatment, revealing the coordinated nature of these two genes. Furthermore, we also found that transcription of the two genes is coordinated in vivo with the highest expression of both genes being found in brown adipose tissue compared to moderate expression in heart and skeletal muscle in mouse tissue samples. This marked a serendipitous revelation that these two genes are coordinately expressed at high levels in metabolically active tissues, but the level of expression in heart and skeletal muscle is dwarfed by the levels observed in brown adipose tissue. We therefore began shifting our focus to study the origins and function of BAT in the context of vertebrate evolution from the lens of the *Chkb-Cpt1b* gene locus. Additionally, BAT serves as a popular target of research and therapeutic intervention for combatting obesity and its related metabolic disorders due to its inherent capability for thermogenic energy expenditure via the short-circuiting of the mitochondrial oxidative phosphorylation to produce substantial amounts of heat from burning glucose and free fatty acids as metabolic fuel.

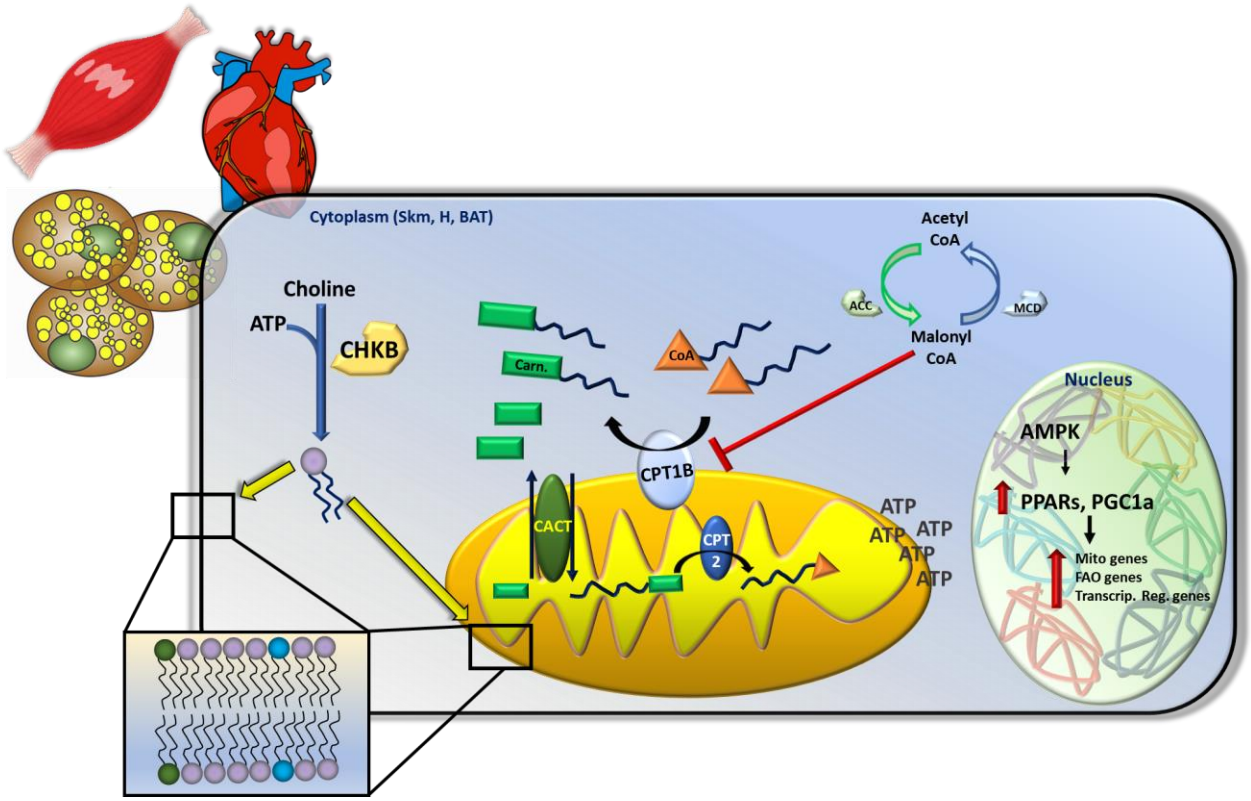

Figure 1. CHKB and CPT1B enzyme locations and functions. CPT1B is the predominant isoform in skeletal muscle, cardiac, and brown adipose cells. It is an important part of the carnitine shuttle system that is crucial for delivering LCFA from the cytoplasm to the mitochondrial matrix. CPT1B temporarily switches the CoA group with carnitine to form fatty acyl carnitines which can then be transported into the matrix via CACT in exchange for free carnitine. Once in the matrix, CPT2 removes the carnitine and adds back the CoA group completing the transport. Malonyl CoA, an intermediate of fatty acid biosynthesis, is an allosteric inhibitor of CPT1B. CHKB catalyzes the first committed step of the CDP-choline/ Kennedy pathway for PC biosynthesis. PC is one of the most abundant phospholipids present in the cell membrane and both inner and outer mitochondrial membranes.

Brown adipose tissue

Brown adipose tissue (BAT) is a uniquely mammalian tissue that helps regulate neonatal body temperature via fatty acid oxidation driven heat production. It is reasonable to argue that BAT acquisition was an important evolutionary event that allowed for mammalian survival and diversification. We therefore wanted to investigate whether the elevated co-expression of both *CHKB* and *CPT1B* genes in this tissue was a product of transcriptional and regulatory linkage between these two genes and whether that linkage was a key driving force in the emergence and formation of this tissue in mammals.

Classical BAT depots are ubiquitous in newborn humans and are vital for maintaining stable body temperature via non-shivering thermogenesis as newborns do not yet have the fully developed skeletal muscle framework to facilitate shivering thermogenesis [50]. In addition, human infants and smaller mammals have a higher surface area to volume ratio, resulting in increased heat dissipation, and thus making BAT a necessary tissue to combat hypothermia. For a long time after its initial discovery, it was believed that the BAT repository and function diminished to negligible levels in adults due to the increasing disparity of BAT mass relative to body mass in a fully grown adult rendering it insufficient to have any significant effects on energy expenditure. However, in 2003, studies using 18-fluorodeoxyglucose-positron emission tomography coupled to computed tomography (^{18}FDG PET/CT) found uptake in response to cold stimulation in the regions of the body where BAT depots were expected to exist in adults [51, 52]. X-ray computed tomography of these BAT regions, for example the supraclavicular area, indicated a similar signature to adipose tissue. Subsequent studies showed that BAT is largely present and more activated by cold exposure in healthy lean and younger individuals but is reduced in overweight and older individuals as well as individuals with metabolic diseases [53, 54].

Mice also possess substantial depots of BAT, with their more prominent repositories located in the interscapular and subscapular regions. There are also numerous smaller BAT

depots scattered deep in the dorsal neck region between the head and scapula, surrounding the aorta, and surrounding the kidneys. In other words, BAT forms a sort of thermal jacket around the vital organs and major blood vessels in mice to protect from hypothermia. The interscapular region, however, is the largest and most widely studied depot of classical BAT in mice and therefore it is the depot that we use for our studies. In one study, BAT from healthy mice, that was implanted in mice fed with a prolonged high fat diet, improved glucose homeostasis and insulin sensitivity [55]. In fact, transplantation of BAT from healthy mice into a diabetic mouse model showed the insulin-independent reversal of Type 1 diabetes (T1D) as well as prevention of T1D if BAT transplantation occurred before induction of T1D [56, 57].

Brown adipocytes and skeletal muscle myoblasts share a common lineage and are postulated to arise from myf5^{+} progenitors. The Spiegelman lab was one of the first to use a myf5^{cre} construct crossed with a cytoplasmic reporter to show that interscapular brown adipocytes and skeletal muscle cells originated from myf5^{+} progenitors while white adipocytes did not seem to share the same lineage origin [58]. However, over time many other groups have demonstrated a clear heterogeneity within BAT depots that showed not all brown adipocytes were derived from myf5^{+} progenitors [59] [60]. In fact, there are even indications for certain groups of white adipocytes that showed my5^{+} origins [61]. There also exists a distinct family of inducible, metabolically active, thermogenic, UCP-1 expressing adipocytes termed “beige” or “brite” adipocytes that exist within WAT depots [62]. There is still much debate on whether beige cells have a similar thermogenic potential and oxidative capacity to classical brown adipocytes, as these cells do not share the same myf5^{+} embryonic lineage. While the interscapular BAT depot represents the classical brown fat made up of myf5^{+} originating brown adipocytes, brown adipocytes from different anatomical locations may have different embryonic origins [59] [60].

Seale *et al.* initially reported that the transcriptional regulator, PRDM16, controls a bidirectional switch of cell type between skeletal myoblasts and brown adipocytes using *in vitro* studies [58]. However, in subsequent studies published by Harms *et al.*, the group was unable

to recapitulate the same findings *in vivo* [63, 64]. PRDM16 showed strong binding to enhancer/promoter regions of BAT-selective genes as well as binding to promoter regions of WAT-selective genes in BAT depots. Furthermore, it was found that PRDM16 promoted the suppression of WAT-selective genes and enhancement of BAT-selective genes via binding and recruitment of mediator complexes that influenced significant changes in chromatin architecture [63, 64]. In fact, knocking out PRDM16 in adult mouse BAT showed a decrease in expression of BAT-selective genes and a concomitant increase in expression of WAT-selective genes resulting in a diminished BAT characteristics of the depots in these mice compared to WT controls [63, 64]. Therefore, although lineage tracing studies using myf5 have indicated a skeletal muscle cell origination for most brown adipocytes, the PRDM16 findings show there may be some relation of these cells to white adipocytes as well. We believe that the evolution of the spatial arrangement of the *CHKB* and *CPT1B* gene loci, and more specifically their coordinated regulation and expression, were a major contributing factor for the development of brown adipocytes and BAT.

Unlike white adipocytes, which possess a large unilocular lipid droplet, brown adipocytes contain numerous small multilocular lipid droplets that store fatty acids in the form of triacylglycerols. The importance of this distinction becomes apparent when the system enters a state, usually induced by cold exposure, where it is necessary to oxidize those fatty acids for heat, named adaptive thermogenesis. Brown adipocytes can rapidly lipolyze and endogenously oxidize the fatty acids while white adipocytes mainly serve as energy reservoirs during fed states and release the hydrolyzed free fatty acids during fasting states into the blood stream to be taken up other cells for ATP production. BAT is highly innervated by the sympathetic nervous system which activates the thermogenic program upon cold stimulation. Briefly, cold exposure leads to release of norepinephrine from sympathetic nerve endings and subsequent binding to and stimulation of the β3-adrenergic receptors present on the surfaces of brown adipocytes commencing a signaling cascade that eventually leads to heat production. Interestingly, studies

have also found diet-induced BAT activity in lean healthy individuals as a response that increases energy expenditure during caloric overload [65-67]. Another distinct characteristic of BAT that separates it from WAT is that they are heavily saturated with numerous UCP1-containing mitochondria that can rapidly oxidize LCFAs to generate a substantial amount of heat. Uncoupling protein 1 (UCP1) is a transmembrane enzyme present in the inner mitochondrial membranes of brown fat cells and its purpose is to uncouple mitochondrial respiration from ATP synthase activity. It does so by essentially acting as a “leak” for protons to travel back into the matrix from the intermembrane space, thereby relinquishing the membrane potential built up by proton pumps (complexes I, III, and IV) of the electron transport chain. In non-brown adipocyte cells, this electrochemical gradient would normally result in hydrogen ions traveling down their gradient into the mitochondrial matrix through the ATP synthase enzyme which uses that proton motive force to phosphorylate ADP into ATP. However, in BAT the energy from the protein motive force dissipated through the UCP1 channel is converted into heat. Of course, this process is energetically costly and highly inefficient as the electron carriers generated from the high rate of FAO produce little to no ATP. This supports the idea that the emergence of BAT was a necessary evolutionary trait that allowed for eutherian mammalian radiation through exploitation of new ecological niches which severely outweighed the energetic costs associated with BAT-mediated NST.

Functional connection of *Chkb* and *Cpt1b* in brown adipocytes

As mentioned earlier, our gene expression analysis showed the highest transcript levels of both *Chkb* and *Cpt1b* genes in BAT tissue compared to all other tissues tested in mice. Our hypothesis is that the juxtaposition of these genes, on an evolutionary timescale, provided the catalyst necessary for the development of this specialized thermogenic adipose tissue. As mentioned, CHKB is responsible for the synthesis of PC, a crucial phospholipid that makes up a

large percentage of cell membrane phospholipids and is specifically required for mitochondrial membrane biogenesis and proper structure and functionality. Brown adipocytes, along with muscle and cardiac cells, contain a high number of mitochondria to fuel their activity via FAO thereby requiring high expression of both *Chkb* and *Cpt1b*. Furthermore, it has been demonstrated that cold exposure also leads to increased brown adipocyte mitophagy due to UCP-1 induced stress and damage coupled with biogenesis to maintain mitochondrial homeostasis with a healthy population [68]. Additionally, Krahmer *et al.* were able to show the importance of PC as a necessary surfactant to prevent coalescence of growing lipid droplets, which would otherwise form large, lipolysis-resistant lipid droplets that store large amounts of triglycerides as in white adipocytes [69]. Brown adipocytes, as described above, require small multilocular lipid droplets with ideal surface area to volume ratios that can be readily accessible for lipolysis and subsequent FAO. Any perturbance in PC levels seems to result in coalescing of lipid droplets into larger droplets which may hinder the rapid lipolysis of TAG and release of fatty acids necessary to fuel nonshivering thermogenesis in brown adipocytes and consequently minimize the thermogenic impact of the tissue. As mentioned earlier, homozygous knockout of *Cpt1b* in mice leads to embryonic lethality; however, a heterozygous knockout, reducing CPT1B protein levels by approximately 50% in striated muscles and BAT, results in the mice succumbing to fatal hypothermia upon 4^{o}C cold challenge. This effect was further amplified with prolonged exposure to suboptimal temperatures, leading to more than 50% of the mutant mice dying compared to WT littermates [12]. This increase in fatality can be associated with reduced FAO in skeletal muscle and BAT and a subsequent hindrance in shivering and nonshivering thermogenesis, respectively.

Chromatin remodeling and transcriptional activation

Our hypothesis for the regulation of a conjoined *Chkb-Cpt1b* gene locus involves the localized targeting of chromatin modifications to the 5' flanking promoter of *Chkb* leading to the transcriptional initiation at this same site that results in two separate *Chkb* and *Cpt1b* transcripts. It is therefore important to understand the chemistry involved in chromatin structure and remodeling. Chromatin is made up of basic repeating units of nucleosomes which comprise of DNA wrapped around histone octamers. This arrangement allows for condensed packaging of DNA in the cell nucleus. The histone core of each nucleosome is made up of two H2A-H2B dimers and a H3-H4 tetramer wrapped by 147 base pairs of DNA and linked to the adjacent nucleosome by approximately 40-60 bp of DNA [70, 71]. It is no surprise that this type of packaging of DNA results in difficulty for transcription to occur since the DNA is not readily accessible to transcription factors and machinery. Therefore, condensed chromatin, otherwise known as heterochromatin, must undergo remodeling for transcriptional processing.

The N-terminal tails of histones contain intrinsically disordered regions that protrude from the nucleosomal unit and can undergo varied and extensive posttranslational modifications to either negatively or positively influence transcription of the genes located in the domain. Some examples of modifications to amino acids within the histone tails are acetylation, methylation, ubiquitination, phosphorylation, and ADP-ribosylation [72, 73]. Transcriptional activation is associated with modifications that result in a more open euchromatin structure where the DNA is loosely bound to histones. These modifications include histone 3 and histone 4 lysine acetylation, H3 lysine 4 di- or trimethylation, and H3 lysine 36 and lysine 79 trimethylation. Epigenetic chromatin modifications that signal gene inactivation include H3 lysine 9 or lysine 27 methylation as well as methylation of H4 at lysine 20 [72]. Acetylation and methylation are the most common forms of modifications, and these are facilitated by histone acetyltransferases (HATs) and histone methyltransferases (HMTs), respectively.

As mentioned previously, a heterochromatin structure is indicative of gene silencing while a euchromatin structure represents transcriptional activation mainly due to how open or tightly packaged the DNA is around the histone core. Lysine acetylation by HAT enzymes neutralizes the positive charges of the histones and result in weakened interactions between the histones and the negatively charged DNA.

CpG islands

DNA methylation is a separate covalent modification that primarily targets cytosines that are followed by guanine, although additional genomic motifs have more recently been identified as methylated. DNA methyltransferases (DNMTs) add a methyl group to the C5 position of cytosine to form 5-methylcytosine. The addition of the methyl group to the cytosine at CpG sites results in gene silencing by either sterically hindering the binding of transcription factors to DNA or by recruiting methyl-cytosine binding proteins which then recruit corepressor complexes including histone deacetylases (HDACs) [74]. There are three types of DNMTs; DNMT1, DNMT3a, and DNMT3b [75, 76] however, DNMT1 is the best characterized and highly expressed mammalian somatic maintenance methyltransferase that prefers to methylate hemi-methylated DNA [74, 77, 78]. During DNA replication, DNMT1 localizes to the replication fork and works to make sure the newly synthesized strand mimics the existing methylation pattern [74, 79].

CpG islands are regions of the genome that are 500 nucleotides or more in length that contain a G+C content equal to or higher than 50% and possess a CpG/GpC ratio higher than 0.6 [80]. Based on the amount of genomic cytosines and guanines, the expected the frequency of a CG dinucleotide in the genome to be around 6.2% of dinucleotides. However, they are only observed at a frequency of around 1%. The reason for this reduction is due to spontaneous deamination of methylcytosine. Spontaneous deamination of cytosine converts it to uracil, which

can be recognized and removed by a special enzyme known as uracil-DNA glycosylase, starting the process of DNA repair [81]. However, in vertebrate genomes with CpG methylation the spontaneous deamination of 5-methyl cytosine converts it to thymine, which goes unrepaired. Over evolutionary time, this has resulted in the aforementioned depletion of CpG dinucleotides in the genome. In contrast, CpG islands in or near gene promoter regions maintain their high relative concentration of CpG dinucleotides due to a historical lack of DNA methylation in these regions, which is thought to be maintained by the binding of transcription factors such as SP1 that have GC-rich binding sites, which competes for DNMT binding [82]. Furthermore, the chromatin structure around these CpG islands is often open and accessible, due to histone acetylation and methylation mediated by histone acetyltransferases and methyltransferases, which is also often occupied by RNA polymerases that are actively transcribing the gene. Thus, the presence of these chromatin modifiers and transcription machinery can act as insulators against DMNT occupancy. The existence of a CpG island in the 5' region of the *Chkb* gene is therefore indicative of a constitutively active housekeeping gene resulting in an easily accessible open chromatin structure that encompasses the *Chkb-Cpt1b* gene locus.

Evolution of endothermy: UCP1 alone cannot account for BAT emergence.

The evolutionary events that led to endothermy are still unclear and we hypothesize that changes in the *Chkb-Cpt1b* gene locus was a crucial driving force for its emergence. It has been demonstrated by several different groups that marsupials and monotremes do not possess classical BAT depots. However, certain marsupials are able to produce transient cold-induced multilocular adipocytes that express a functional UCP1 within WAT depots but are not thermogenic [83-88]. In fact, BAT has only unequivocally been detected in eutherian (placental) mammals, evoking the question of what the driving force was for the development of BAT. Was it simply the acquisition of UCP1, which has been lauded as the main marker for brown

adipocytes? First discovered in the late 1970s, UCP1 was initially believed to be a monophyletic trait that originated with the evolution of brown adipose tissue only in eutherian mammals [89-91]. However, in 2005 using comparative genetics, Jastroch *et al.* discovered the presence of a *ucp1* orthologue in certain teleost fish species that had high sequence similarity to eutherian *ucp1* [84]. What this meant was the *ucp1* gene was likely present farther back than 420 mya in common ancestor of lobe finned and ray finned fish before their divergence. Furthermore, the amino acid sequence comparison suggests that this “ectothermic *ucp1*” should still have the same proton translocation characteristic that is seen in eutherians. Of course, teleosts are still ectotherms and therefore are unable to utilize the *ucp1* gene/protein for nonshivering thermogenesis and we argue that in part they are lacking the elevated FAO required to fuel UCP1-mediated heat production. Additionally, while marsupials and monotremes are endothermic they present lower body temperatures that are subject to quite severe fluctuations dependent on the external conditions, compared to the higher and more stable nature of body temperature in placental mammals. These observations are compounded by the lack of reported classical BAT depots or UCP-1 mediated NST, even though the *ucp1* gene is present, in species of these clades [83, 86]. Therefore, it is more appropriate to classify the species of these two clades as “heterothermic endotherms” or “metatherians" due to their lower basal metabolic rate and lower and fluctuating body temperature. It is also important to mention the presence of a skeletal muscle based NST program that exists in most vertebrates and may have predated BAT-mediated thermogenesis. This system involves the uncoupling activity of the SERCA calcium pump which results in heat production from ATP hydrolysis while myofibril contractions do not take place. It has been postulated that this method of NST is prominent in marsupials and monotremes. Observational data from hibernating Echidna indicated an increase in body temperature with seeming no muscular movements [92] so, while this species does not contain BAT, perhaps this warming was due to SERCA NST. However, there are obvious drawbacks to using muscle NST which include perturbance in locomotion which does

not bode well for survival from predation. This method of NST was perhaps insufficient to protect from cold stress at birth, or to allow for high body temperature for extended parental care [91, 93-95]. Comparing rewarming rates from torpor between short-beaked echidnas and alpine marmot, a similarly sized eutherian, showed lower rate of rewarming for the echidnas compared to the marmots [96]. This indicates that BAT-based NST was an important acquisition for rapid arousal from hibernation that certainly was more effective than any other mode of NST. These observations provide more evidence against UCP1 being the sole driving force for BAT emergence and function and therefore may have required additional mechanisms to evolve. We have discovered a phylogenetic pattern of the increased condensation of the *Chkb* and *Cpt1b* intergenic region coinciding with elevated basal metabolic rate and body temperature as well as BAT-based endothermy in placental mammals. Therefore, we postulate that the juxtaposition of the *Chkb* and *Cpt1b* genes was necessary to elevate basal expression of both genes and provide the phosphatidylcholine for cellular and mitochondrial membrane biogenesis and fuel from FAO driven by CPT1B for UCP1-mediated thermogenesis in the newly acquired BAT.

CHAPTER II: EMERGENT COORDINATION OF THE *CHKB* AND *CPT1B* GENES IN EUTHERIAN MAMMALS: IMPLICATIONS FOR THE ORIGIN OF BROWN ADIPOSE TISSUE[1]

[1]This research was originally published in the Journal of Molecular Biology: Bhavin V. Patel, Fanrong Yao, Aidan Howenstine, Risa Takenaka, Jacob A. Hyatt, Karen E. Sears, and Brian M. Shewchuk Emergent Coordination of the *CHKB* and *CPT1B* Genes in Eutherian Mammals: Implications for the Origin of Brown Adipose Tissue. *Journal of Molecular Biology.* 2020; 432. 6127–6145.

Summary

Mitochondrial fatty acid oxidation (FAO) contributes to the proton motive force that drives ATP synthesis in many mammalian tissues. In eutherian (placental) mammals, brown adipose tissue (BAT) can also dissipate this proton gradient through uncoupling protein 1 (UCP1) to generate heat, but the evolutionary events underlying the emergence of BAT are unknown. An essential step in FAO is the transport of cytoplasmic long-chain acyl-coenzyme A (acyl-CoA) into the mitochondrial matrix, which requires the action of carnitine palmitoyltransferase 1B (CPT1B) in striated muscle and BAT. In eutherians, the *CPT1B* gene is closely linked to the *choline kinase beta* (*CHKB*) gene, which is transcribed from the same DNA strand and terminates just upstream of *CPT1B*. CHKB is a rate-limiting enzyme in the synthesis of phosphatidylcholine (PC), a predominant mitochondrial membrane phospholipid, suggesting that the coordinated expression of *CHKB* and *CPT1B* may cooperatively enhance mitochondrial FAO. The present findings show that transcription of the eutherian *CHKB* and *CPT1B* genes is linked within a unitary epigenetic domain targeted to the *CHKB* gene, and that that this regulatory linkage appears to have resulted from an intergenic deletion in eutherians that significantly altered the distribution of *CHKB* and *CPT1B* expression. Informed by the timing of this event relative to the emergence of BAT, the phylogeny of *CHKB-CPT1B* synteny, and the insufficiency of UCP1 to account for eutherian BAT, these data support a mechanism for the emergence of BAT based on the acquisition of a novel capacity for adipocyte FAO in a background of extant UCP1.

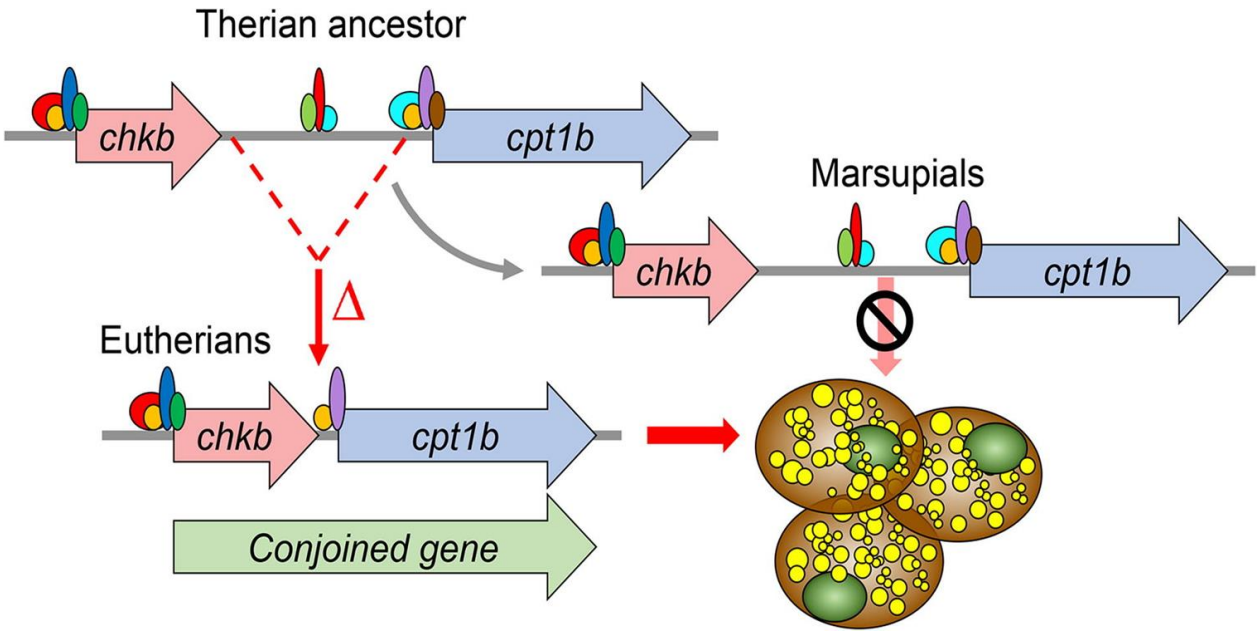

Figure 2. Graphical abstract. The transcription of the eutherian *CHKB* and *CPT1B* genes is linked within a unitary epigenetic domain targeted to the *CHKB* gene, and this regulatory linkage appears to have resulted from an intergenic deletion in eutherians that significantly altered the distribution of *CHKB* and *CPT1B* expression. Informed by the timing of this event relative to the emergence of BAT, the phylogeny of *CHKB-CPT1B* synteny, and the insufficiency of UCP1 to account for eutherian BAT, these data support a mechanism for the emergence of BAT based on the acquisition of a novel capacity for adipocyte FAO in a background of extant UCP1.

Keywords

Brown adipose tissue, CHKB, CPT1B, fatty acid oxidation, conjoined gene

Abbreviations

ATP, adenosine triphosphate

FAO, fatty acid oxidation

NADH, nicotinamide adenine dinucleotide

$FADH_2$, flavin adenine dinucleotide

CoA, coenzyme A

TCA, tricarboxylic acid

PC, phosphatidylcholine

BAT, brown adipose tissue

WAT, white adipose tissue

MRCA, most recent common ancestor

CMD, Congenital Muscular Dystrophy

Introduction

Lipid metabolism contributes significantly to homeostasis in mammals and is regulated at multiple levels in response to diet, stress, physical activity, reproduction, and the environment. The mitochondrial β-oxidation of long chain fatty acids derived from both dietary lipids and triglycerides synthesized *de novo* is the predominant pathway for the generation of ATP from lipids. Fatty acid oxidation (FAO) produces NADH and $FADH_2$ both directly and through the complete oxidation of the acetyl-coenzyme A (acetyl-CoA) product in the TCA cycle. The oxidation of NADH and $FADH_2$ in turn generates a proton motive force that fuels ATP synthesis in a broad range of tissues which spares systemic glucose, and tissues like striated muscle rely heavily on FAO to provide enough ATP for the elevated demand in these cells. In brown adipose tissue (BAT), the FAO derived proton flux also fuels the thermogenic leakage of protons back into the mitochondrial matrix through inner membrane localized uncoupling protein 1 (UCP1), which contributes significantly to thermoregulation uniquely in eutherian (placental) mammals.

A rate-limiting reaction in the pathway for FAO is the synthesis of acylcarnitine from cytoplasmic long chain fatty acyl-CoA as required for transport into the mitochondrial matrix for β-oxidation [7]. The exchange of the CoA moiety for carnitine is mediated by carnitine palmitoyltransferase 1 (CPT1), which exists as three independently encoded isoforms with differing tissue predominance: the liver isoform CPT1A expressed in many non-muscle tissues, the muscle isoform CPT1B expressed in striated muscle and BAT, and the brain isoform CPT1C. As skeletal muscle represents a major site of FAO, regulation of the *CPT1B* gene has been a focus of research on the systemic control of lipid metabolism and its derangement in metabolic disease. *CPT1B* is regulated by long chain fatty acid levels and other signals that govern FAO in skeletal muscle congruent with substrate availability and energetic needs, which is mediated at the gene level through multiple characterized gene-proximal elements [15, 97-

108]. Perhaps significantly, the regulation of *CPT1B* transcription by proximal complexes occurs in the vicinity of an additional gene transcribed from the same DNA strand that terminates just a few hundred nucleotides upstream of *CPT1B* and encodes choline kinase beta (CHKB), a rate limiting enzyme in the CDP-choline pathway (Kennedy pathway) for phosphatidylcholine (PC) biosynthesis (Figure 3a) [109-111]. However, little is known about the regulation of the *CHKB* gene [112].

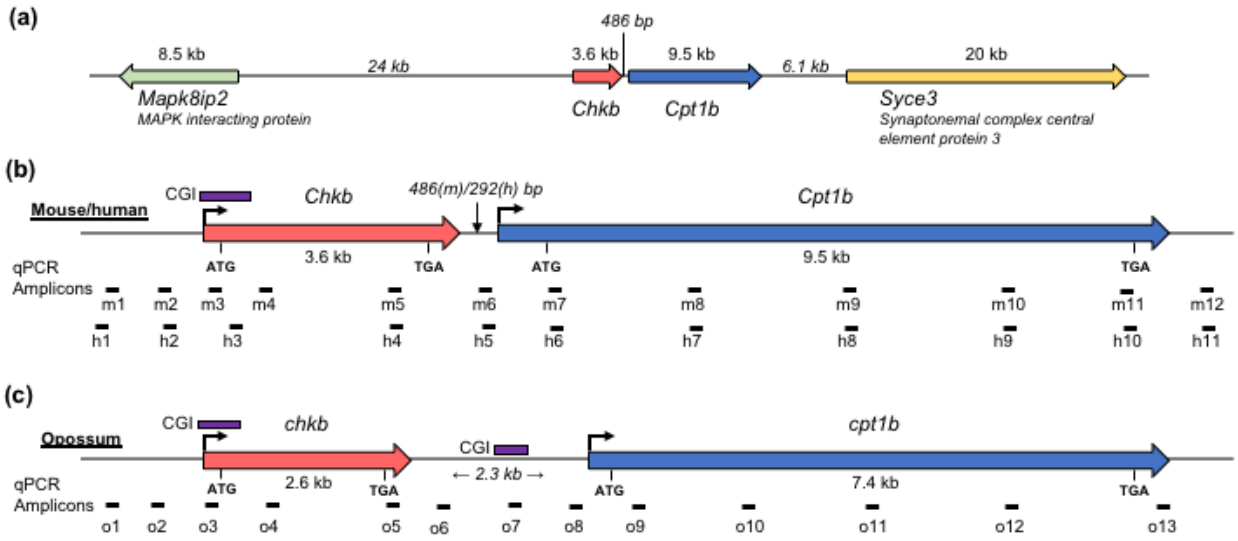

Figure 3. Structure of the CHKB-CPT1B locus. (a) Map of ~75 kb of mouse chromosome 15qE3 containing the *Chkb* and *Cpt1b* genes, indicated by arrows showing transcriptional orientation. (b) Map of the mouse and human *CHKB-CPT1B* locus. (c) map of the opossum *chkb-cpt1b* locus. Transcription start sites are indicated by bent arrows. Locations of start and stop codons are indicated below the map. CpG islands (CGI) are delimited by purple bars above the map. The locations of ChIP assay amplicons are indicated below the maps.

The significance of CPT1B activity in homeostasis is illuminated by the phenotype of homozygous CPT1B knockout (*Cpt1b*$^{-/-}$) mice, which die *in utero* before embryonic day 9.5-11.5 [12]. An absolute requirement for CPT1B expression is corroborated by the lack of severe *CPT1B* mutations in humans. However, a common *CPT1B* haplotype is associated with multiple markers of metabolic syndrome, consistent with the importance of CPT1B activity in maintaining metabolic homeostasis [113]. Previous studies demonstrated that a normal upregulation of the *CPT1B* gene in response to fatty acids observed in primary human skeletal myotube cultures was blunted in association with obesity, accompanied by repressive epigenetic changes and shifts in transcription factor occupancy at the *CPT1B* gene, and a reduction in cellular FAO [14, 15]. It was concluded from these findings that part of the basis for the positive lipid balance associated with obesity is the stable epigenetic programming of the *CPT1B* gene in the obesogenic environment, which reduces its capacity for lipid-mediated activation, and that changes in *CPT1B* expression can affect cellular oxidative capacity. This model for an essential, rate-limiting role of CPT1B in skeletal muscle lipid metabolism that is relevant to metabolic disease has been reinforced in several experimental systems [10, 12, 114]. As such, the regulation of this gene locus continues to be an attractive target for approaches to modify FAO therapeutically, and to understand the regulation of lipid metabolism in physiology and disease more broadly.

Consistent with the elevated FAO capacity of BAT, brown adipocytes possess distinct mechanisms for endogenous lipid storage and catabolism. Unlike white adipocytes, which store triglycerides in a single large cytoplasmic lipid droplet from which fatty acids are released into circulation by hormone-activated lipases in response to fasting and stress, brown adipocyte triglycerides are stored in multilocular lipid droplets as a source of fatty acids to be oxidized endogenously in numerous mitochondria to fuel UCP1-mediated thermogenesis, largely in response to β-adrenergic signaling stimulated by cold exposure [115-118]. The elevated β-

oxidation of fatty acids and resulting ATP synthesis can also fuel UCP1-independent thermogenesis in skeletal muscle in a mechanism involving the sarcoendoplasmic reticulum Ca^{2+}-ATPase (SERCA), indicating that the significance of FAO to non-shivering thermogenesis may extend beyond BAT [119-122].

In addition to UCP1 expression, and the enhanced lipolysis provided by multilocular lipid droplets, a defining parameter of the eutherian BAT phenotype is the elevated expression of genes associated with the oxidation of endogenously stored lipids to fuel thermogenesis. This is accompanied by the induction of mitochondrial biogenesis, collectively increasing the capacity for oxidative metabolism in BAT [115-118]. The activation of endogenous lipid catabolism also occurs in 'beige' adipocytes, an inducible phenotype recruited from white adipose tissue (WAT) that attains BAT-like characteristics in response to certain stimuli. Because of their elevated FAO capacity, brown and beige adipocytes contribute to the depletion of systemic fat and are thus a target for therapeutic approaches to counter obesity [116, 123, 124]. Informatively, heterozygous CPT1B knockout mice, which display a 50% reduction in carnitine palmitoyltransferase activity specifically in skeletal muscle, heart, and BAT, die of fatal hypothermia following a three-hour cold challenge [12], illuminating the significance of CPT1B activity in thermogenesis. Considering the similarities in the FAO functions of BAT and skeletal muscle, understanding the regulation of *CPT1B* in BAT would complement investigation of the *CPT1B* regulation in skeletal muscle in the context of thermoregulation and metabolic homeostasis.

Interestingly, despite its central function in eutherian non-shivering thermogenesis, UCP1 alone cannot account for the evolutionary emergence of BAT, as the functional UCP gene family (*UCP1*, *UCP2*, *UCP3*) has been extant since the bony fish common ancestor, with significant amino acid sequence identity among endothermic and ectothermic species. Furthermore, marsupials can acquire multilocular adipocytes that express UCP1 upon cold stimulation, suggesting that a pathway exists in marsupials to establish the framework for

augmented adipocyte lipolysis and thermogenic proton leakage. However, marsupials cannot form thermogenic BAT despite the presence of functional UCP1 [84-86, 125-127]. Indeed, phylogenetic analysis of UCP1 amino acid sequences among eutherian and non-eutherian species has confirmed that UCP1 has undergone neutral evolution, displaying no purifying adaptive changes in eutherians and supporting the observation that the biochemical properties of UCP1 in endothermic and ectothermic species are likely the same, mediating proton flux in common non-thermogenic processes [87, 88]. These observations show that UCP1 is necessary but not sufficient for canonical eutherian BAT and suggest that what had yet to evolve in the eutherian ancestor was a mechanism for enhanced endogenous mitochondrial FAO in extant cold-stimulated multilocular adipocytes that could supply UCP1-mediated proton flux at thermogenic levels [87].

In the present study, we expand the analysis of *CPT1B* gene regulation, informed by its central rate-limiting activity in striated muscle and BAT FAO. Specifically, we address the role that the closely apposed *CHKB* gene may play in the regulation of *CPT1B*. Beyond the potential for a transcriptional regulatory interaction due to their genomic proximity, a primary impetus for this study is the potential for the *CHKB* and *CPT1B* gene products to cooperatively reinforce cellular FAO. As mentioned above, CHKB is a rate-limiting enzyme in the synthesis of PC, a major membrane phospholipid and the predominant phospholipid comprising the inner mitochondrial membrane. Consistent with this essential function, both naturally occurring *CHKB* alleles in humans and induced mutations of *Chkb* in mice result in a megaconial mitochondrial muscular dystrophy due specifically to a deficiency of PC [31, 32, 36, 37, 39, 128-130]. Thus, an upregulation of CHKB has the potential to increase PC availability for mitochondrial membrane synthesis, which could reinforce mitochondrial FAO in concert with an increase in CPT1B protein levels such that the coordinated expression of the *CHKB* and *CPT1B* genes plays an important role in maintaining cellular FAO levels in muscle and BAT. The findings presented here show that the expression of the *CHKB* and *CPT1B* genes is linked within a unitary domain

of histone modifications and RNA polymerase II (RNAPII) transcription initiation targeted to the *CHKB* gene, and that this transcriptional and epigenetic linkage appears to have resulted from an intergenic deletion that occurred at the point of marsupial/eutherian divergence in mammalian evolution. Informed by the timing of this event relative to the emergence of BAT, the phylogeny of the *chkb-cpt1b* linkage, and the inability of functional changes in extant UCP1 to account for the acquisition of the eutherian BAT phenotype, these data support an evolutionary path for the emergence of BAT and homeothermy in eutherian mammals based on the serendipitous acquisition of a novel capacity for adipocyte FAO in a background of extant functional UCP1.

Results and Discussion

Transcription of the *Chkb* and *Cpt1b* genes is coordinated in mice.

As noted above**,** a series of early transfection studies identified several transcription factors that affect *CPT1B* transcription through proximal elements. However, the regulation of *CPT1B* in its natural chromatin locus *in vivo* has not been well examined, particularly in the context of the regulation of the nearby *CHKB* gene (Figure 3a). One possible outcome of the close apposition of the *CHKB* and *CPT1B* genes is an overlap in the *cis*-acting regulatory domains of the two genes, such that their transcription is linked by a common regulatory mechanism that controls the expression of both genes in concert. Spliced readthrough fusion transcripts that contain exons from both *CHKB* and *CPT1B* are readily observable in both mice and humans [131, 132], indicative of a potential transcriptional coincidence. In order to assess the coordinated regulation of the *Chkb* and *Cpt1b* genes, transcripts were measured in mouse tissues to determine whether they are transcribed in concert. As shown in Figure 4a, both genes are transcribed at a notably high level in BAT, followed by moderate relative levels in skeletal muscle and heart. The role of CPT1B in mediating a high capacity for FAO in these tissues has been established and is consistent with the unique demand. In contrast, they were both transcribed at low levels in WAT where FAO is limited, and liver, where CPT1A is the predominant CPT1 isoform. The *Chkb-Cpt1b* fusion transcript showed the same general distribution. The congruent tissue specificities of the *Cpt1b* and *Chkb* genes indicates the potential for the coordinated transcriptional regulation of these genes and suggests that *Chkb* and *Cpt1b* expression support a common process. In particular, the highest expression levels of both genes in BAT suggests that the coordinated expression of these genes may be especially relevant in this tissue. Immunoblot analysis of mouse tissues lysates corroborated the transcript levels, showing the highest CPT1B protein levels in BAT, followed by skeletal muscle and heart

(Figure 4b). The distribution of the additional fatty acid translocation pathway enzyme transcripts*, carnitine-acylcarnitine translocase (Slc25a20)* and *carnitine palmitoyltransferase 2 (Cpt2)*, were congruent with the relative levels of *Chkb* and *Cpt1b*, corroborating the elevated capacity for FAO in heart, skeletal muscle, and BAT (Figure 4a).

The *CPT1B* gene is regulated by long chain fatty acids as ligands for peroxisome proliferator-activated receptors (PPAR) that transactivate the *CPT1B* gene through fatty acid response elements (FARE) in the *CPT1B* proximal 5'-flanking region [97, 98, 101]. In order to further test whether regulation of the *CHKB* and *CPT1B* genes is coordinated, the effects of palmitate exposure on *Chkb* and *Cpt1b* transcript levels in cultured mouse myotubes was determined. As shown in Figure 4c, *Chkb* and *Cpt1b* mRNA levels increased in parallel in C2C12 myotubes in response to palmitate. These results indicate that transcription of the *Chkb* and *Cpt1b* genes can be regulated in concert by fatty acids, consistent with a regulatory linkage due to their genomic apposition. The level of the fusion transcript was also responsive to palmitate. These results, showing coordinated regulation and the formation of a spliced *Chkb-Cpt1b* fusion transcript, meet the operational classification of a 'conjoined' gene, a recognized phenomenon predicted at several hundred loci in the human genome [45-47].

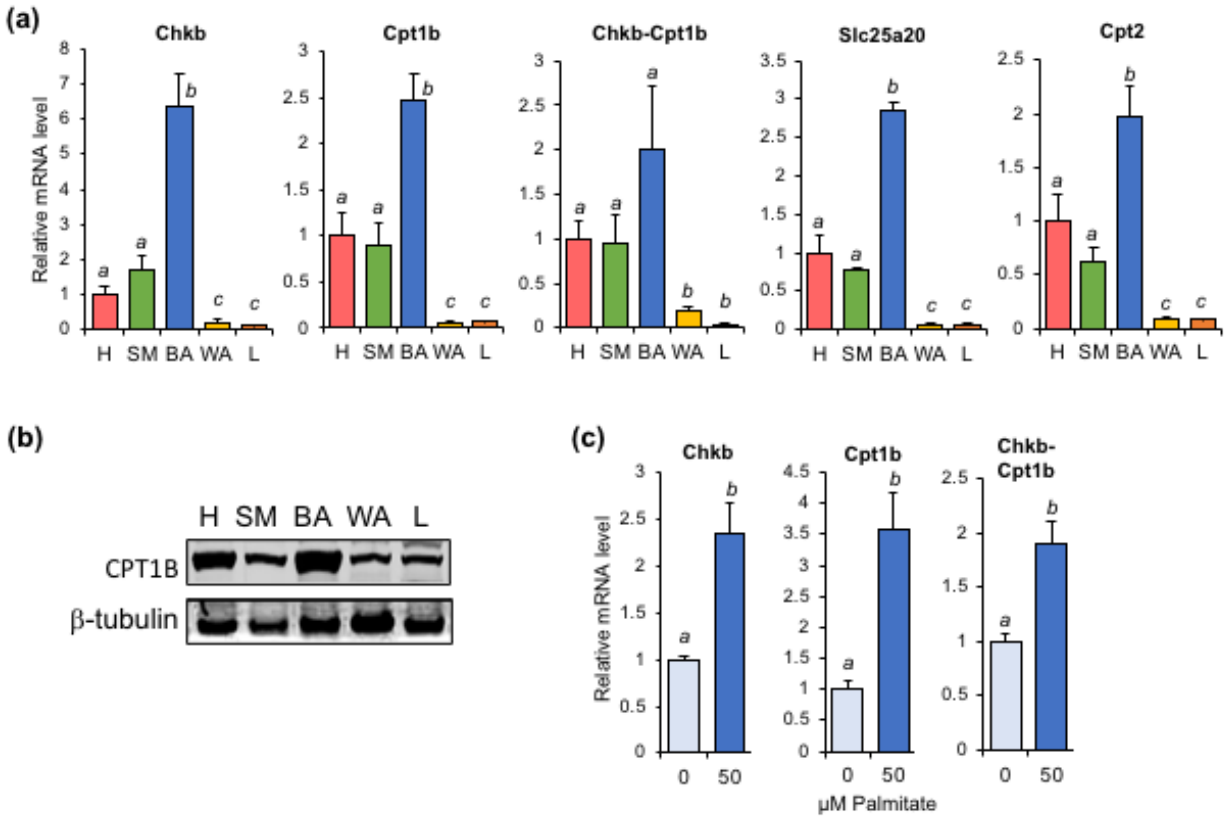

Figure 4. *Chkb* and *Cpt1b* expression are linked in mouse tissues and cell lines. (a) *Chkb* and *Cpt1b* are differentially expressed in concert in mouse tissues: heart (H), skeletal muscle (SM), brown adipose tissue (BA), white adipose tissue (WA), liver (L). Internally controlled mRNA levels (vs. *β2-microglobulin* mRNA) were normalized to the level in heart. The data indicate the mean ± *SE* (*n*=3). ANOVA indicated significant differences among the tissues (P<0.001). Levels not connected by the same letter are significantly different as determined by a Tukey-Kramer multiple comparison test (P<0.05). (b) Immunoblot of CPT1B protein levels mirrors the relative transcript levels. (c) *Chkb, Cpt1b* and *Chkb-Cpt1b* fusion transcript levels are modulated in concert by palmitate. The data indicate the mean ± *SE* (*n*=3).

The mouse *Chkb* and *Cpt1b* genes are encompassed in a unitary domain of histone modification and RNA polymerase II initiation targeted to the *Chkb* gene.

Gene activation and repression are mechanistically linked to specific chromatin modifications that contribute to the establishment and maintenance of gene activity states. In addition to DNA methylation, these epigenetic mechanisms involve nucleosomal histone modifications that have structural and cofactor recruitment functions in modulating gene activity [133, 134]. Of these modifications, the acetylation of histone 3 (H3) and histone 4 (H4) at multiple lysine residues alters electrostatic interactions leading to a more transcriptionally permissive chromatin state, and also functions to recruit nucleosome remodeling complexes. Histone acetyltransferases (HATs) are recruited to genes by DNA-binding transcription factors as components of coactivator complexes, and foci of histone acetylation above background levels are thus a marker for *cis*-acting regulatory sequences to which HAT activities are recruited, such as enhancers, promoters and transcription start sites. To map the pattern of histone acetylation across the contiguous *Chkb-Cpt1B* gene locus, a chromatin immunoprecipitation (ChIP) assay was performed with chromatin prepared from mouse C2C12 cell myotubes using a panacetyl-H3 antibody. Surprisingly, despite the parallel expression of the *Chkb* and *Cpt1b* genes, there was a single peak of histone acetylation at the 5' end of the *Chkb* gene (amplicons m3 and m4, Figure 5b), coincident with a known CpG island (CGI) [112] (Figure 5a). There was not a separate domain of histone acetylation targeted to the putative *Cpt1b* promoter, which indicates the possibility of a single *cis*-acting region at *Chkb* affecting the transcription of both genes in concert. A similar pattern of H3 acetylation was identified in differentiated primary mouse myotubes (Figure 5b), the mouse brown adipocyte cell line DE2.3 (Figure 5c) and differentiated primary mouse brown adipocytes (Figure 5d). These findings indicate that HAT activity is being recruited to a single site in the mouse *Chkb-Cpt1b* locus near the transcription start site of the *Chkb* gene in skeletal muscle and BAT.

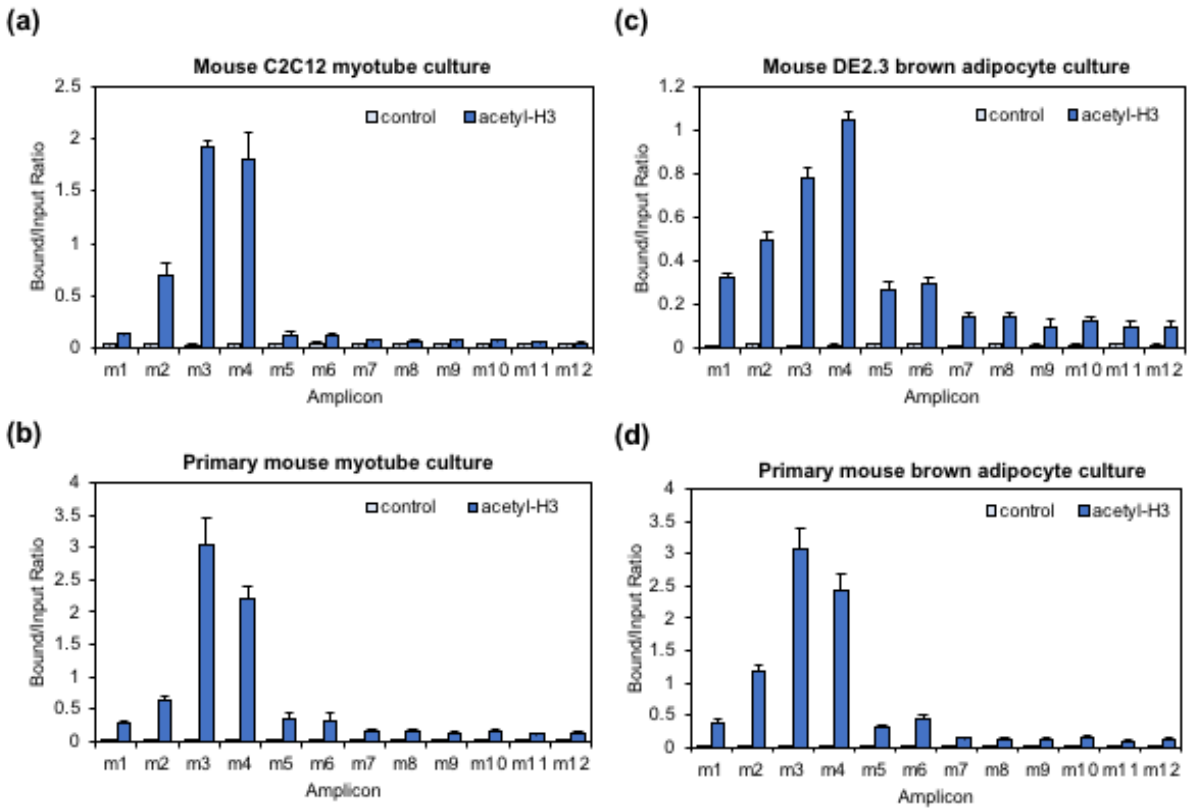

Figure 5. Elevated H3 acetylation is focused near the CHKB transcription start site in mice. Results of ChIP assays with a panacetyl-H3 antibody. ChIP assay enrichment levels determined by qPCR, normalized to 1% v/v of input chromatin. The exponential qPCR data (C_t values) was antilog transformed to a linear expression of bound/input ratios as Bound/Input=$2^{(\text{input Ct-bound Ct})}$. (a) Mouse C2C12 myotubes. (b) Primary mouse myotubes. (c) Mouse DE2.3 brown adipocytes. (d) Primary mouse brown adipocytes. Amplicons are mapped in Figure 3b.

In addition to histone lysine acetylation, foci of specific histone lysine methylation mediated by recruited histone methyltransferases are associated with domains of transcriptional activation. One such mark, H3 lysine 4 trimethylation (H3K4me3), is enriched around transcription start sites, and its distribution is informative in identifying sites of RNAPII initiation [134]. Similar to what was observed for histone acetylation, a ChIP assay of chromatin from mouse DE2.3 brown adipocytes using an antibody specific for H3K4me3 revealed a single peak at the *Chkb* gene (amplicon m3). No discrete peak of H3K4me3 was localized to a separate putative *Cpt1b* promoter. The colocalization of single foci of H3 hyperacetylation and H3K4me3 at the 5' end of the *Chkb* gene suggests that elements there are recruiting coactivator complexes to regulate both the *Chkb* and *Cpt1b* genes, considering their coordinated transcription (Figure 4). These observations also raise the possibility that *Chkb* may serve as the predominant site of RNAPII transcription initiation that leads ultimately to the production of both *Chkb* and *Cpt1b* mRNA. Indeed, the readily detected spliced readthrough transcript that contains exons from both genes provides direct evidence for the processivity of RNAPII through both genes (Figure 6). In order to directly map the distribution of RNAPII transcription initiation throughout the *Chkb-Cpt1b* locus, a ChIP assay was performed using antibodies specific for the initiating (serine 5-phosphorylated) and elongating (serine 2-phosphorylated) forms of RNAPII with differentiated mouse DE2.3 brown adipocytes. The results indicated that the initiating serine 5-phosphorylated form of RNAPII was localized predominantly at the 5' end of *Chkb* (Figure 6b), while the elongating serine 2-phosphorylated form of RNAPII was distributed downstream in a continuous domain across both the *Chkb* and *Cpt1b* genes (Figure 6c). Significantly, there was not an independent peak of serine 5-phosphorylated RNAPII observed at the putative *Cpt1b* promoter, indicating a lack of independent transcription initiation there. Considered in conjunction with the patterns of histone modification, these data reinforce the possibility that the *Chkb* promoter and associated regulatory sequences are controlling the transcription of both the *Chkb* and *Cpt1b* genes. However, the responsiveness of *Chkb* to

palmitate indicates that a reciprocal regulatory crosstalk may also result from the apposition, in light of the established fatty acid responsiveness of the proximal *Cpt1b* 5'-flanking region.

Considering the high level of mouse *Chkb* and *Cpt1b* transcription in BAT (Figure 4), it would be informative to determine if this locus is regulated by known mediators of the BAT phenotype. The transcriptional coactivator PRD1-BF1-RIZ1 homologous domain containing 16 (PRDM16) has a central function in controlling a bi-directional switch in cell fate between brown adipocytes and skeletal myoblasts in development. PRDM16 manifests this effect through the activation of a suite of BAT selective genes, while also repressing the expression of white adipocyte specific genes [58, 63]. The seminal role of PRDM16 in the establishment of the BAT lineage raises the question of whether the *Chkb-Cpt1b* locus is also a direct target of PRDM16 action. However, a ChIP assay for PRDM16 showed no evidence for localization to the *Chkb-Cpt1b* locus in mouse BAT or DE2.3 adipocytes (data not shown), suggesting that any involvement of PRDM16 in the regulation of this locus is likely indirect, mediated by downstream effectors of PRDM16 activity, or through elements beyond the genomic boundaries of the current analysis. Future studies will focus on the dynamics and molecular details of *Chkb-Cpt1b* locus regulation during brown adipocyte, white adipocyte, and myocyte ontogeny to address this question more rigorously.

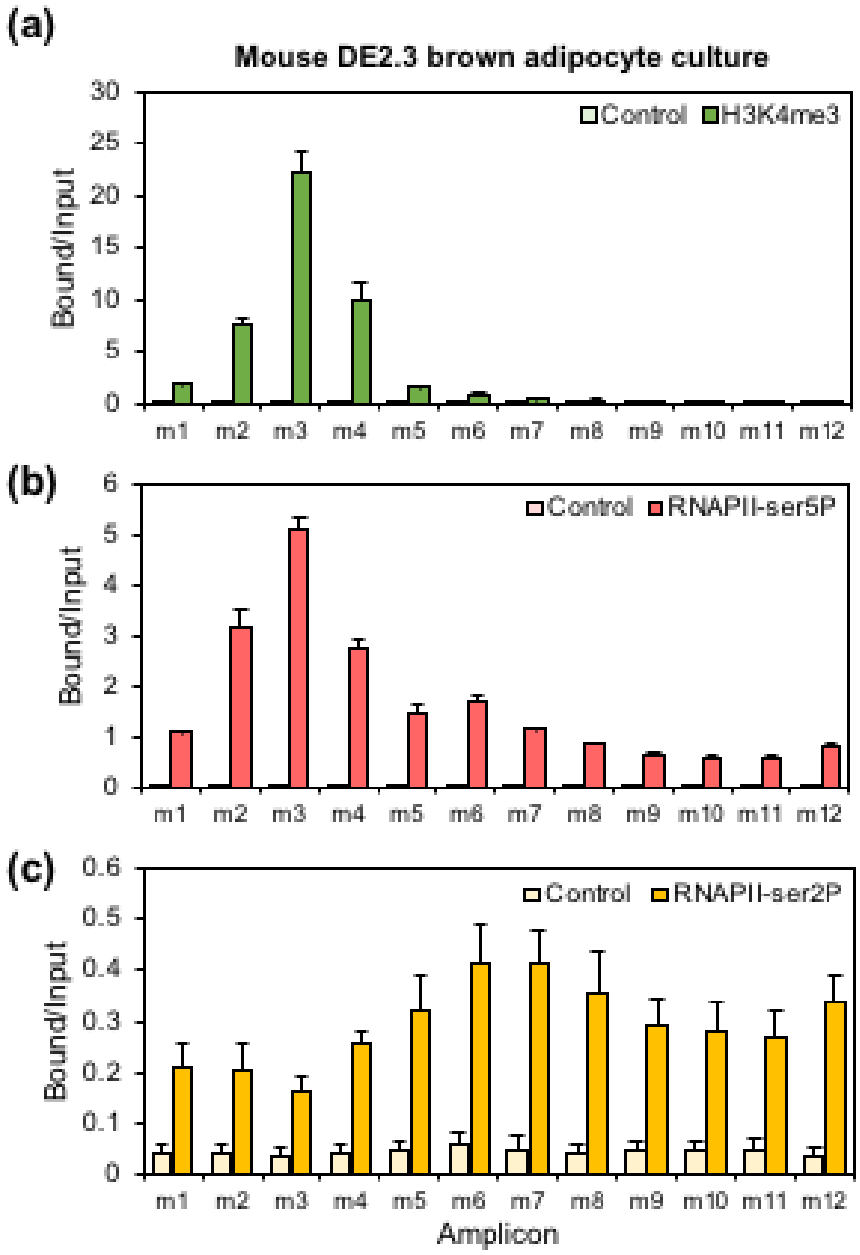

Figure 6. H3 lysine 4 trimethylation and RNAPII initiation are focused near the CHKB transcription start site in mouse DE2.3 brown adipocytes. Results of ChIP assays. (a) H3 lysine 4 trimethylation. (b) RNAPII serine 5 phosphorylation (initiating form). (c) RNAPII serine 2 phosphorylation (elongating form). Amplicons are mapped in Figure 3b.

A single domain of histone modification is targeted to the *CHKB* gene in primary human myotubes.

If the unitary epigenetic domain associated with conjoined *chkb* and *cpt1b* genes is functionally significant to FAO in muscle and BAT, the conservation of a tight intergenic distance in humans predicts that single foci of histone modifications targeted to *chkb* may be a common feature of eutherians. To begin to address this, primary human myoblasts were expanded and differentiated into myotubes *ex vivo*, and ChIP assays performed with acetyl-H3 and H3K4me3 antibodies. Single foci of H3 acetylation (Figure 7a) and H3 lysine 4 trimethylation (Figure 7b) were localized to the CGI at the 5' end of *CHKB* (amplicons h2 and h3, Figure 1), similar to the pattern observed in mice (Figure 5, 6). This result suggests that conjoined *chkb* and *cpt1b* genes within a unitary epigenetic domain targeted to *chkb* may be common to eutherians.

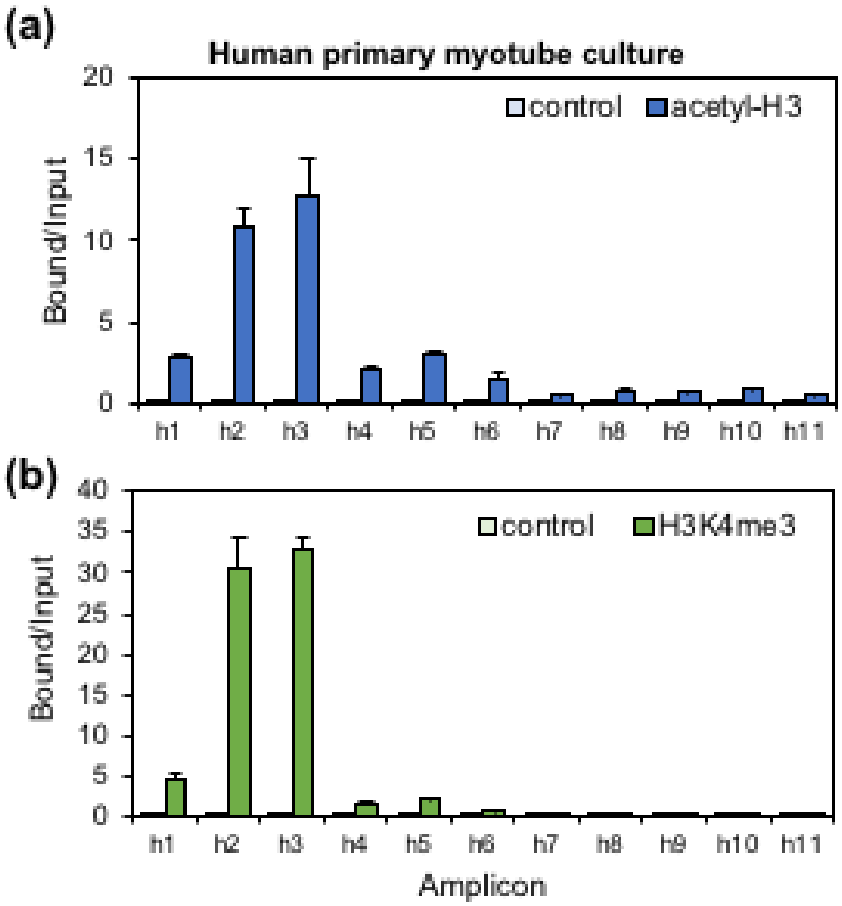

Figure 7. A single peak of histone modification is localized to the 5' end of the human *CHKB* gene in primary skeletal muscle myotubes. Results of ChIP assays for (a) H3 acetylation and (b) H3 lysine 4 trimethylation. Amplicons are mapped in Figure 3b.

The apposition of the *chkb* and *cpt1b* genes occurred in mammals subsequent to the divergence of marsupials.

The activities of CHKB and CPT1B, rate-limiting for PC and acylcarnitine synthesis respectively, have the potential to mutually reinforce the capacity for mitochondrial FAO, which is supported by the outcomes of naturally occurring and induced mutations of these genes. Thus, it is possible that this regulatory linkage is an important component of lipid metabolism in mammals. In order to address this possibility, a phylogenetic approach was used to map the genomic arrangement of the *chkb* and *cpt1b* genes among extant animals with sequenced and annotated genomes that span vertebrate evolution. As shown in Figure 8a, the *chkb* and *cpt1b* genes are unlinked in lamprey (Petromyzon marinus) and a shark (Callorhinchus milii) and became distally syntenic prior to the divergence of bony fish, indicated by the 9.3 kb intergenic distance in zebrafish (Danio rerio) and coelacanth (Latimeria chalumnae). This distal synteny is conserved in amphibians (e.g. Xenopus tropicalis) but was lost in birds (e.g. Gallus gallus) and reptiles (e.g. Anolis carolinensis) [131, 132].

The intergenic distance is reduced in a marsupial (*Monodelphis domestica*) to 2.3 kb, and in the earliest diverging placental mammals, represented by extant Xenarthrans such as the armadillo *Dasypus novemcinctus*, the intergenic distance is about 1 kb. An average intergenic distance of about 500 bp between syntenic *chkb* and *cpt1b* genes appears to be highly conserved among eutherian mammals, suggestive of positive selection. Thus, the timing of this transition to a close intergenic distance coincides with the emergence of eutherian mammals. The *chkb-cpt1b* loci of two additional marsupials, a wallaby (*Notamacropus eugenii*) and Tasmanian devil (*Sarcophilus harrisii*), are similar to *Monodelphis*, supporting a marsupial/eutherian dichotomy with respect to the *chkb-cpt1b* locus [131, 132]. Considering that many of the physiological and allometric differences between marsupials and eutherians are connected to distinctions in basal metabolic rate and energy expenditure [135-138], a modification of lipid metabolism at this point in mammalian evolution due to a novel *chkb-cpt1b* linkage is potentially significant. Consistent

with this prediction, the most phylogenetically ancient mammal in which canonical thermogenic BAT depots have been identified is the tenrec *Echinops telfairi*, a protoendothermic eutherian mammal that maintains elevated body temperatures during reproduction through transient BAT formation around reproductive organs [94]. Thus, as the *E. telfairi* genome contains ~700 bp *Chkb-Cpt1b* intergenic distance (Figure 8a), the emergence of BAT occurred relatively soon after the condensation of the *chkb-cpt1b* locus that coincides with the divergence of marsupials and eutherians.

***Chkb-cpt1b* synteny is constrained by functional UCP1.**

In considering the evolutionary emergence of BAT, it's intriguing to note that the *UCP1* gene was also lost in in the common saurian ancestor of birds and reptiles [2] in conjunction with the loss of *chkb-cpt1b* synteny in this clade (Figure 8a). While reptiles remain as ectothermic heterotherms, relying on environmental warming to increase body temperature, birds have taken a distinct evolutionary route to achieve an elevated capacity for endothermy based on UCP1-independent thermogenesis in hypertrophic skeletal muscle depots [2, 139, 140]. The coincident loss of both *chkb-cpt1b* synteny and the *UCP1* gene in this clade, in contrast to the conserved proximal *chkb-cpt1b* synteny in eutherians, supports a model in which the loss of the *UCP1* gene (and thus the potential for thermogenic proton leakage) eliminated the selective pressure to maintain conjoined *chkb* and *cpt1b* genes. In further support of this, despite the nearly complete loss of *chkb-cpt1b* synteny among saurians, *chkb-cpt1b* synteny with ~8 kb intergenic distance was retained in two distantly related members of this clade, the golden eagle (*Aquila chrysaetos canadensis*) and American alligator (*Alligator mississippiensis*) (Figure 8b) [131, 132]. This indicates that *UCP1* gene loss in the saurian ancestor preceded the loss of c*hkb-cpt1b* synteny, which then occurred independently in multiple lineages, suggesting that UCP1 function generates significant selective pressure to maintain *chkb-cpt1b* linkage.

As an additional assessment of the apparent correlation between functional UCP1 and *chkb-cpt1b* synteny, we compared the distribution of intergenic distance in UCP1-expressing eutherians with those that have lost UCP1 function (and thus the ability to develop BAT) due to genomic loss of the *UCP1* gene [141]. As shown in Figure 8c, the intergenic distance in $UCP1^{-}$ eutherians is significantly greater than in $UCP1^{+}$ eutherians, consistent with the relaxation of a constraint for tight *chkb-cpt1b* synteny resulting from the loss of UCP1 function in these lineages. These observations of a strong correlation between *chkb-cpt1b* apposition and UCP1 activity provide significant phylogenetic support for the causal relevance of conjoined *chkb* and *cpt1b* genes to the emergence of eutherian BAT.

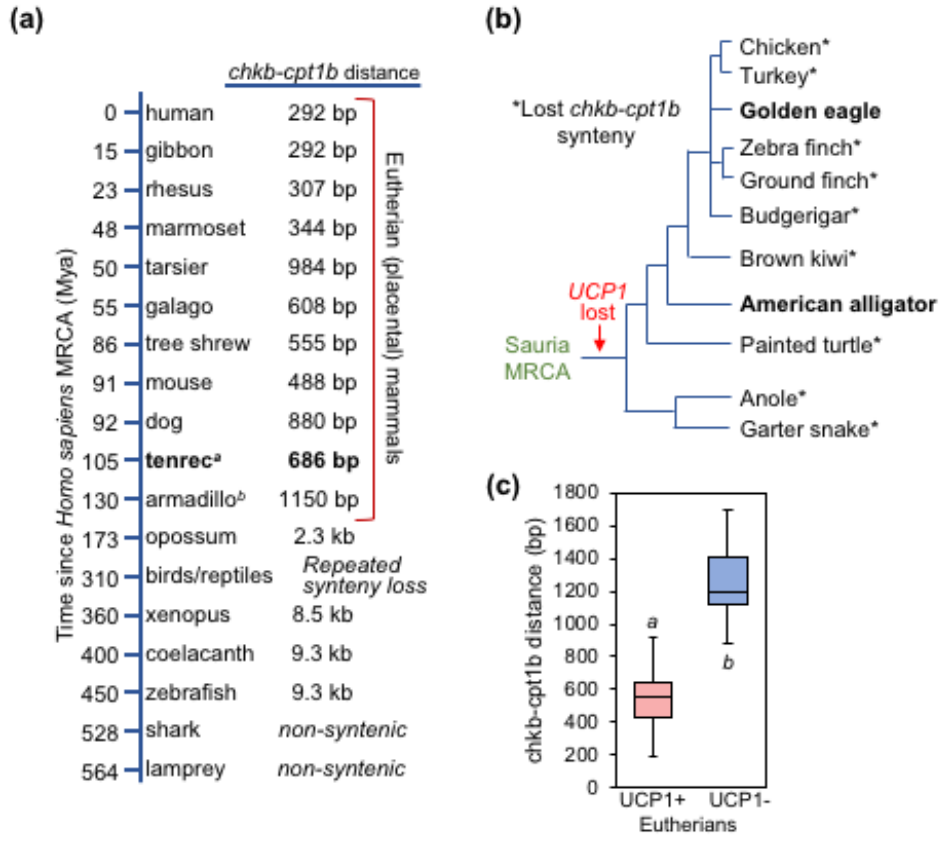

Figure 8. *chkb-cpt1b* synteny is constrained by functional UCP1. (a) Cladogram of *chkb-cpt1b* locus structure. *H. sapiens* MRCA (most recent common ancestor) divergence times were adapted from a molecular timescale of vertebrate evolution [1]. [a]Tenrec represents the oldest eutherian lineage in which canonical BAT has been shown to develop. [b]Armadillo lost the *UCP1* gene and is included in the data shown in panel (c). (b) Phylogenetic tree of *chkb-cpt1b* synteny in Sauria (extant birds and reptiles). Species that have lost *chkb-cpt1b* synteny are indicated by an asterisk, and species retaining *chkb-cpt1b* synteny are indicated in bold. UCP1 was lost in the saurian common ancestor [2]. (c) Box and whisker plot comparing the *chkb-cpt1b* intergenic distance in *UCP1*$^{+}$ and *UCP1*$^{-}$ eutherians (species listed in Materials and Methods). The intergenic distances were significantly different as determined by ANOVA ($P<0.05$).

Expression of the *chkb* and *cpt1b* genes is not coordinated in the marsupial Monodelphis domestica.

The working hypothesis that emerged from the presented findings is that the apposition of the *chkb* and *cpt1b* genes resulted in a functional transition from their independent expression to a coordinated mode of regulation due to the loss of functional intergenic sequences. This model predicts that the expression of *chkb* and *cpt1b* would be unlinked in species possessing a pre-eutherian locus structure, in contrast to what was observed in mice. Fortunately, the existence of an experimentally accessible marsupial provides a natural experiment to test this hypothesis, given the dichotomy of eutherian and marsupial *chkb-cpt1b* locus structure. The small laboratory opossum *Monodelphis domestica*, which possesses a 2.3 kb intergenic distance (Figure 5c, 6a), was employed to assess this prediction by surveying the tissue-specific expression of *chkb* and *cpt1b*. In contrast to the colocalization of *Chkb* and *Cpt1b* transcripts and the *Chkb-Cpt1b* fusion transcript observed in mice (Figure 4b), the expression of the *chkb* and *cpt1b* genes was unlinked in the marsupial, and the relative levels of expression of the genes among the tissues was different from that in mice with no fusion transcripts detected (Figure 9a). While both *chkb* and *cpt1b* were transcribed at high levels in opossum skeletal muscle, their mRNA levels in heart were low, in contrast to the more similar expression in these tissues in mice. The distribution of the opossum *cpt2* and *slc25a20* transcripts parallels the elevated expression of *cpt1b* in skeletal muscle, consistent with their roles in the same fatty acid transport pathway. In addition, while both genes are expressed at low levels in mouse liver (Figure 4), *chkb* (but not *cpt1b*) is transcribed at a moderately high level in opossum liver. However, the most striking difference is the highest level of *chkb* transcripts (and concomitant low level of *cpt1b* mRNA) in opossum WAT, which contrasts with the low level of transcription of both genes in mouse WAT (Figure 4b). Collectively, these observations reinforce the possibility that the deletion of the *chkb-cpt1b* intergenic sequence precipitated the coordination of *chkb* and *cpt1b* expression in eutherians, such that the remaining intergenic sequence in opossum,

albeit reduced in size from the ~9 kb span in fish and amphibians, is sufficient to maintain independent *chkb* and *cpt1b* expression likely through the action of specific retained *cis*-acting regulatory sequences in this region.

These observations also suggest that the opossum *chkb* gene is associated with a strong WAT-specific enhancer, which may have directed novel *cpt1b* expression in this compartment following the intergenic deletion that brought *cpt1b* closer to the *chkb* 5'-flanking region if it was functional in the MRCA. Notably, the intergenic sequence in *Monodelphis* contains a CGI that may contribute to the expression of marsupial *cpt1b* in skeletal muscle (and its suppression in WAT and liver) and is encompassed in the intergenic deletion in eutherians (Figure 5c). The loss of putative functional elements in this region may have facilitated the co-option of *cpt1b* regulation by *chkb* 5'-flanking sequences, leading to the novel co-expression of *chkb* and *cpt1b* in an adipocyte lineage in the eutherian ancestor that provided for the subsequent acquisition of the BAT phenotype, where both genes are highly expressed (Figure 4b). Furthermore, the high degree of apparent positive selection for *chkb-cpt1b* apposition among eutherian mammals (in light of the repeated loss of synteny in saurians) suggests that this event may have had significant implications for mammalian physiology and the emergence of placental mammals.

Independent domains of histone modification and RNA polymerase II initiation are targeted to the *chkb* and *cpt1b* genes in opossum.

The unlinked expression of the *chkb* and *cpt1b* genes in opossum, in contrast to the parallel tissue distribution of their expression in mice, raises the question of whether this disparity was also associated with a distinction in chromatin modification at the *chkb-cpt1b* locus that could reflect distinct mechanisms of gene regulation at the eutherian and marsupial loci. To address this, ChIP assays for acetyl-H3 were performed with chromatin prepared from opossum heart and skeletal muscle. The results showed a marked distinction from the patterns of H3

acetylation observed in mouse cells. In contrast to a single peak of histone acetylation localized to the 5' end of *Chkb* in mice (Figure 5), distinct peaks of H3 acetylation and were localized to the *chkb* gene (amplicons o3 and o4) and to the CGI located upstream of the *cpt1b* gene in opossum (amplicon o7) (Figure 9 b, c). The distinct peak upstream of *cpt1b* displayed the highest level of H3 acetylation in the opossum locus and was absent in the mouse locus (Figure 5). Since this sequence was lost in the eutherian intergenic deletion it may have eliminated a putative enhancer or chromatin domain boundary function, or elements associated with the independent transcription of *cpt1b*, allowing the *chkb* 5'-flanking region to co-opt the transcriptional regulation of *cpt1b*. Concomitantly, this may have also allowed conserved functional sequences in the *cpt1b* proximal 5'-flanking region, such as the FARE and other elements, to reciprocally affect the transcription of *chkb* as supported by the coordinated response of mouse *Chkb* and *Cpt1b* to palmitate (Figure 4c).

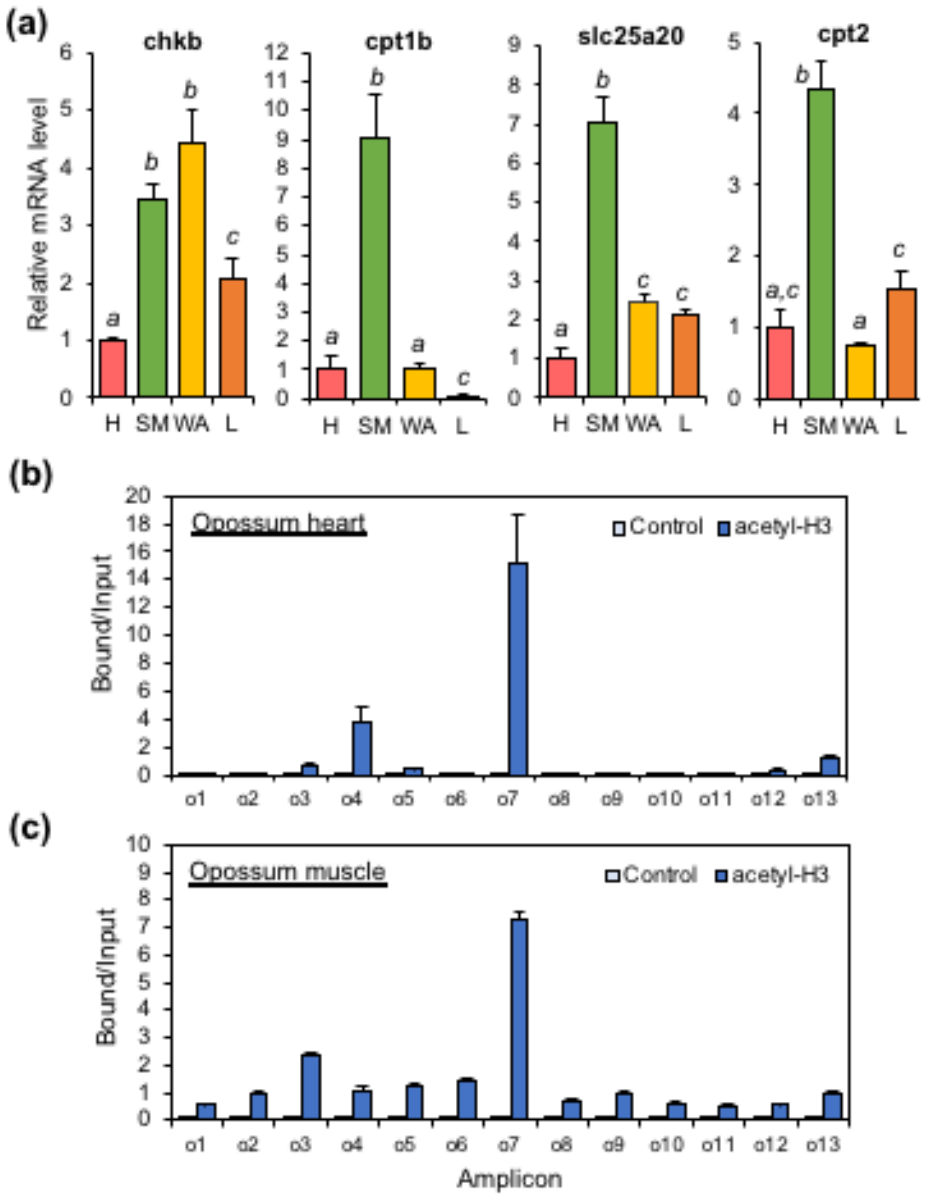

Figure 9. Chkb and cpt1b expression are unlinked in the opossum *M. domestica*. (a) Tissue distribution of *chkb* and *cpt1b* transcription in opossum: heart (H), skeletal muscle (SM), white adipose tissue (WA), liver (L). Internally controlled mRNA levels (vs. *β2-microglobulin* mRNA) were normalized to the level in heart. Results of ChIP assays with a panacetyl-H3 antibody in opossum (b) heart and (c) skeletal muscle.

In order to further assess the potential for independent transcription of the opossum *cpt1b* gene, a ChIP assay for H3K4me3 was performed as a marker for sites of transcription initiation in opossum skeletal muscle where both genes are highly expressed (Figure 9). As was seen in mice (Figure 4a), a peak of H3K4me3 was localized at the 5' end of the opossum *chkb* gene (amplicon o3; Figure 10a). However, in contrast to the mouse locus, a second robust peak of H3K4me3 was also localized upstream of *cpt1b* at the CGI (amplicon o7). A ChIP assay for serine 5-phosphorylated RNA Pol II in opossum skeletal muscle also showed a second domain of recruitment adjacent to the intergenic CGI island (amplicon o6; Figure 10b), consistent with intergenic transcription initiation. Concomitantly, a ChIP for serine 2-phosphorylated RNA Pol II revealed discrete domains of elongating RNAPII downstream of these sites of initiation (Figure 10c). Considered together, the patterns of H3 acetylation and methylation and RNA Pol II recruitment support the inference from the disparate tissue specific expression of *chkb* and *cpt1b* that these genes are independently transcribed in opossum. Significantly, the sites of intergenic histone modification and RNAPII initiation in the opossum locus are colocalized with the CGI, suggesting that specific sequences in this region are involved in the independent transcription of *cpt1b* that were lost in the eutherian locus. It should be noted that the precise 5' ends of opossum *cpt1b* transcripts have not been mapped, such that transcription initiation in the CGI region is possible. Our interpretation of these findings is that it is the loss of specific functional elements in the intergenic region, not the reduction in intergenic distance *per se*, that distinguishes the eutherian and marsupial *chkb-cpt1b* loci. However, as *cpt1b* then became dependent on the *chkb* promoter for transcription in this model, the short intergenic distance became functionally significant and selected for in eutherians.

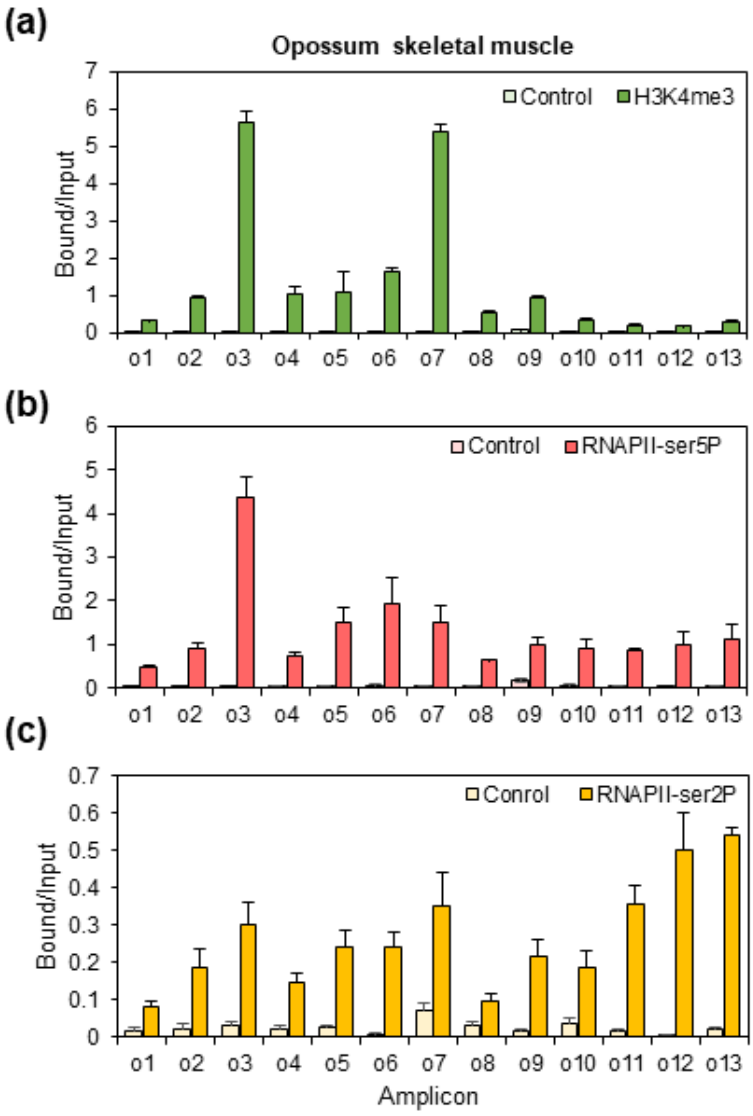

Figure 10. H3 lysine 4 trimethylation and RNAPII initiation are focused near the *Chkb* transcription start site and the *Cpt1b* 5' CGI in opossum skeletal muscle. Results of ChIP assays. (a) H3 lysine 4 trimethylation. (b) RNAPII serine 5 phosphorylation (initiating form). (c) RNAPII serine 2 phosphorylation (elongating form). Amplicons are mapped in Figure 3c.

A model for the emergence of BAT in a background of neutral UCP1 evolution.

In extant eutherians, the distinguishing phenotype of brown adipocytes includes the coincidence of multilocular lipid droplets, UCP1 expression, numerous mitochondria, and endogenous FAO. However, UCP1 has been under neutral selection since the emergence of bony fish, and a persistent paradox is the ability of marsupials to produce multilocular adipocytes that express functional UCP1 but cannot form thermogenic BAT. Indeed, the *ex vivo* induction of multilocular adipocytes from embryonic limb bud mesenchyme has been demonstrated in chicken [2], and cold-induced multilocular adipocytes have been shown in Muscovy ducks, but with few mitochondria compared to mammalian BAT [142]. These observations show that the underlying mechanism for inducing multilocular adipocytes, which increases the surface area for lipolysis, has been in place since before the emergence of mammals and suggest that the defining characteristic of eutherian BAT is thus an elevated capacity for FAO in these cells. We propose that the coordinated high levels of co-expression of *chkb* and *cpt1b* in a multilocular adipocyte lineage due to an intergenic deletion in the eutherian ancestor was a precipitating event that provided for the elevated mitochondrial FAO that caused these cells to become BAT and may have thus contributed significantly to the evolution of eutherian mammals.

In this model (Figure 11), *chkb* and *cpt1b* were initially independently transcribed from discrete insulated regulatory domains in the therian ancestor, as is retained in marsupials. Upon deletion of the majority of intergenic sequence, a putative regulatory region associated with the intergenic CGI in opossum that maintained independent *cpt1b* expression was lost, and dominant regulation of *cpt1b* was subsequently co-opted by *chkb*-associated regulatory sequences (with potential reciprocal effects of *cpt1b*-associated elements on *chkb* transcription), which drove novel elevated *cpt1b* expression in adipocytes in parallel with extant *chkb* expression. The resulting increase in both mitochondrial fatty acid flux and PC synthesis for mitochondrial membrane biogenesis drove an increase in endogenous FAO, providing for

thermogenic levels of proton flux and the emergence of the brown adipocyte phenotype. The co-expression of *chkb* and *cpt1b*, likely through independently functioning factor binding sites at each gene, may also be important for muscle FAO, and appears to have been in place prior to the marsupial/eutherian divergence, evidenced by their relatively high co-expression in opossum skeletal muscle (Figure 9a).

While the juxtaposition of the *cpt1b* gene to a putative strong WAT-specific enhancer (as appears to be associated with the opossum *chkb* gene) provides a potential molecular mechanism for the novel expression of *cpt1b* in an adipocyte lineage (and thus the capacity for endogenous FAO), this also raises the question of the role of elevated *chkb* expression observed in marsupial WAT (Figure 9). An intriguing possibility is that it is causally linked to the ability of marsupials to generate multilocular adipocytes in WAT, as is also characteristic of eutherian beige and brown adipocytes. Cytoplasmic lipid droplets are composed of a triglyceride core encapsulated by a phospholipid monolayer comprised predominantly of PC [143] such that the formation of multilocular droplets with greater total lipid droplet surface area may require elevated *de novo* PC synthesis, accounting for the high level of *chkb* expression in opossum WAT. This would provide for elevated lipolysis to generate secreted fatty acids for FAO in skeletal muscle, where *cpt1b* expression is high (Figure 9a). Thus, the primordial marsupial BAT may only be lacking the capacity for endogenous FAO provided by *cpt1b* expression to become thermogenic. *Chkb* expression was subsequently maintained at a high level in eutherian BAT upon the linkage of the *chkb* and *cpt1b* genes (Figure 4b), providing PC for both lipid droplet and mitochondrial membrane synthesis. The low expression levels of *Chkb* and *Cpt1b* in mouse WAT (in contrast to their highest expression in BAT) further suggests the maintenance of a distinct pathway in eutherians for the ontogeny of white adipocytes with a transcriptionally inactive *chkb-cpt1b* locus (Figure 4b), consistent with current models for the distinct developmental origins of white and brown preadipocytes [144].

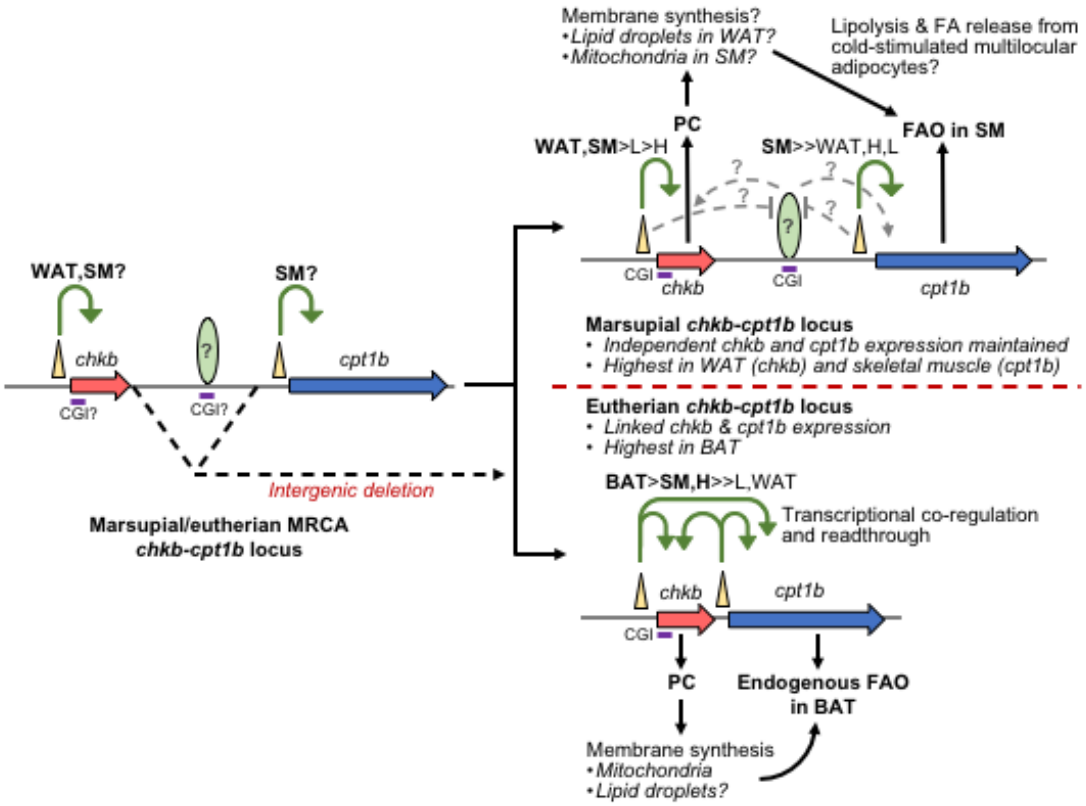

Figure 11. Model for the structural and functional divergence of the marsupial and eutherian *chkb-cpt1b* loci. In the therian MRCA, the distally syntenic *chkb* and *cpt1b* genes were both expressed in skeletal muscle (SM), but only *chkb* was expressed in WAT as mediated by proximal elements (gold triangles, bent arrows). This pattern was maintained in marsupials due to the retention of the CGI associated with the marked peak of acetylation upstream of *cpt1b*, where putative regulatory complexes (green oval) may be insulating *cpt1b* from *chkb*-associated elements (gold triangles), and/or providing independent enhancer function (dashed arrows). In eutherians, the deletion of the intergenic CGI eliminated the associated putative insulator and/or activating function, allowing the regulation of *cpt1b* to be co-opted by *chkb*-associated sequences, with reciprocal crosstalk from *cpt1b*-associated elements, which resulted in novel co-expression of these genes in an adipocyte lineage that became BAT.

Conclusions

The acquisition of BAT was a major evolutionary achievement in mammals that significantly affected the scope of mammalian physiology, life history, and geography. In addition, the FAO capacity of BAT has made it a potential therapeutic target to counteract a positive lipid balance in the treatment of obesity. The central role of UCP1 in BAT function has been firmly established, both biochemically and genetically, but the mechanistic basis for the emergence of BAT in eutherian mammals has remained elusive. While previous studies have characterized the functional evolution of the UCP gene family, they have also revealed the insufficiency of UCP1 evolution alone to account for BAT function, and led to the prediction that the acquisition of elevated mitochondrial oxidative capacity likely played a predominant role [87]. Consistent with this prediction, we conclude from the findings presented here that the convergence of the *chkb* and *cpt1b* genes in a common transcriptional regulatory domain in the eutherian ancestor resulted in their coordinated expression, and provided for the novel attainment of elevated mitochondrial FAO in an extant multilocular adipocyte lineage that became canonical thermogenic BAT as a direct result of this event. Ongoing studies will resolve the molecular mechanisms of the transcriptional and epigenetic co-regulation of the *chkb* and *cpt1b* genes, and the significance of their co-expression to mitochondrial structure, cellular FAO capacity, and thermogenesis in eutherian striated muscle and BAT to test this hypothesis.

Materials and Methods

Cell culture.

Mouse C2C12 myoblast cells (ATCC, Manassas, VA), DE2.3 brown preadipocyte cells [145] (a gift of Bruce Spiegelman, Harvard University), and primary female human myoblasts (a gift of Joseph Houmard, East Carolina University) were maintained in Dulbecco's modified Eagle's medium (DMEM, Mediatech, Manassas, VA) supplemented with 10% fetal bovine serum (Gemini Bio Products, West Sacramento, CA), 100 U/ml penicillin, 100 µg/ml streptomycin, and 250 ng/ml amphotericin B (Gibco, Gaithersburg, MD). Upon reaching 80-90% confluency, differentiation was induced by replacing growth medium with low-serum DMEM containing 1% FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, and 250 ng/ml amphotericin B. On day six of differentiation verified microscopically by the formation of syncytial myotubes and multilocular lipid droplets, respectively, myotubes and brown adipocytes were harvested for analysis. To prepare BSA-conjugated palmitate, a solution of 1.7 mM fatty acid free BSA (Sigma-Aldrich, St. Louis, MO) and 750 mM NaCl was prepared. Separately, 61.2 mg of sodium palmitate (TCI America, Portland, OR) was dissolved in 17.6 ml of 750 mM NaCl at 70^{o}C with stirring until clear. To make BSA-conjugated 5 mM palmitate, the palmitate solution was mixed with 20 ml of pre-warmed BSA solution at 37^{o}C with stirring until clear, then diluted 1:2 with 750mM NaCl. A control solution containing only BSA was prepared in parallel. Solutions were adjusted pH 7.4 with 1N NaOH, sterile filtered, and stored -20^{o} C. For the palmitate treatment experiment, cells were washed with PBS and treated with fresh media containing 50 µM BSA-palmitate or BSA vehicle control for 4 h prior to transcript analysis.

Mouse and opossum tissue harvest and primary cell culture.

Tissues from female C57BL/6 mice (a gift of Kym Gowdy, East Carolina University) were snap frozen in liquid nitrogen and stored at -80^{o} C until analysis of mRNA and chromatin.

Cultures of primary mouse myotubes were prepared as previously described [146, 147]. In short, freshly harvested gastrocnemius muscle was digested in 0.05% trypsin with 0.53 mM EDTA and cultured in DMEM containing 4.5 g/L glucose, 10% FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, and 250 ng/ml amphotericin B. After repeated selective passage to remove fibroblasts, enriched myoblasts were differentiated in DMEM containing 1.0 g/L glucose, 1% FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, and 250 ng/ml amphotericin B. Cultures of primary mouse brown preadipocytes were prepared as previously described [148]. Briefly, intrascapular brown adipose tissue was digested in 1.5 mg/ml collagenase A and cultured in the same growth medium as primary skeletal muscle cells. Differentiation was induced with DMEM containing 4.5 g/L glucose, 10% FBS, 0.5 µM IBMX, 5 µM dexamethasone, 20 nM insulin,1 nM T3 , 125 µM indomethacin (each from Sigma-Aldrich, St. Louis, MO), 100 U/ml penicillin, 100 µg/ml streptomycin, and 250 ng/ml amphotericin B. The Sears Lab (UCLA) maintains a breeding colony of opossums (*Monodelphis domestica*) using established protocols [149, 150]. This colony is comprised of pedigreed animals from the Texas Biomedical Research Institute (San Antonio, TX) and their descendants. Adult female opossums were euthanized by carbon dioxide inhalation to effect followed by cervical dislocation, and appropriate tissues dissected. All procedures have been approved by the ECU and UCLA IACUC.

mRNA purification and qRT-PCR.

All primers for quantitative reverse transcription and PCR (qRT-PCR; Table 1) and genomic quantitative PCR (qPCR; Table 2) were designed and validated as previously described to ensure that the requirements for quantification by the comparative threshold cycle (C_t) method were met [151]. RNA purification, cDNA synthesis, and qPCR were performed as previously validated and described in detail [151]. Briefly, frozen tissue was finely ground in liquid nitrogen and RNA was purified with Trizol reagent according to the manufacturer's instructions, and reverse transcribed using the Verso cDNA synthesis kit (Thermo Fisher

Scientific, Waltham, MA). QPCR was performed with Power SYBR Green PCR Master Mix (Thermo Fisher Scientific) in accordance with manufacturer's instructions using a C1000 Touch™ Thermal Cycler (Bio-Rad, Hercules, CA). The specified primer pairs are shown in Table 1. Transcript levels were measured by the comparative C_t method using beta-2 microglobulin (B2M) as an internal control.

Chromatin immunoprecipitation assay (CHIP).

ChIP assays and analysis of fractions by qPCR were performed as previously validated and described in detail [151]. Primer sequences are shown in Table 2. Frozen tissue samples were finely ground in liquid nitrogen prior to the fixation and lysis procedure. Antibodies against the following epitopes were used: normal rabbit IgG control (Millipore, Burlington, MA, #12-370), panacetyl-H3 (Millipore #06599MI), trimethyl-H3Lys4 (Millipore #0473MI), RNA polymerase II CTD repeat (phospho-Ser5, Abcam, Cambridge, MA, #ab5131), RNA polymerase II CTD repeat (phospho-Ser2, Abcam #ab5095) and PRDM16 (Sigma Aldrich #SAB3500989). Bound:input ratios were determined by the comparative C_t value method as Bound/Input=$2^{(\text{input Ct-bound Ct})}$ using a 1% v/v aliquot of input chromatin.

Western blot of CPT1B.

Finely ground frozen tissue was used to prepare whole cell lysates for resolution by SDS-PAGE, transfer to PVDF membrane, and immunodetection as previously described in detail [152]. The anti-CPT1B rabbit polyclonal antibody (Thermo Fisher Scientific #PA5-79065) was used at a 1:500 dilution. Beta tubulin was detected as a loading control using a previously described monoclonal antibody [152].

Phylogenetic analysis of *chkb-cpt1b* loci.

The intergenic distances between *chkb* and *cpt1b* genes in the indicated species were determined using the publicly available Genome Browser [131] at the University of California Santa Cruz Genomics Institute (http://genome.ucsc.edu) and the associated annotation tracks [132]. *Chkb-cpt1b* intergenic distances were calculated from the endpoints of the most proximate mapped reference transcripts for the genes. In species for which the *chkb* and *cpt1b* transcripts have not been directly mapped, the endpoints of the alignments with the orthologous mouse transcripts were used, which was corroborated by gene prediction algorithms. The set of UCP1^{+} eutherians analyzed in Figure 8c consists of human (*Homo sapiens)*, tarsier (*Tarsus syrichta*), tree shrew (*Tupaia belangeri*), mouse (*Mus musculus*), squirrel (*Spermophilius tridecemlineatus*), pika (*Ochotona princeps*), cat (*Felis catus*), dog (*Canis lupus*), microbat (*Myotis lucifugus*), and tenrec (*Echinops telfairi*). The set of UCP1^{-} eutherians consists of pig (*Sus scrofa*), rock hyrax (*Procavia capensis*), elephant (*Loxodonta africana*), horse (*Equus caballus*), cow (*Bos taurus*), dolphin (*Tursiops truncatus*), minke whale (*Balaenoptera acutorostrata scammoni*), manatee *(Trichechus manatus latirostris*), Chinese pangolin (*Manis pentadactyla*), and armadillo (*Dasypus novemcinctus*) [141].

Statistical analysis. Statistical significance of observed differences was determined by ANOVA and a *post hoc* Tukey-Kramer multiple comparison test to identify highly significant differences (α=0.05).

Table 1. Primer sequences for mRNA quantification by qRT-PCR.

Transcript	Forward Primer	Reverse Primer
Mouse		
Chkb	TCCCTGAGATGAACCTGCTG	GGCGATGGGGTAGACTCTAG
Cpt1b	CAACTCCTGGAAGAAACGCC	TCCACCTTGCAGTAGTTGGA
Chkb-Cpt1b	CCCCACCTCAAGTTTCTCCT	TACTGCCTGGTGTGCTTCC
Cpt2	ACTAAGAGATGCTCCGAGGC	AACAAGTGTCGGTCAAAGCC
Slc25a20	CTCTGGGACCTTGGACTGTT	AAACCCAAAGAAGCACACGG
B2m	CCCTGGTCTTTCTGGTGCTT	CGTAGCAGTTCAGTATGTTCG
Opossum		
chkb	GCGCCTTTGGATAAGCTGAT	TGGTTCCCAATGTCAAAGCC
cpt1b	GGGGTAGACTTCCAGCTCAG	CATGACAACGACCAGCCAAC
cpt2	AGATACCTCAGTGCCCAGAG	CTTTCAAAGGACCGGCAGAG
slc25a20	ACTCTGGGACGTTTGACTGT	CCCCAATCCAAAGCCAAAGA
b2m	TTGTGCTTCCTTCCCTACCT	TGGAATCCTGACACGTAGCA

Table 2. Primer sequences for ChIP assay fraction analysis by qPCR.

Amplicon	Primer Sequences	
m1	CCCACCTGCACGTTTATCTG	CACTGAGAGCTGGAGAGAGG
m2	CTCTTCAGCCCCATCCCTAG	GAAAGGGCCTCAAAGCTACG
m3	TAAACCTAACGGAGCCCACG	CATCCTGCAAACCGTCCTTG
m4	AGGTGCTGCTACGACTCTAC	TTGCTCCAGATGTCCGAACT
m5	ATTCCAATCTCCCCAGTCCC	CCTTCTGAACCTCTGCCAGA
m6	CTCCTGGTGACCTTTTCCCT	CACAAGGTTGCTGGAAGGTC
m7	GGCACACTAAGCACCTTCTG	CTCAGAGCCTCCCGACTAAG
m8	CTCAAGTCATGGTGGGCAAC	TCTTCCCACCAGTCACTCAC
m9	GAAGAGAGCCCATCCCTTGT	TGTCTAAGCCAGTCCCTGTG
m10	CGATAGGGCTCTGTGGGATT	TCTAACTTGCCCCTCCTTGG
m11	AAAGAGATACCCCAGCTGCC	CACTGCCTCAAGAGCTGTTC
m12	GATGCCTGTAGTGTGCAAGG	TCCCTTACTGAAGCCCTTGG
o1	ACCCATTCACCTCTCCCAAG	TGTGCTCAACCCTCTTCTCC
o2	GCCTCAGCCAAGAAAACCCT	TGGAAAGATGCCGAGGACTG
o3	ACGTGACCTATGAGCAGACG	CAGTCAGATACCCTCCGCAG
o4	CCGGCTTTATGGGGTCTTCC	ACTAAGGTGGGAGATTGGGG
o5	AAACAGGTAAAGGCCGAGGA	CCAGAGGCCCCAGAATATGT
o6	CACACCCAGATGTCTCCCCA	GTGTGAAGGTGGAGGGTACA
o7	GAAGGGTCTGAGGGTCTGTG	CAGCCAGCTTCAGGACAAAG
o8	GGACCTTTCTGCCCTCATCTA	CAGGAGCAGAAAGGAGCACT
o9	GTTGGCTGGTCGTTGTCATG	ATGCAGCAGATCATCCCCAT
o10	ATGTCATGGTAAGAGCCGGG	CGTCCCTGACTGTGACTTCT
o11	CGGAATGGATTTGGAGCTGG	CCTTGGATTTCAGCCTGCTC
o12	AGATCTTCCCTCCACCCTCT	CTGCTTTCTCGGTGGGGATA
o13	CCTCCTTCCTCTCAATGCCA	AGCTTGGCCCTAAACTCACT
h1	CCCCAGCCAGAATCAGTTTT	CTTGAACCCAGGCCTTGTTG
h2	GGAATGAATGAAGACGCGCA	TGCTGAAGGAATGTGACCCT
h3	GCGATGCGGGTAGTATTGTT	AGGCTTCGGTGTAATGAGGT
h4	TCCTGACTCTTGGCCATTGA	CCTTCCTCACCGACTGACTT
h5	ACATCGGTGACCTTTTCCCT	GGGTTTCTGTGCGGTGAAG
h6	GAGGCCCTGAAACACGTCTA	CTCTGGGAGAAGCACCTGTG
h7	CTCAAGTCATGGTGGGCAAG	GGGGAGGTGGAAGGTTAGAG
h8	TTCCCTGCTTCTGACACTGT	CTCTCACGCAACACCAACAA
h9	AAGTGTCCTCTGCCCATGTG	GCGGGGCAGAAGTAAAGG
h10	TTTTCCAAGTTCCCAAGGCC	GTGTCCTTGGAGCCTGGG
h11	GTGTCTGTTGAACACCCACC	CGACATGGGCAGGAATAGGA

CHAPTER III: DELETION OF THE CHKB PROMOTER AND CPG ISLAND IN SKELETAL MUSCLE AND BROWN FAT CELLS LEADS TO A SIGNIFICANT LOSS OF CPT1B TRANSCRIPTS

Introduction

Phylogenetic analysis of the *CHKB-CPT1B* gene locus throughout vertebrate evolution indicates that the genomic convergence of these two genes into close proximity coincides with the emergence of mammals with higher basal metabolic rate and body temperature. CPT1B is a crucial enzyme catalyzing the rate limiting step in the pathway of mitochondrial LCFA oxidation and is thus an ideal candidate for a transcriptional regulatory path for the evolution of BAT. The hypothesis at this point was that the reduction of intergenic sequence throughout vertebrate evolution resulted in the co-option of *CPT1B* regulation and transcription by the promoter and flanking region of *CHKB* in eutherian mammals. As shown in Figure 5 and 6, chromatin immunoprecipitation assays show that histone 3 acetylation and trimethylation are largely localized at the 5' region of *Chkb* in murine DNA. These chromatin modifications are marks of transcriptional activation; the addition of an acetyl group on the lysine residues of histone tails works to neutralize their positive charges thereby weakening the interaction between the histone proteins and the negatively charged DNA. H3K4Me3 is a strong mark of transcriptional activation that correlates closely with the binding of transcriptional machinery. The unraveling of the DNA from the tightly joined nucleosome structures allows for increased accessibility to transcription factors and transcriptional machinery for gene expression, such that these modifications are generally found at promoter regions of active genes.

Aoyama *et al.* were one of the first groups to characterize potential transcription factor binding sites within the putative promoter region of *Chkb* in mice and showed the presence of two CCAAT boxes (in inverse orientation), one SP1 binding site and one CREB binding site all within the proximal promoter region situated around 55 bp upstream of the TSS [16, 17]. A

functional TATA box was not present in the proximal promoter region immediately upstream of the TSS. Instead, a TATA-like sequence was found 135 bp upstream of the TSS that seems to have no TBP binding and therefore no influence on transcription of the gene. Additionally, the region within -1 to -500 (relative to *Chkb* TSS) showed to be heavily saturated with predicted binding elements for MYOD, SRY, GATA-1/2/3, HSF2, C/EBPb, CRE-BP, and OCT-1 transcription factors. *In silico* analysis of the 2000 bp *CHKB* promoter region in humans also showed a multitude of possible binding sites for SP1, Ets, GATA factors, SREBP, MZF1, NF-κB, AP-1, and E2 transcription factors [112]. Considering that much of the *CHKB* genomic sequence is conserved between humans and other mammals such as mice, we were confident in using this information to direct our Cas9 directed mutagenesis in murine C2C12 myoblasts and DE 2.3 brown fat cells.

Interestingly, we discovered that a 475bp deletion in the promoter region upstream of the *Chkb* TSS site showed no significant effects on *Chkb* or *Cpt1b* transcript levels in C2C12 skeletal muscle cells. Instead, the inclusion of the *Chkb* CpG island along with the upstream promoter deletion was necessary to detect significant reduction in not only *Chkb* transcripts but *Cpt1b* transcripts as well in both C2C12 skeletal muscle and DE 2.3 brown fat cells compared to wildtype (WT).

Results and Discussion

Isolation of a C2C12:*Chkb*Δ(promoter) deletion clone.

Using a multiplexed CRISPR Cas9 approach, adapted from a published 2013 protocol by Ran *et al.* [153], we initially planned to introduce two simultaneous double stranded breaks that targeted a 475bp region within the proximal promoter of *Chkb* to be deleted. The region was chosen as it was shown to be heavily saturated with a number of potential transcription factor binding sites [16, 17]. Importantly, the selected deletion also encompassed both CCAAT boxes (inverse orientation) and SP1 binding site which represent cis promoter elements with strong influence on RNA transcription.

Deletion of the previously defined *Chkb* promoter does not affect *Chkb* or *Cpt1b* transcript levels.

The genotyping of the isolated clones showed a 475 bp *Chkb*ΔPromoter deletion indicated by the presence of a single deletion band compared to the single WT band after genomic PCR and gel electrophoresis (Figure 12). Interestingly, the C2C12 cells with the *Chkb* upstream promoter deletion showed no significant differences in transcript levels of *Chkb* or *Cpt1b* compared to wildtype cells (Figure 13 A-B). We also did not observe a significant difference in readthrough transcript levels between WT and mutant cells (Figure 13 C). This observation may be explained by a dominant activity of sequences within the CpG island that spans from upstream of the TSS to 5' half of the first intron of the *Chkb* gene in the initiation of transcription. We already indicated that the highest peaks of both acetylation of histone 3 and trimethylation of histone 3 at lysine 4 are located at the *Chkb* CpG island (ChIP primer m3) indicating this region can act as a platform with ample accessibility for the transcription initiation complex to assemble compared to any other region within the *Chkb-Cpt1b* locus. Accordingly,

ChIP assays using the total histone 3 antibody showed nucleosome occupancy to be the lowest at the CpG island (data not shown) which suggests that the DNA sequence in this region has less affinity for histone binding and may allude to the inability of CpG islands to form stable nucleosomes. This suggests that *Chkb* seems to function without a functional TATA-box and may instead rely on SP1-mediated transcriptional activation. SP1 binding sites are plentiful within CpG islands given the high propensity for GC rich sequences. SP1 is a very well characterized mammalian transcription factor that belongs to the 26-member rich Sp/KLF family of ubiquitously expressed transcription factors. It functions as a transcriptional activator of a range of genes including housekeeping genes, cell cycle regulators, and tissue-specific genes. SP1 therefore serves as a constitutive activator of housekeeping genes, by binding to the GC boxes within the putative CpG islands and recruit TFIID via a tethering factor that is physically associated with TATA binding proteins. Therefore, in this mechanism, the TFIID complex is recruited to the promoter via protein-protein interactions between its tethering factor and the DNA-bound Sp1, instead of a direct protein-DNA interaction [154]. There have been chromatin looping mechanisms postulated to deliver the RNA polymerase complex to the upstream of the TSS site to begin transcription, however assessment of the exact mechanism within our gene locus would require future chromatin conformation capture (3C) assays. Additionally, SP1 sites may also be prevalent in the region of DNA that comes in contact with the TSS after deletion of the proximal promoter sequence along with RNA pol II occupancy which could also explain why we don't see a significant drop of *Chkb* transcript levels with this deletion. Alternatively, we've consistently observed a strong H3K4me3 signal at the *Chkb* CpG island which has been demonstrated to recruit chromatin remodeling proteins that result in an open accessible chromatin for transcription machinery, however there is also evidence for a direct relationship between RNA Pol and H3K4me3. [155-157]. Furthermore, the binding of TFIID has also been shown to precede TATA-binding protein (TBP) interaction with DNA [158]. Additionally, Acetylation of histones has also been demonstrated to increase binding of TFIID to H3K4me3

chromatin regions [156, 159]. ChIP assays using antibodies against TBP were unsuccessful in showing any strong interactions with the *Chkb-Cpt1b* gene locus (data not shown).

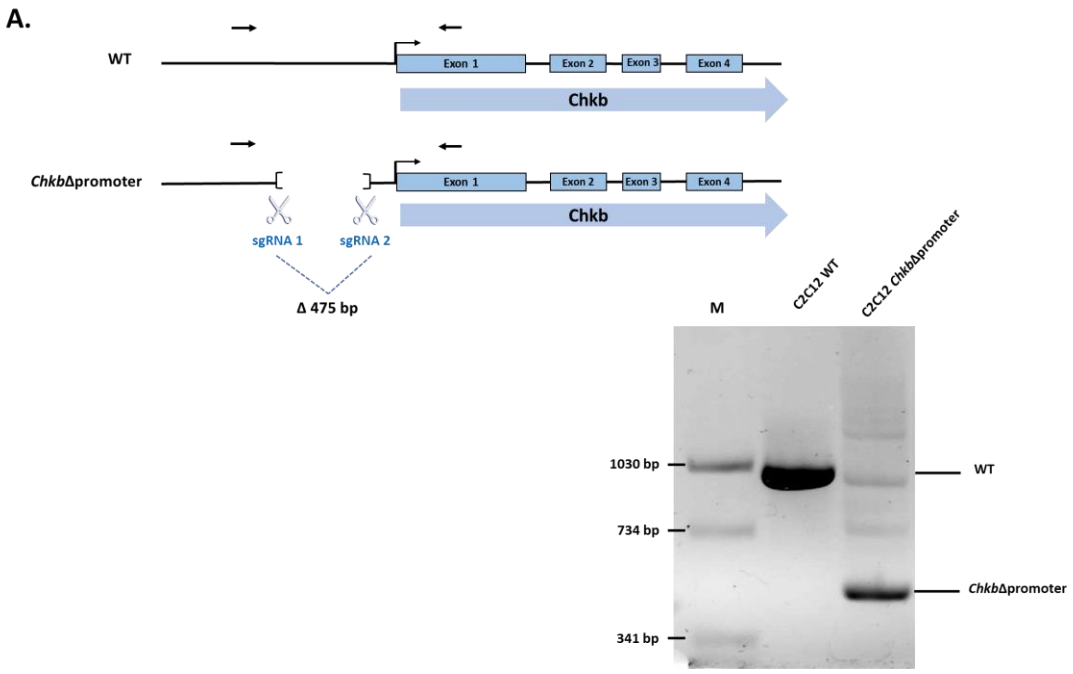

Figure 12. C2C12 *Chkb*ΔPromoter mutant locus. A) The *Chkb*ΔPromoter deletion (475 bp) removes the promotor region upstream of the *Chkb* TSS. sgRNA sequences are listed in table 3. Agarose gel shows genomic PCR at the *Chkb* locus from wild type and deletion mutant. Primers used for PCR are shown as black arrows for reference. The expected WT amplicon is 979 bp while expected deletion amplicon band is 504 bp. Black arrows represent forward and reverse primers for endpoint PCR and are labeled “E1” in table 5.

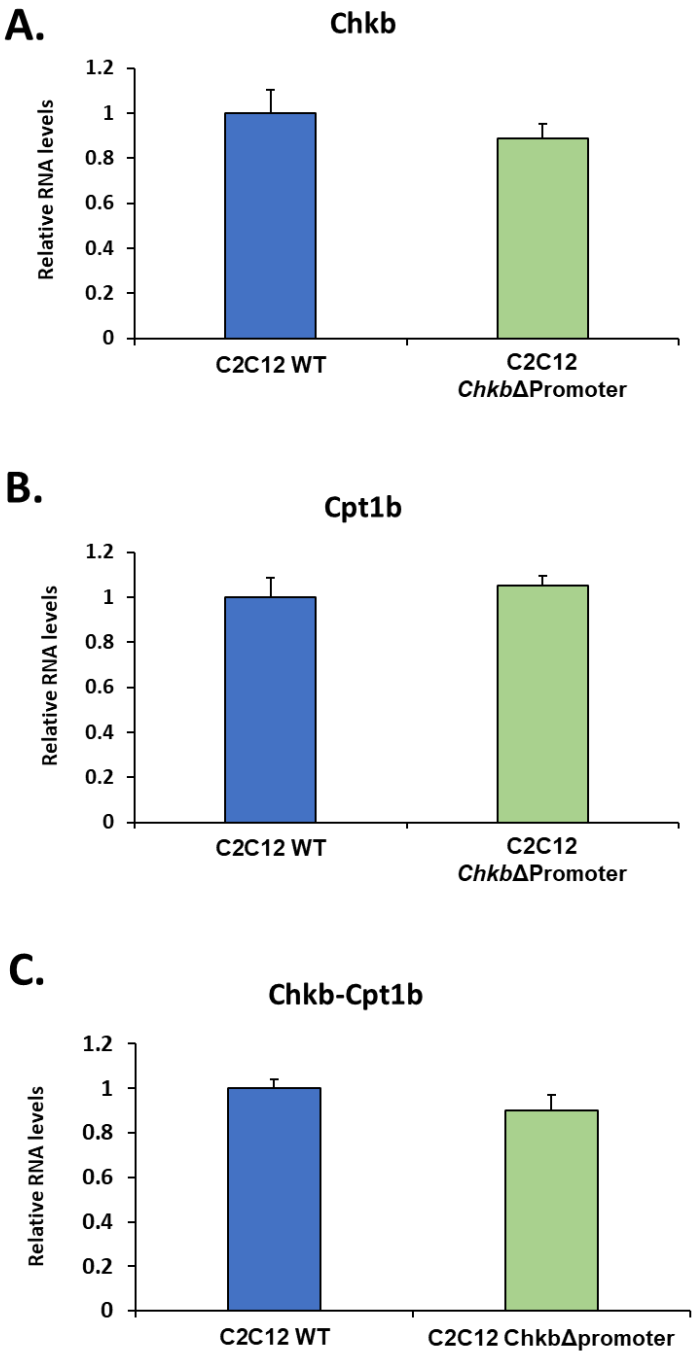

Figure 13. Transcript levels of *Chkb*, *Cpt1b* and *Chkb-Cpt1b* are similar in C2C12 *Chkb*ΔPromoter deletion mutants compared to WT. (A) *Chkb* transcripts are slightly lower in

C2C12 mutants with proximal promoter deletion compared to WT cells. (B) *Cpt1b* transcripts show no difference in quantity between mutants and WT cells. (C) *Chkb-Cpt1b* readthrough transcripts show no significant difference in quantity between mutants and WT cells. Internally controlled mRNA levels (vs. *β2-microglobulin* mRNA) were normalized to the WT controls. The data indicates the mean ± *SE* (*n*=3). Student's t-test determined no statistical significance between wild-type and mutant cells ($P > 0.05$).

The lack of effect of a deletion of the putative *Chkb* 5'-flanking promoter suggests that the sequences involved in RNAPII recruitment and transcription initiation lie elsewhere. One possibility is the presence of an intragenic promoter, which has been a longstanding observation at specific loci in multiple species. There are several published phenomena known as intron-mediated enhancement (IME) occurring in a wide array of eukaryotes, including vertebrates, invertebrates, fungi, and plants [160-166]. Similarly, an exon mediated activation of transcription has also been found in mammalian cells [167, 168]. These types of introns and exons are transcriptional enhancers that can directly influence transcription initiation. One of the most widely studied IME introns is the *UBQ10* intron 1 that has been shown to maintain transcription of *TRP1* gene when inserted within 1 kb downstream of the transcription start site of *TRP1*. The most surprising finding of this study was that even after the deletion of the *TRP1* proximal promoter, transcript levels remained unchanged when the *UBQ10* intron was present within the 5' region of the gene, indicating that this type of intron is capable of recruiting and guiding transcriptional machinery to the TSS of the gene [169]. IMEs have been well studied mainly in *Arabidopsis thaliana* and *Oryza Sativa* and are identified using a bioinformatics approach implemented by Rose *et al.*, however we cannot rule out the possibility that such a transcriptionally regulatory intron or exon could be present within the 5' of the *CHKB* gene especially when we consider the existence of the CpG island that encompasses the first exon and part of the first intron.

Inclusion of the *Chkb* CpG island (CGI) and 5'-flanking promoter in a new deletion.

In light of the lack of effect of the 5'-flanking promoter deletion of *Chkb* and *Cpt1b* transcript levels, we expanded the deletion to encompass not only the proximal promoter region but also the 0.7 kb CpG Island which extends from -86 bp upstream of the TSS to +589 bp downstream. To target this deletion, the 5' sgRNA-directed cut remained the same, while a new

3’ sgRNA was designed to induce a double stranded break 776 bp downstream of the TSS located within the second exon. The new 1296 bp deletion would also encompass the *Chkb* translation start codon, so we expected *Chkb* activity to be completely knocked out. Additionally, we previously observed a low cell viability in both C2C12 and DE2.3 cells after FACS sorting in the prior transfections signaling that this process for monoclonal isolation may be too harsh for these cells. Therefore, considering that the pSpCas9 (BB) vectors contained a puromycin resistance gene, we decided to use puromycin selection from this point onward for the full *Chkb* Δ(promoter-CGI) deletions.

Isolation of C2C12*:Chkb*Δ(promoter-CGI) deletion clones

Using genomic qPCR screening with the qPCR primers listed in table 4, we were able to identify two clones of interest in the C2C12 cell line and one clone of interest in the DE 2.3 brown fat cell line. We utilized the comparative C_t method and anti-log transformation to visualize loss of sequence based on amplicon enrichment levels relative to the internal control amplicon g11. Amplicons are mapped for the WT and mutant C2C12 locus in figure 14A. QPCR analysis of genomic DNA from C2C12 clone 1 showed a complete deletion of the intended target indicated by low levels of amplicons g3, g4, g5, and g6 relative to g11 as opposed to nearly equal levels of enrichment for amplicons outside of the intended deletion which are similar to WT (Figure 14B). qPCR analysis of genomic DNA from C2C12 clone 2 however, shows either a partial deletion or a heterozygous mixture consisting of a full deletion in one allele and a partial deletion in the second allele. Because we see intermediate levels of g3 and g4 amplicons and a complete loss of amplicons g5 and g6 relative to g11 compared to WT we propose that one allele of this clone received the full intended deletion of the promoter and CpG island while the other allele still maintains some sequence near the 5’ cut site indicated by ~40% of g3 and g4 amplicons relative to g11 compared to WT, and a complete loss of g5 and

g6 amplicons relative to g11 compared to WT (Figure 14C). Therefore, C2C12 clone 2 has complete loss of CpG island in both alleles but still contains a partial upstream promoter sequence.

Gel electrophoresis of endpoint PCR products from C2C12 *Chkb*Δpromoter:CGI clones 1 and 2 confirm our qPCR findings. Gel electrophoresis of genomic PCR products using primers flanking the deletion (primer pair E2) shows a single band in clone 1 approximately 412 bp in size indicative of a full deletion and the presence of 4 separate bands in clone 2 indicative of differential cutting (Figure 15B). An additional genomic PCR using a primer pair (primer pair E3) encompassed within the deletion confirmed the deletion as we observe a single band for WT controls and no bands present for either deletion clone (Figure 15C). To confirm the exact size and sequence of the amplicons for both C2C12 deletion clones, the bands were gel extracted and submitted for next gen sequencing (NGS) which confirmed that C2C12 clone 1 had a 1296 bp deletion and the resulting band was 412 bp in size with the intended sequence (Figure 16A). Only bands 1,3 and 4 were able to be resolved using NGS in C2C12 clone 2. Band 1 was 768 bp and the 940 bp deletion initiated 356 downstream of the intended 5' sgRNA 1 cut site while the 3' cut site remained as expected for sgRNA 3 (Figure 16B). Band 3 had a 1296 bp deletion and the resulting band was 412 bp in size with the intended sequence (Figure 16B). Band 4 was 379 bp in size and seemed to initiate 33 bp upstream of the 5' sgRNA 1 cut site indicating a 1329 bp deletion (Figure 16B). Bands 2 and 4 could just be artifacts of genomic PCR due to nonspecific primer binding and we believe that bands 1 and 3 are the true amplicons representative of the differential cutting sites between the two alleles in C2C12 clone 2.

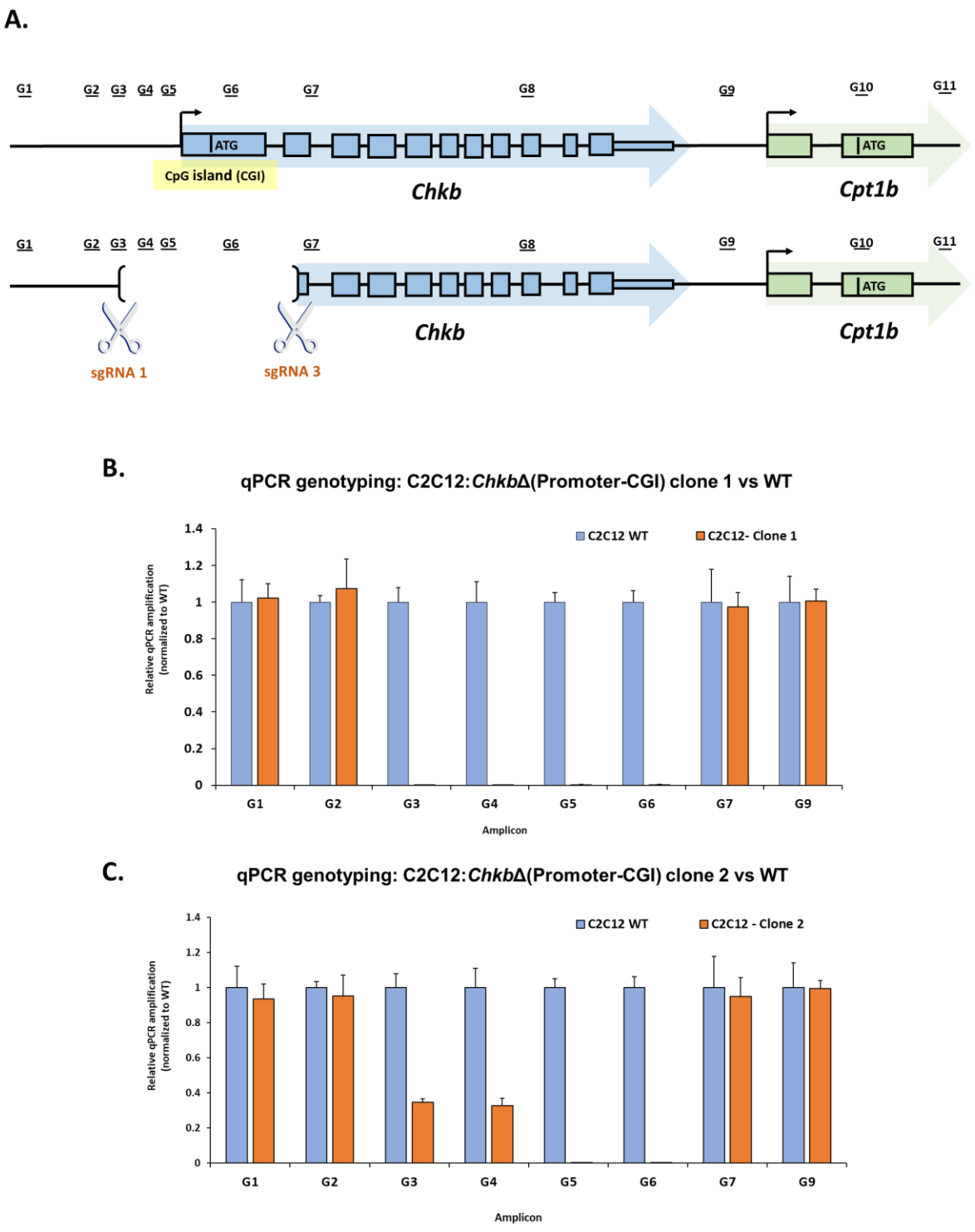

Figure 14. qPCR genotyping of C2C12 *Chkb*Δ(Promoter-CGI) mutant Clones 1 & 2 using qPCR primers. (A) Amplicons, labeled G1-10, are mapped throughout the *Chkb-Cpt1b* locus.

Primer sequences for G1-G10 are listed in table 4. (B) Genomic DNA qPCR results of C2C12 deletion clone 1 using primers that amplified 100-150bp segments. (C) Genomic DNA qPCR results of C2C12 deletion clone 2 using primers that amplified 100-150bp segments. Delta- Ct values of G1-10 were acquired using a positive control qPCR primer pair (g11) within the *Cpt1b* gene. The exponential qPCR data (C_t values) was antilog transformed to a linear expression $=2^{(g11\ Ct-\ (g1-g10)\ Ct)}$ and normalized to WT.

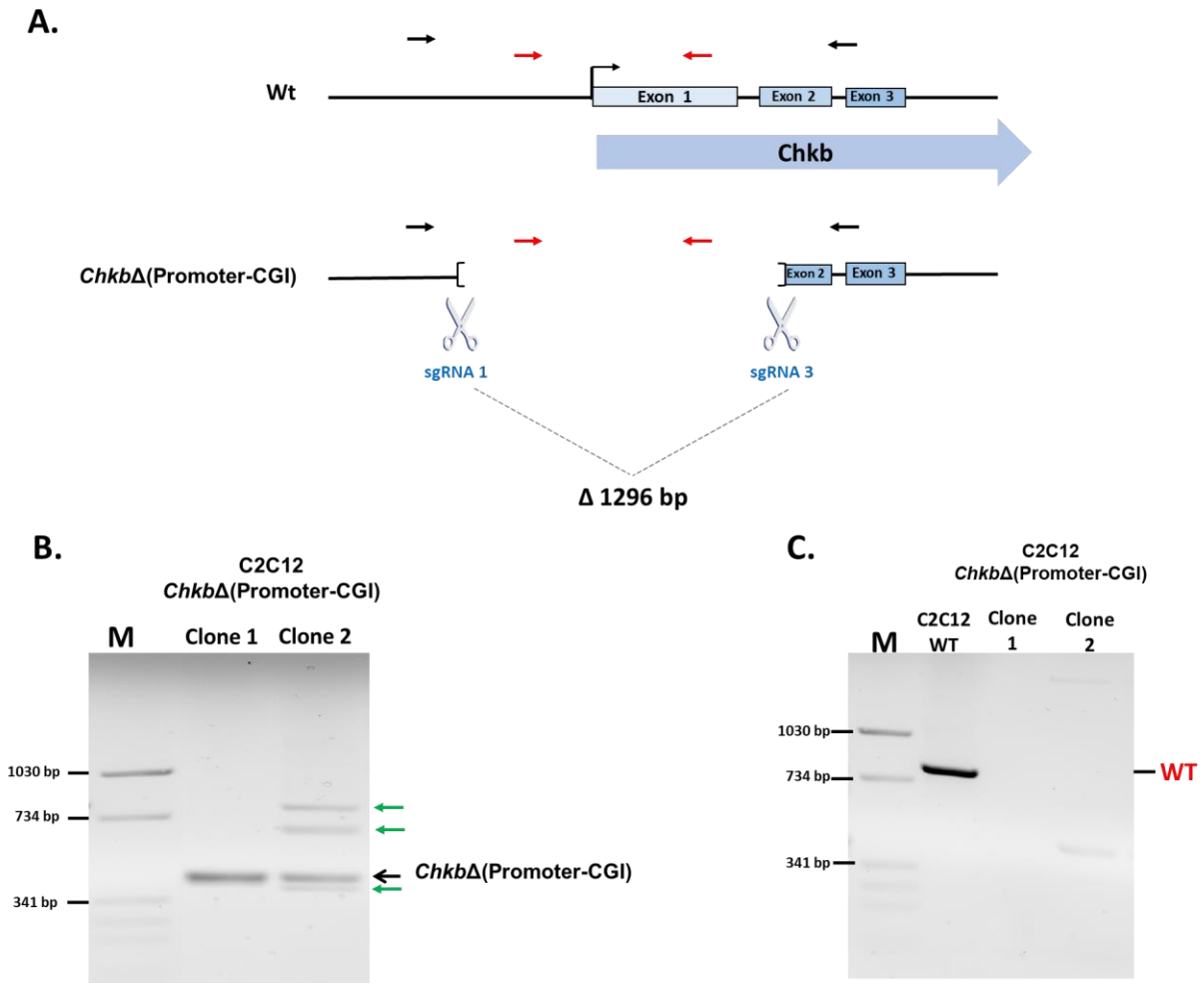

Figure 15. C2C12 *Chkb*Δ(Promoter-CGI) mutant locus. (A) The *Chkb* Δ(Promoter-CGI) deletion (Δ 1296 bp) removes the promotor region of *Chkb* along with the entirety of exon 1, intron 1, and the 5' of exon 2 in C2C12 mutants. sgRNA sequences are listed in table 3. (B) Agarose gel shows genomic PCR at the *Chkb* locus from C2C12 *Chkb* Δ(Promoter-CGI) deletion clones 1 and 2. Primers used for endpoint PCR in are shown as black arrows for reference. Expected length of Wt amplicons is 1708 bp and expected length of deletion amplicon band is 412 bp. Clone 1 shows a single band of the expected size, clone 2 shows multiple bands (green arrows) including the expected band (black arrow). (C) Agarose gel shows genomic PCR at the *Chkb* locus from C2C12 WT and *Chkb* Δ(Promoter-CGI) deletion clones 1 and 2. Primers used for endpoint PCR in are shown as red arrows for reference. Expected length of Wt amplicons is 805 bp while no amplicons are expected in deletion

mutants. The marker is DPNI-digested pBlueScript(KS-) plasmid fragments. Black endpoint PCR primers are labeled “E2” in table 5. Red endpoint PCR primers are labeled “E3” in table 5.

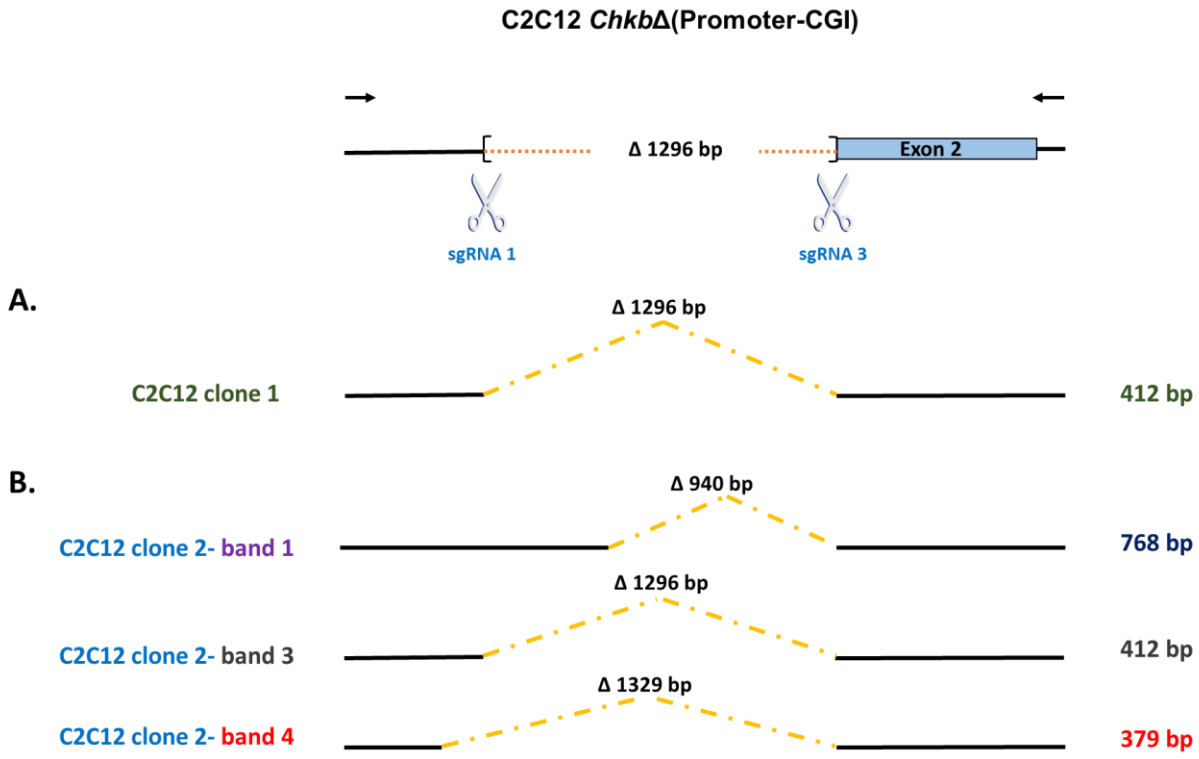

Figure 16. Sequencing results of C2C12 *Chkb*Δ(Promoter-CGI) deletion clones 1 and 2 amplicons. Next gen sequencing of the gel extracted genomic PCR products from figure 15. (A) C2C12 clone 1 amplicon was confirmed to be cut exactly at the predicted sgRNA cut sites resulting in a 414 bp amplicon with a 1296 bp deletion. (B) C2C12 clone 2 genomic PCR showed 4 bands in figure 15, however only the 1st , 3rd, and 4th bands were resolved during sequencing. Sizes and exact cut sites are shown for the three C2C12 clone 2 amplicons.

Genotyping indicates the presence of a DE 2.3 mutant with a *Chkb*Δ(Promoter-CGI) deletion

Genomic qPCR screening amplicons are mapped across the WT and mutant DE 2.3 locus in figure 17A. Analysis of DE2.3 *Chkb*Δ(Promoter-CGI) deletion mutant indicates a larger deletion than anticipated as we see low levels of amplicons g2 and g7 along with g3, g4, g5, g6 relative to g11 compared to WT (Figure 17B). As g2 and g7 are just upstream and downstream of the proposed deletion, respectively. We theorize that this was a result of non-homologous end joining (NHEJ) repair process that removed an additional amount of DNA sequence in the DE 2.3 mutant.

To account for the additional DNA sequence loss in DE 2.3 *Chkb*Δ(promoter-CGI) mutant we used a different set of forward and reverse flanking primers (primer pair E4 in table 5) for endpoint PCR that would amplify 2587 bp in WT cells (Figure 18A). Gel electrophoresis indicated a single amplicon band around slightly above the 1030 bp marker band, possibly a 1050 bp amplicon (Figure 18B). We can therefore hypothesize that these mutants had a deletion around 1540 bp instead of the anticipated 1296 bp deletion. DNA repair is not an exact process; therefore, we theorize that the NHEJ repair process resulted in the removal of an additional ~ 241 bp of DNA sequence. An additional genomic PCR using a primer pair encompassed within the deletion confirmed the loss of sequence as we observe a single band for WT controls and no bands present for either deletion clone (Figure 18C). To confirm the exact size and sequence of the DE deletion amplicon, the bands were gel-extracted and submitted for next gen sequencing which showed a 1070 bp amplicon indicating that this mutant received a 1517 bp deletion (Figure 19). Based on NGS we were able to determine that an additional 179 bp of sequence was removed upstream of the 5’ sgRNA1 cut site and an additional 42 bp of sequence was removed downstream of the 3’ sgRNA3 cut site.

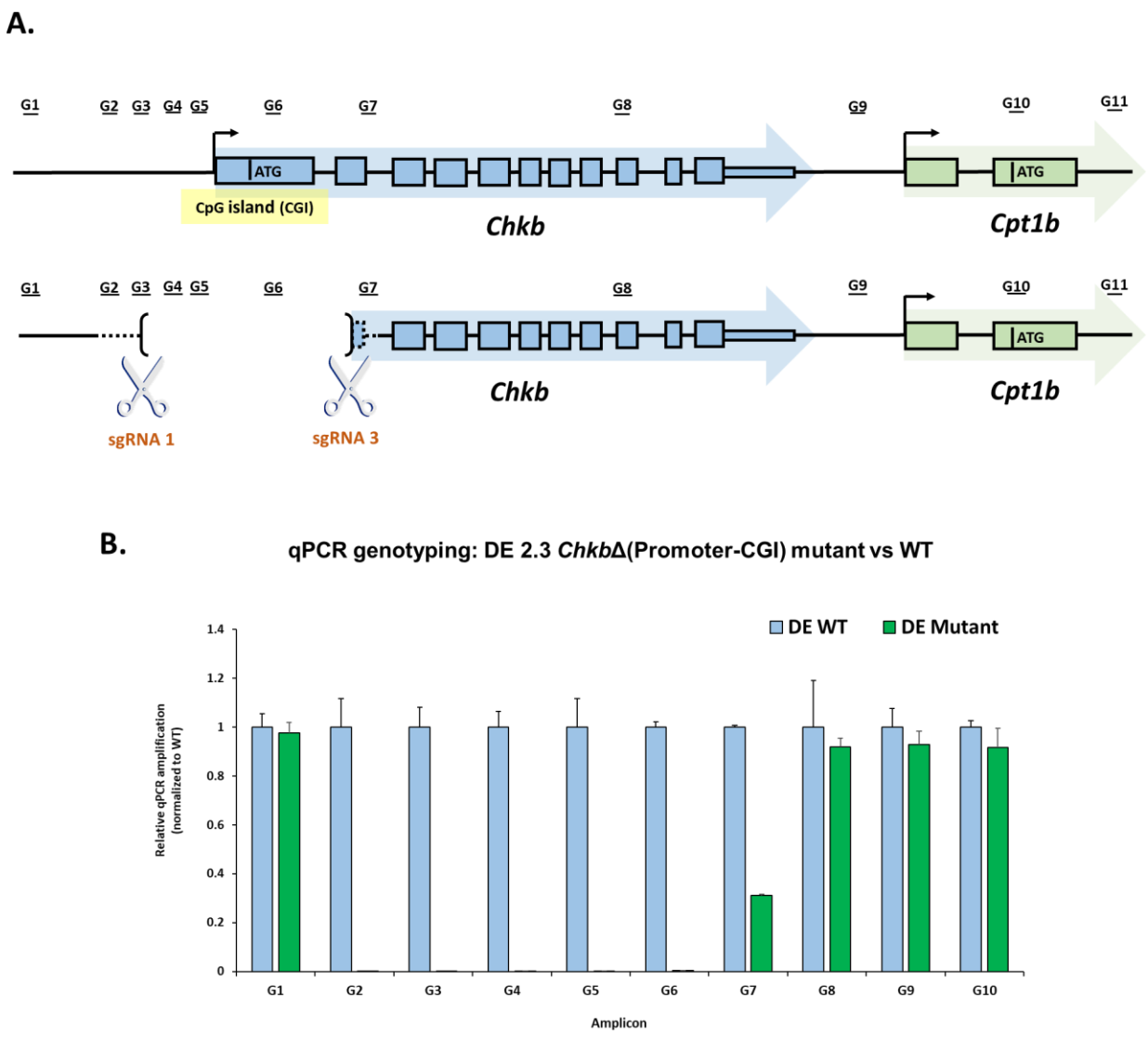

Figure 17. qPCR genotyping of DE2.3 *Chkb*Δ(Promoter-CGI) deletion mutant using qPCR primers. (A) Amplicons, labeled G1-10, are mapped throughout the *Chkb-Cpt1b* locus. sgRNA cute sites are portrayed and dotted lines represent possible additional DNA sequence that may have been deleted by NHEJ repair. Primer sequences for G1-G10 are listed in table 4. (B) Genomic DNA qPCR results using primers that amplified 100-150bp segments. Delta- Ct values of G1-10 were acquired using a positive control qPCR primer pair within the *Cpt1b* gene. The exponential qPCR data (C_t values) was antilog transformed to a linear expression $=2^{(\text{control Ct- g1-g10 Ct})}$ and normalized to WT.

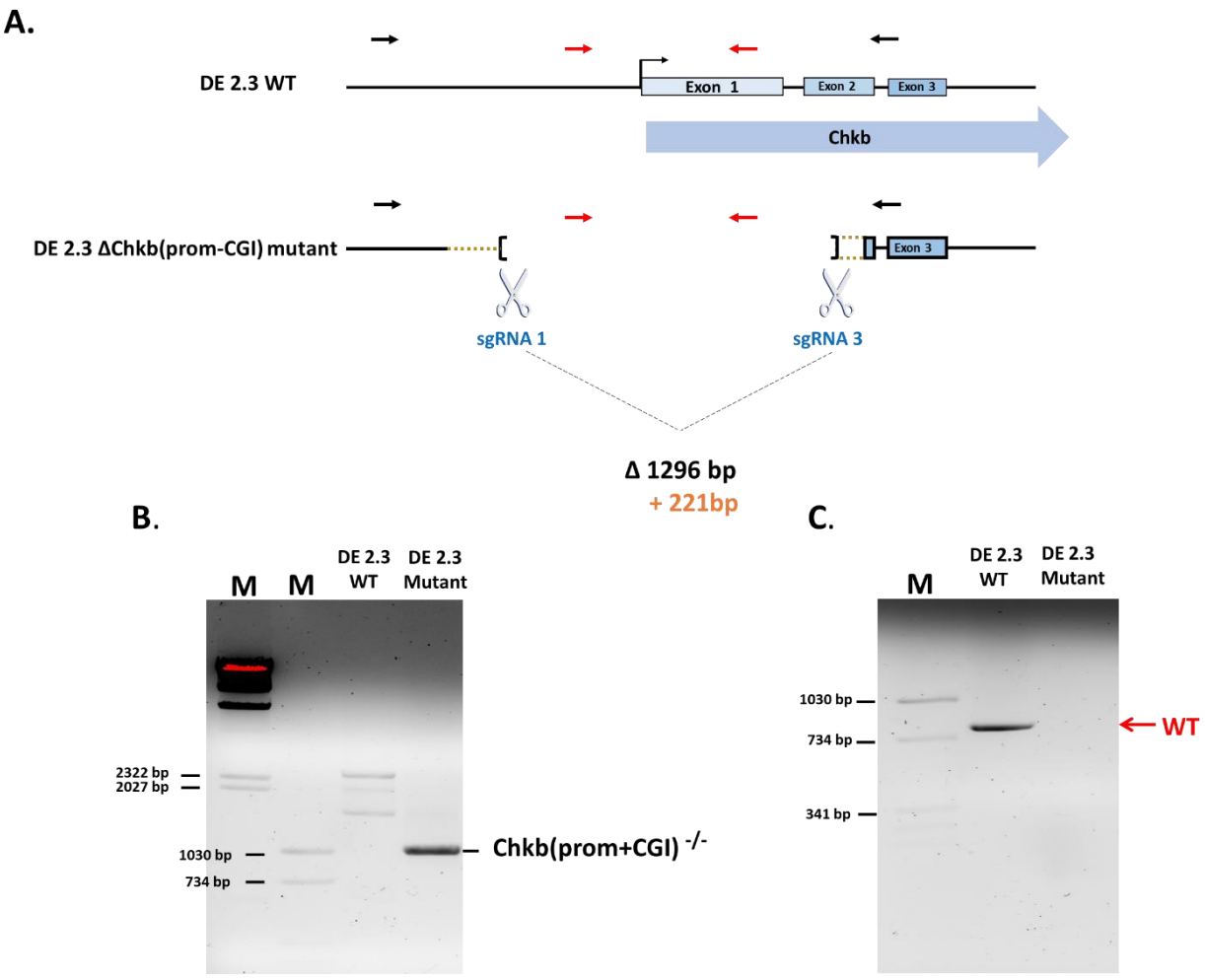

Figure 18. DE2.3 *Chkb*Δ (promoter-CGI) mutant locus. (A) The *Chkb*Δ (promoter-CGI) deletion (1296 bp) removes the promotor region of *Chkb* along with the entirety of exon 1, intron 1, and the 5' of exon 2 in DE2.3 mutants along with an additional 221 bp sequence via NHEJ confirmed via fragment sequencing. sgRNA sequences are listed in table 3. (B) Agarose gel shows genomic PCR at the *Chkb* locus from WT and the DE2.3 *Chkb*Δ (promoter-CGI) deletion mutant. Primers used for endpoint PCR in are shown as black arrows for reference. Expected length of Wt amplicons is 2587 bp and expected length of deletion amplicon band is 1291 bp. A single band is observed for the deletion mutant (indicated by black arrow) however it is shorter

than expected size. (C) Agarose gel shows genomic PCR at the *Chkb* locus from DE WT and *Chkb*Δ (promoter-CGI) deletion mutant. Primers used for endpoint PCR in are shown as red arrows for reference. Expected length of Wt amplicons is 805 bp while no amplicons are expected in deletion mutants. The first marker is HindIII- digested lambda DNA and the second marker DPNI-digested pBlueScript(KS-) plasmid fragments. Black endpoint PCR primers are labeled E4 in table 5. Red endpoint PCR primers are labeled E3 in table 5.

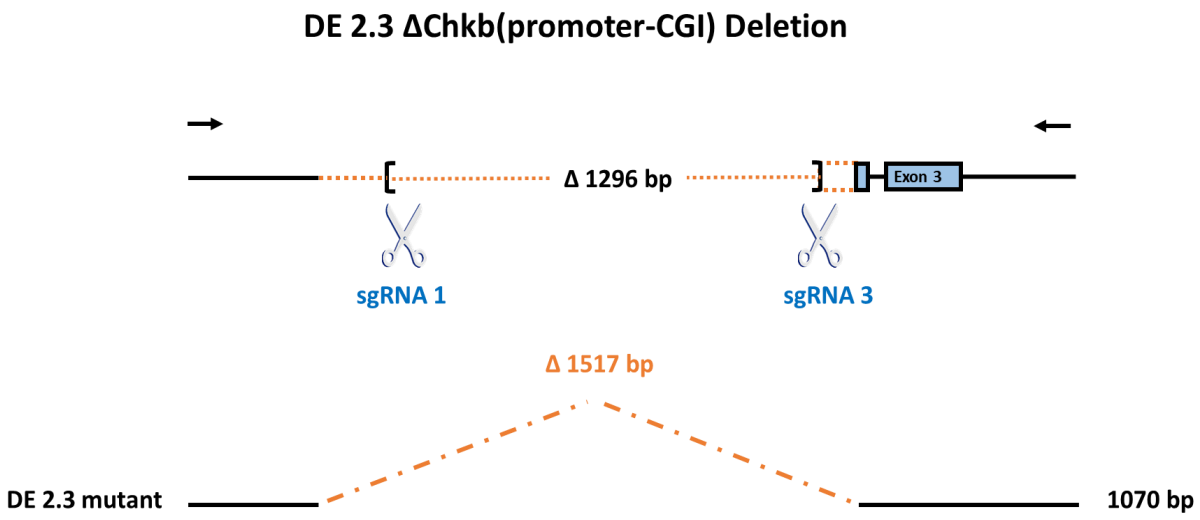

Figure 19. Sequencing result of DE 2.3: *Chkb*Δ(Promoter-CGI) deletion mutant amplicon. Next gen sequencing of the gel extracted DE 2.3 mutant band from figure 18 confirms a deletion of 1517 bp.

***Chkb*, *Cpt1b*, and readthrough transcripts are undetectable in C2C12 *Chkb*Δ(promoter-CGI) mutants.**

QRT-PCR analysis of C2C12 mutants with the *Chkb*Δ(promoter-CGI) deletion, unsurprisingly, showed severe diminishment in *Chkb* transcripts (Figure 20A) that was statistically significant ($P<0.05$). Surprisingly, we also observed a similar degree in diminishment in *Cpt1b* transcripts levels rendering them virtually undetectable ($P<0.05$) (Figure 20B). Transcript levels of the major read through *Chkb-Cpt1b* transcripts also followed the same trend of significant reduction in both mutants compared to WT ($P<0.05$) (Figure 20C). Conversely, we observed minor to significant increases in the transcript levels of other mitochondrial enzymes also involved in the carnitine-acyl-carnitine shuttle system. Transcripts of *Slc25a20*, the carnitine-acylcarnitine translocase, were slightly increased in both clones compared to WT, however, were not statistically significant (Figure 20 D). CPT2, located within the inner mitochondrial membrane, replaces the carnitine with a CoA to complete the transport of LCFA-CoA into the mitochondrial matrix. We found a significant increase in *Cpt2* transcripts in clone 1 compared to WT ($P<0.05$) while there were no significant changes between clone 2 and WT (Figure 20 E). Transcript levels of *Cpt1a*, the predominant liver isoform of the CPT1 family, were increased in clone 1 compared to WT, although not quite statistically significant ($P= 0.0585$) (Figure 20F). *Cpt1a* transcripts were not significantly different between clone 2 and WT ($P=0.2058$) (Figure 20F). in deletion mutants compared to wild type cells (Figure 20 F) indicating an upregulation in expression, potentially to compensate for the loss of *Cpt1b* expression to maintain FAO levels.

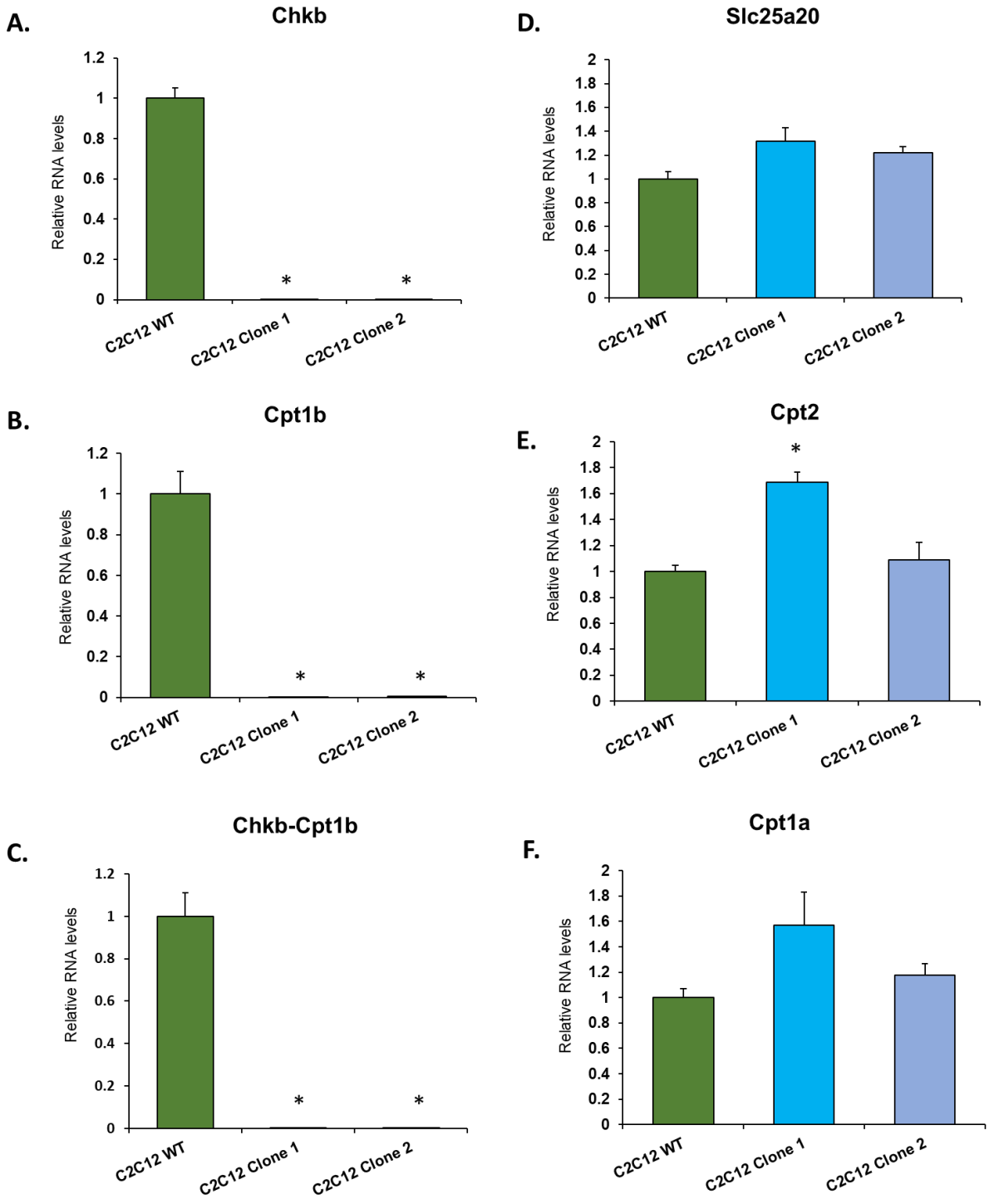

Figure 20. *Chkb*, *Cpt1b,* and readthrough transcripts are undetectable in the both C2C12: *Chkb* Δ(Promoter-CGI) deletion clones compared to WT. (A) *Chkb* transcripts are

undetectable in mutants compared to WT. (B) *Cpt1b* transcripts are undetectable in mutants compared to WT. (C) *Chkb-Cpt1b* readthrough transcripts are undetectable in mutants compared to WT. (D) *Slc25a20* transcripts show a slight increase in deletion mutants compared to WT. (E) *Cpt2* transcripts show a slight increase in deletion mutants compared to WT. (F) *Cpt1a* transcripts are increased in both deletion mutants compared to WT. Internally controlled mRNA levels (vs. *β2-microglobulin* mRNA) were normalized to the C2C12 WT. The data indicate the mean ± *SE* (*n*=3). Student's t-test was used to determine statistical significance between wild-type and mutant cells (*P < 0.05).

Loss of histone acetylation is observed across the entire *Chkb-Cpt1b* locus in C2C12 *Chkb*Δ(promoter-CGI) mutants.

Chromatin immunoprecipitation assays using antibodies to detect histone 3 acetylation (anti-pan Acetyl-H3) show a reduction in acetylation across the entire *Chkb-Cpt1b* gene locus for both C2C12 clones (Figure 21). C2C12 clone 1 shows that with the deletion of the CpG island and promoter there is a drastic reduction in acetylation signal in all regions around the deletion as well (Figure 21A). Interestingly, we do observe a slight increase in acetylation of histones occupying the *Cpt1b* gene locus, represented by amplicons 7-10, in C2C12 clone 1 compared to WT (Figure 21B). However, increase in modification at the *Cpt1b* locus does not amount to any significant independent expression of *Cpt1b* as shown by our qRT-PCR data (Figure 20B). AcH3 ChIP on C2C12 clone 2 shows a relatively higher AcH3 signal at the m1 and m2 amplicon compared to all other sites of the gene locus (Figure 21C). Considering that Clone 2 still possesses one allele containing the part of the upstream *Chkb* promoter where the m2 amplicon is located, we do see a bound/Input ratio >1 (figure 21C). However, due to the loss of the CpG island in both alleles we still notice a drastic reduction in histone acetylation across the entire gene locus which becomes more apparent when we compare the Ach3 Chip assay to WT C2C12 cells (Figure 21D). Therefore, we can confidently say that the chromatin modifications localized to the *Chkb* 5' region, specifically the CpG island, are crucial in maintaining an open/active chromatin domain that encompasses both genes, and the loss of the active site results in an inactive unacetylated domain. Accordingly, ChIP assays using RNA Pol II Ser5P (initiation form) and RNA Pol II Ser2P (elongation form) provide more evidence for a transcriptionally inactive gene locus (Figure 22).

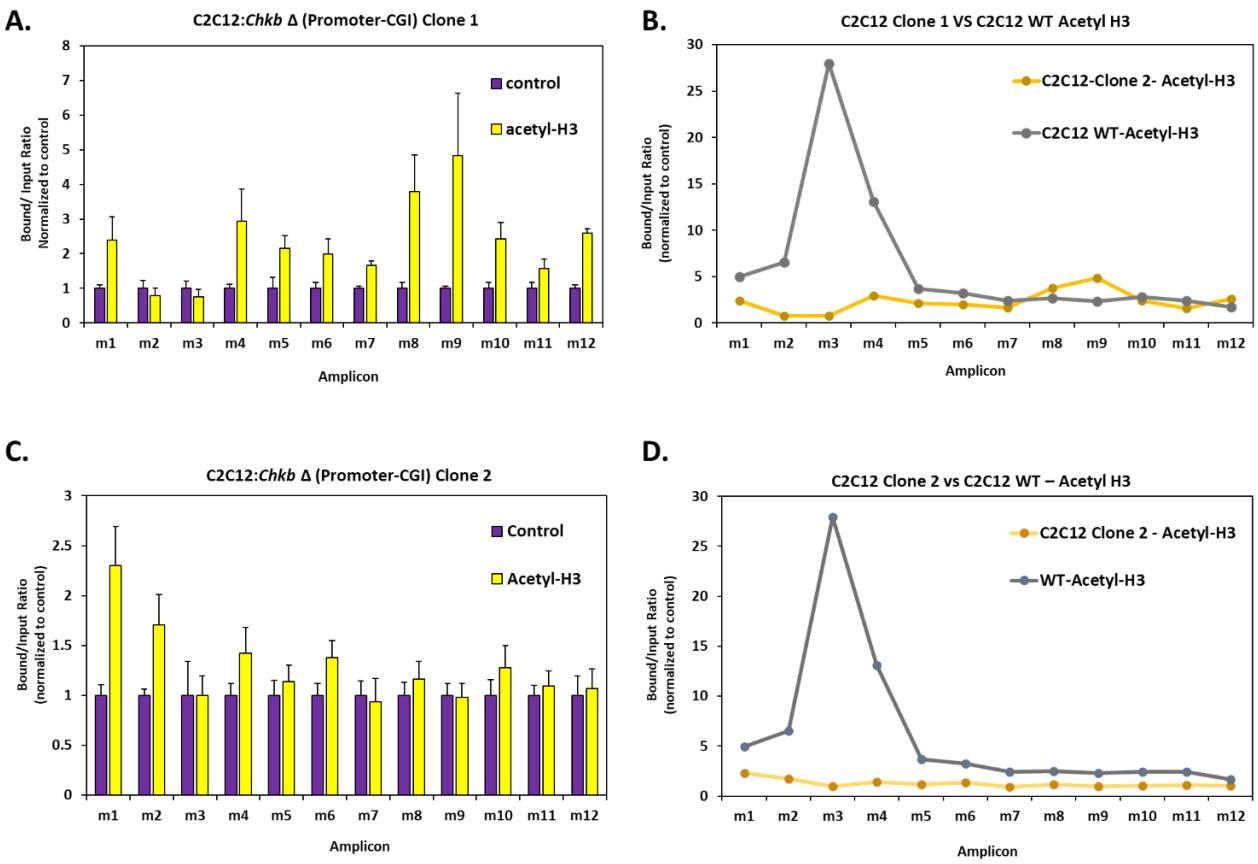

Figure 21. H3 acetylation in both C2C12 Δ*Chkb*(promoter-CGI) clones is reduced throughout the *Chkb-Cpt1b* locus compared to WT. Results of ChIP assays with a panacetyl-H3 antibody. ChIP assay enrichment levels determined by qPCR, normalized to 1% v/v of input chromatin. The exponential qPCR data (C_t values) was antilog transformed to a linear expression of bound/input ratios as Bound/Input=$2^{(\text{input Ct-bound Ct})}$. (A) C2C12 Δ*Chkb*(promoter-CGI) clone 1. (B) Comparing Acetyl-H3 Chip of C2C12 Clone 1 and C2C12 WT. (C) C2C12 Δ*Chkb*(promoter-CGI) clone 2. (D) Comparing Acetyl-H3 Chip of C2C12 Clone 2 and C2C12 WT. Amplicons are mapped in Figure 3b.

Loss of the CpG island in C2C12 mutants results in diminished RNA Pol II occupancy.

Serine-5 and Serine-2 phosphorylated RNA Pol II ChIP assays of C2C12 *Chkb*Δ(prom-CGI) clone 1 shows little apparent occupancy of either form of RNA Pol II as we notice Bound/Input DNA ratios close to 1 suggesting not much enrichment in antibody-containing ChIP samples compared to no antibody controls (Figure 22A-B). This loss in RNA Pol II processing is more apparent when we compare the RNA Pol II Ser5P and RNA Pol II Ser2P Chip assays between C2C12 clone 1 and WT C2C12 cells, respectively (figure 22C-D). As a result, transcription of *Cpt1b* is fully dependent on the recruitment of transcriptional machinery to the *Chkb* 5'- flanking region and specifically to the CpG island in C2C12 skeletal muscle cells.

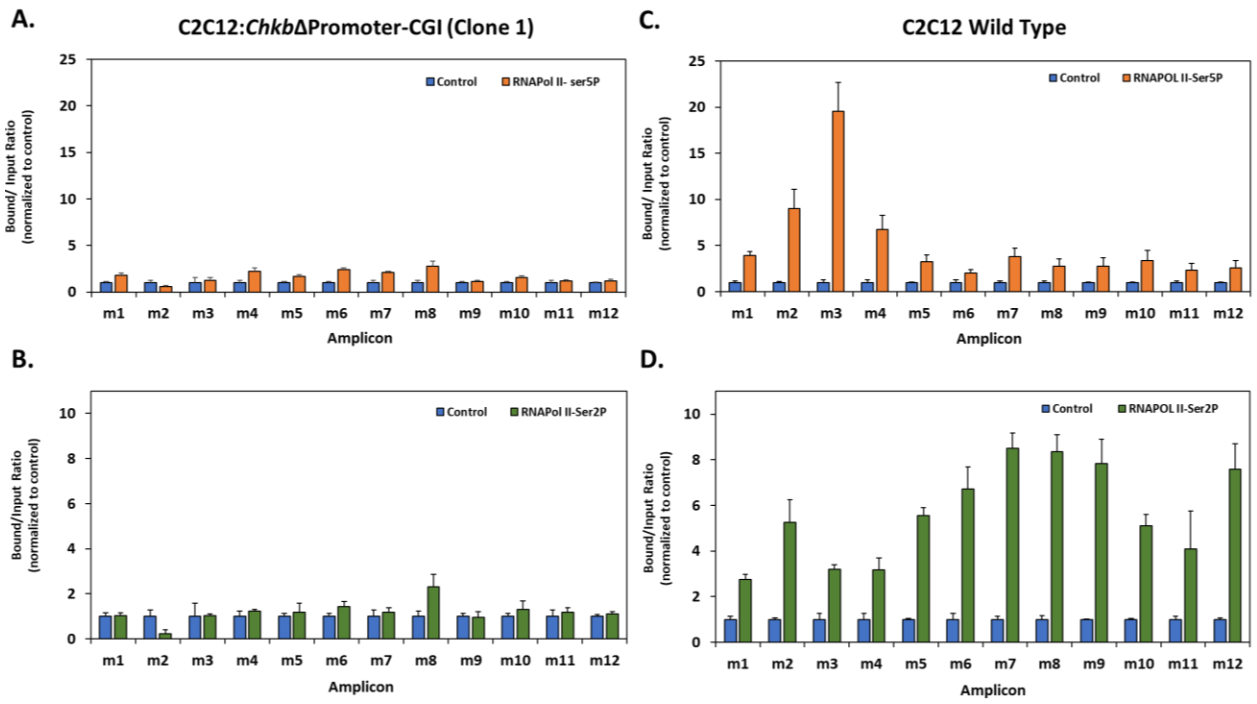

Figure 22. Deletion of *Chkb* CpG island in C2C12 cells results in little to no transcriptional activity throughout the *Chkb-Cpt1b* locus. Results of ChIP assays. **C2C12: *Chkb* Δ(promoter-CGI) clone 1**: (A) RNAPII serine 5 phosphorylation (initiating form), (B) RNAPII serine 2 phosphorylation (elongating form). **C2C12 wild type**: (C) RNAPII serine 5 phosphorylation (initiating form), (D) RNAPII serine 2 phosphorylation (elongating form). Amplicons are mapped in figure 3B.

Chkb, _Cpt1b_, and readthrough transcripts are reduced in the DE 2.3 _Chkb_ Δ(promoter-CGI) mutant.

Chkb transcripts in brown adipocyte mutants with a full deletion of *Chkb* promoter and CpG island were expectedly reduced by a significant margin compared to WT DE 2.3 ($P < 0.05$) (Figure 23A). However, unlike the skeletal muscle C2C12 cell mutants, *Cpt1b* transcript levels in brown adipocyte mutants were only reduced to a quarter of the levels observed in wild type brown fat cells (Figure 23B) and readthrough transcripts were reduced but not undetectable compared to WT (Figure 23C). This was a surprising observation and what it meant was that perhaps the regulation and activation of the *Cpt1b* gene in brown fat cells was not completely dependent on *Chkb,* but rather the evolutionary cooption resulted in elevated levels of *Cpt1b* transcripts that supplemented the basally present *Cpt1b* transcripts produced via an independent mechanism targeted at the promoter of *Cpt1b*. While transcript levels of *Slc25a20* and *Cpt2* are relatively similar in mutants compared to WT DE 2.3 cells (Figure 23 D-E), we notice a greater than 3-fold increase in *Cpt1a* transcripts in DE 2.3 mutants vs WT that was statistically significant ($P>0.05$) (Figure 23F).

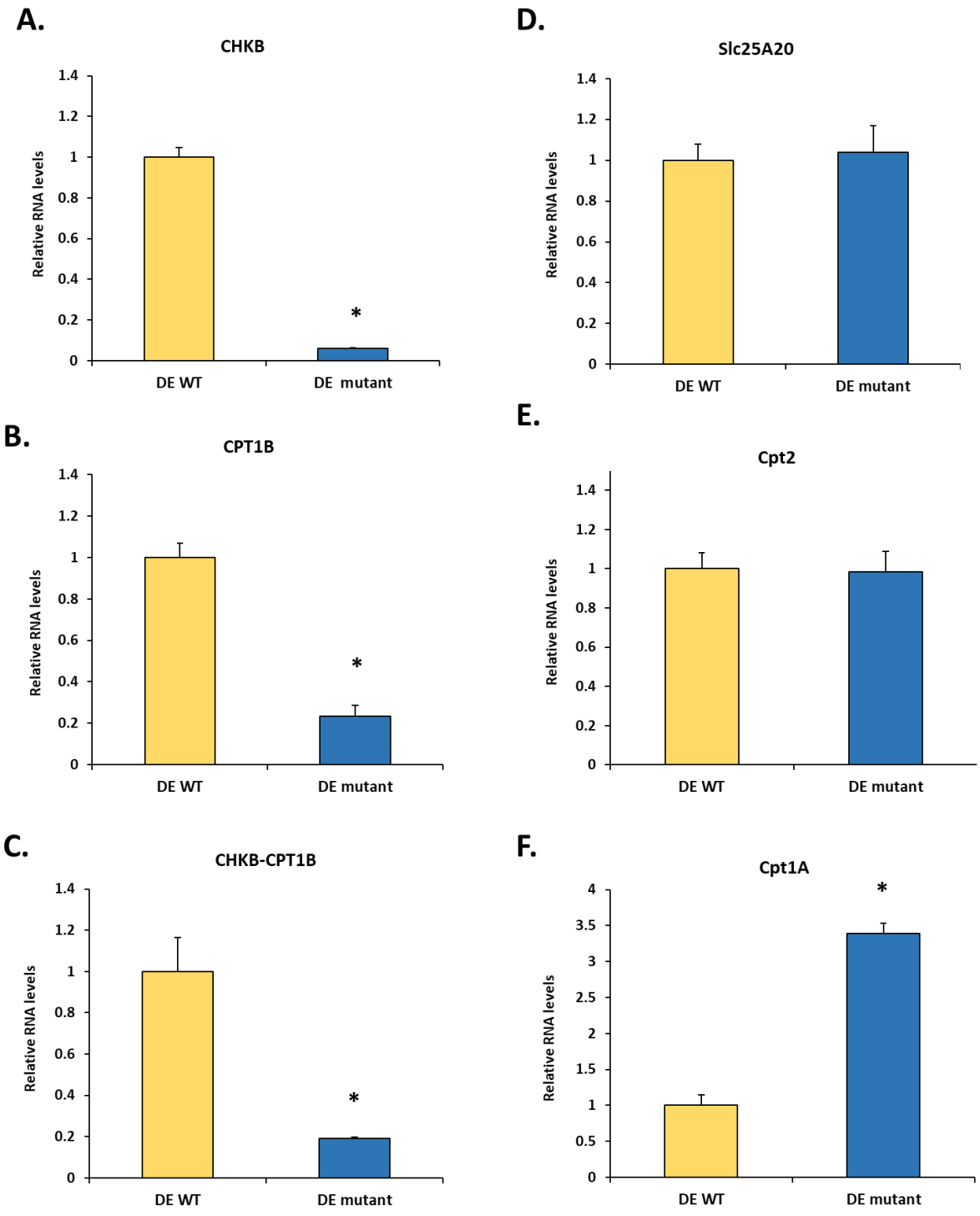

Figure 23. *Chkb* and *Cpt1b* expression are reduced in the DE 2.3 *Chkb* CpG island deletion clone compared to WT. (A) *Chkb* transcripts are reduced by approximately 17-fold

compared to WT. (B) *Cpt1b* transcripts are reduced by approximately 4.5-fold compared to WT. (C) *Chkb-Cpt1b* readthrough transcripts are reduced by approximately 5-fold compared to WT. (D) *Slc25a20* transcripts show no significant difference between the WT and mutant. (E) *Cpt2* transcripts show no significant difference between the WT and mutant. (F) *Cpt1a* transcripts are increased by approximately 3.5-fold compared to WT. Internally controlled mRNA levels (vs. *β2-microglobulin* mRNA) were normalized to the WT controls. The data indicate the mean ± *SE* (*n*=3). Student's t-test was used to determine statistical significance between wild-type and mutant cells (*P < 0.05).

H3K4me3 is diminished in DE 2.3: *Chkb*Δ(Promoter-CGI) deletion mutants compared to WT

To further characterize the activity at the *Chkb-Cpt1b* locus within DE2.3 mutants, we performed ChIP assays using H3K4me3 antibody as well as serine-2 and serine-5 phosphorylated RNA Pol II antibodies. HeK4me3 antibody ChIP revealed that while the deleted sequence encompassing primers m2 and m3 was expectedly not present, there was still a relatively higher trimethylation signal present at the m1 primer position and a moderate signal at the m5 position relative to the other sites on the mutant *Chkb-Cpt1b* locus (Figure 24A). However, when we compare the H3K4me3 ChIP assay in De2.3 deletion mutants to WT DE 2.3 cells, we see a drastic reduction in H3K4me3 signal at all sites of the locus and therefore we can conclude that deletion of the CpG island results in a significant loss of the “open” chromatin structure driven by the *Chkb* CpG site (Figure 24B). However, it is important to note that we still see moderate enrichment of H3K4me3 at the *Cpt1b* promoter which may be contributing to the remaining transcript pool in these mutants.

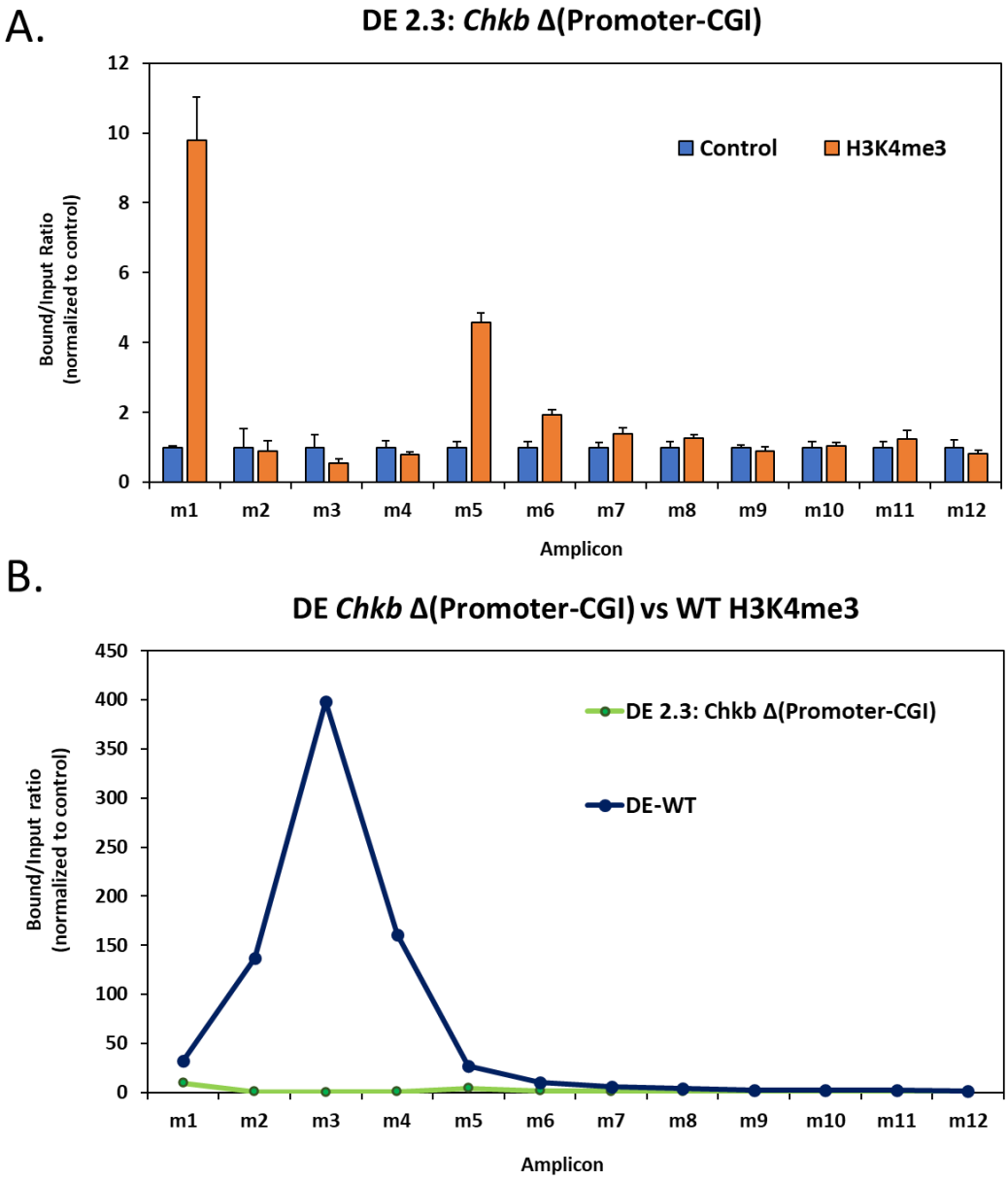

Figure 24. Histone 3 lysine 4 trimethylation (H3K4me3) is drastically reduced in DE 2.3 deletion mutants compared to WT. (A) DE 2.3:*Chkb*Δ(Promoter-CGI) deletion. (B) Comparison of H3K4me3 ChIP between DE 2.3 deletion mutant and De 2.3 WT. Amplicons are mapped in figure 3B.

Pol II occupancy is drastically reduced at all sites throughout the locus in DE 2.3 Δ*Chkb*(promoter-CGI) deletion mutants.

ChIP assays on the DE 2.3 mutant using antibodies for the initiating (serine-5 phosphorylated) form of RNA Pol II revealed maintenance of independent recruitment near the *Cpt1b* promoter represented by amplicon m5 (Figure 25A). When we compare the RNA pol II Ser5P ChIP of the mutants to wild type De2.3 cells and adjust the scale for bound/input DNA enrichment we notice that while Pol II initiation is virtually undetectable at the *Chkb* 5' region the occupancy of RNA Pol II ser5p at the 3' half of *Chkb* (m5 amplicon) *Cpt1b* promoter (m6 amplicon) in deletion mutants is comparable to WT (Figure 25C) and may explain that the 20-25% of *Cpt1b* transcripts still being produced in DE2.3 mutants are likely a result of this separate recruitment of RNA Pol II to transcribe *Cpt1b* independent of *Chkb* CGI influence. This finding is further corroborated by Bound/Input ratio of DNA being >1 for Ser2P RNA Pol II occupying amplicons of *Cpt1b* (amplicons m6-m12) compared to controls (Figure 25B) however the quantity of RNA Pol II Ser2P is drastically diminished when compared to WT DE gene loci (Figure 25D). Taken together, RNA Pol II Chip assays along with qRT-PCR transcript data for this mutant suggest that the ablation of the *Chkb* CpG island and 5'-flanking region results in a 4 to 5-fold reduction of *Cpt1b* transcripts and the residual transcripts that are still being produced is a result of transcription initiation still occurring near the *Cpt1b* promoter. Therefore, we can conclude that in brown fat cells, the co-option of *Cpt1b* transcription and regulation by the *Chkb* 5' flank and CpG island resulted in 4-5 fold increase of Cpt1b activity to fuel BAT function. Furthermore, ChIP assays using PRDM16 and PGC1a antibodies indicated independent binding at the *Chkb* 5' region (especially the CpG island) and the 5' of the *Cpt1b* gene (Figure 26 A-B). Briefly, PRDM16 has been shown to control a bi-directional switch between skeletal muscle and brown fat cell fate as well as promoting BAT-specific genes and downregulating

expression of WAT specific genes through chromatin mediator complex recruitment [58, 63, 64, 170]. PGC1a is a master regulator of mitochondrial biogenesis and coregulator of genes involved in energy metabolism. It's important to note that ChIP assays using anti-PRDM16 and anti-PGC1a showed a more robust interaction at the *Chkb* 5' region compared to the 5' region of *Cpt1b*, suggesting that while *Cpt1b* transcription may not be completely dependent on the *Chkb* 5' region, a significantly higher contribution to the *Cpt1b* transcript pool is made by initiation at the *Chkb* 5' region. The interactions of a brown fat specific TF in Prdm16, and a transcriptional coregulator of FAO genes in PGC1a with the *Chkb-Cpt1b* gene locus is quite significant and provides further evidence for their coordinated nature in BAT.

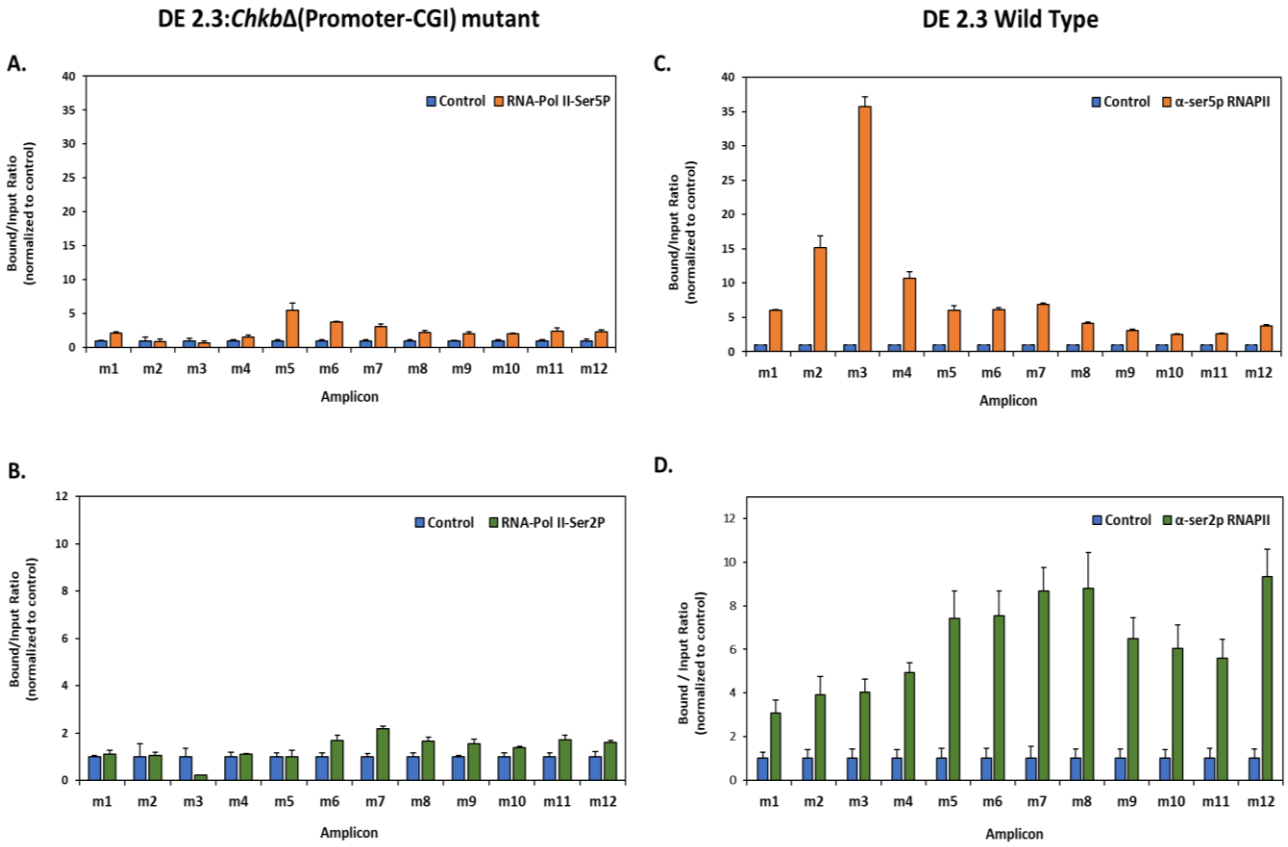

Figure 25. Independent RNA Pol II initiation at the intergenic region and elongation across *Cpt1b* may be contributing to the transcript pool that is still detectable in brown fat cells with the *Chkb*Δ(promoter-CGI) deletion. Results of ChIP assays. **DE2.3: *Chkb*Δ(promoter-CGI) mutant**: (A) RNAPII serine 5 phosphorylation (initiating form), (B) RNAPII serine 2 phosphorylation (elongating form). **DE 2.3 wild type**: (C) RNAPII serine 5 phosphorylation (initiating form), (D) RNAPII serine 2 phosphorylation (elongating form). Amplicons are mapped in figure 3B.

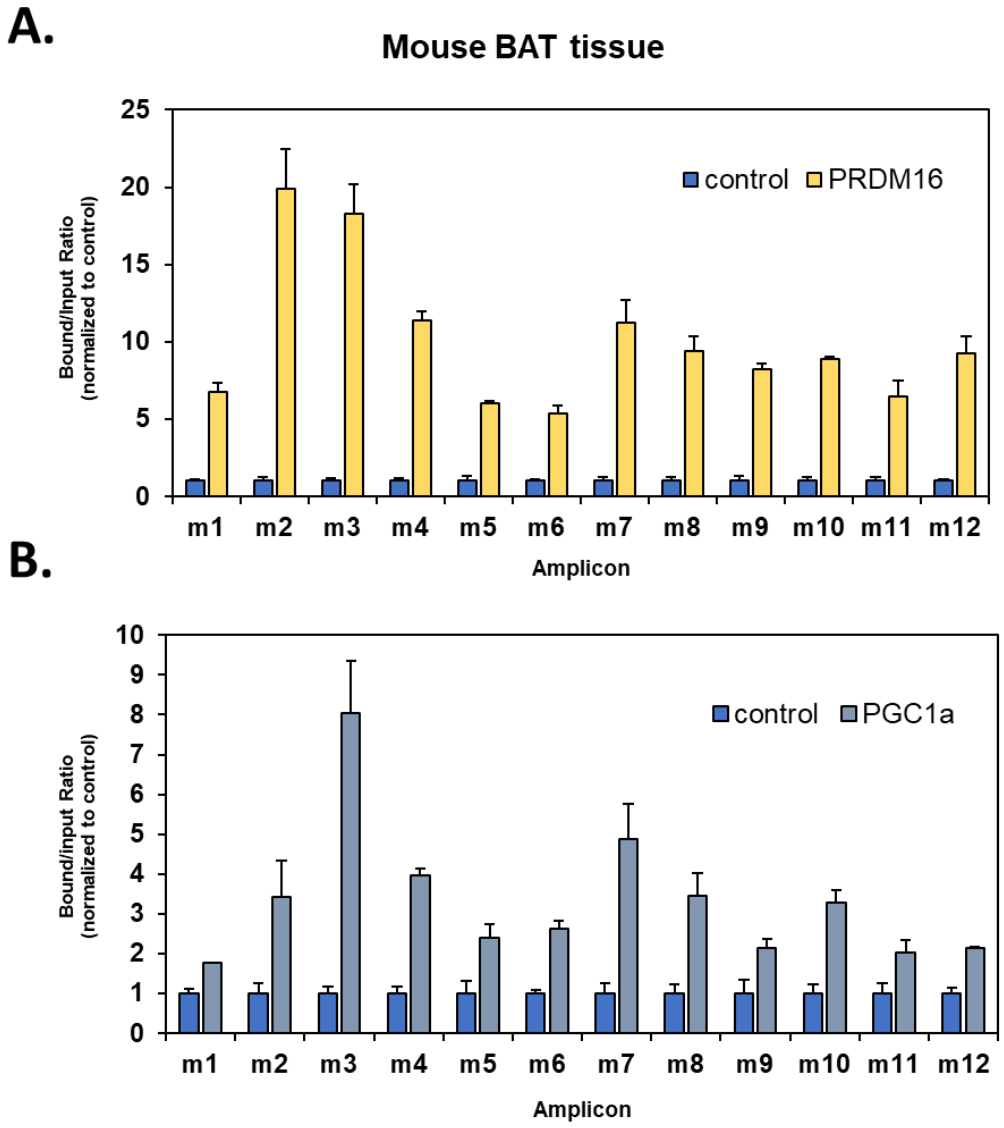

Figure 26. Independent binding of PRDM16 and PGC1a at the *Chkb* 5' region and *Cpt1b* gene. (A) Anti-PRDM16 ChIP qPCR results in murine BAT shows robust binding at the Chkb 5' flank and CpG island along with additional recruitment at the beginning of the *Cpt1b*. (B) Anti-PGC1a ChIP qPCR results in murine BAT show binding at the *Chkb* CpG island along with an additional peak of localization at the beginning of *Cpt1b* gene

Respirometry assays indicate a similar low rate of FAO in WT and mutant cells, but an overall decrease in mitochondrial oxidative function in mutant cells.

Figure 27 shows Oxygen consumption rates of C2C12 wt and *Chkb*Δ(promoter-CGI) deletion clone 1 using Oroboros O2K respirometer. Both samples were grown and differentiated into myotubes as previously described. Cells were collected and washed with 1xPBS after which point the cells were suspended in Buffer Z (1mM EGTA) to a final concentration of 0.5 million cells/mL. Digitonin was used at a final concentration of 30ug/mL as a mild detergent to permeabilize the cells for 10 minutes with gentle agitation. Cells were centrifuged at 1000 RPM for 10 minutes and washed once with fresh buffer Z. The cell pellet was suspended in the same volume of buffer Z (0.5 million cells/ mL) before running on the O2K respirometer. 1mM malate and 0.5mM ADP were added together first which caused an increase in O_2 consumption signaling the presence of an endogenous substrate (figure 27 A-B). Addition of 20uM Palmitoyl-CoA resulted in no increase in respiration in WT or mutant cells which persisted even after the addition of 5mM carnitine (figure 26 A-B). This finding would not be surprising in Cpt1b KO mutants however the same finding in WT cells indicates that perhaps C2C12 cells are not utilizing LCFAs at a high level. To assess cell viability and mitochondrial function, PMGSO (5mM pyruvate, 1mM Malate, 5mM glutamate, 5mM succinate, and 250uM of octanoyl-carnitine) was added together which indicated a dampened change in oxygen consumption rate (ΔJO_2) in response to the PMGSO substrates in C2C12 *Chkb*Δ(promoter-CGI) deletion clone 1 compared to WT. The ΔJO_2 after addition of PMGSO in WT C2C12 myotubes was 201 [pmol/(s*ml)] (figure 27A), while the ΔJO_2 after addition of PMGSO in C2C12 deletion mutants was only 97 [pmol/(s*ml)] (Figure 27B). This indicates that while the cells from both sample groups were still viable, the depletion of *Chkb* and *Cpt1b* in the mutant resulted in an overall dampened mitochondrial response compared to WT cells. As a control, the sample C2C12 wt

and mutant samples were assayed using palmitoyl-carnitine as a substrate, thereby bypassing CPT1B activity, however we again detected no changes in oxygen consumption in either cell type in response to only PC as the substrate (figure 27 C-D). However, we did again notice a diminished ΔJO_2 after addition of PMGSO in C2C12 *Chkb*Δ(promoter-CGI) compared to WT. The ΔJO_2 after addition of PMGSO in WT C2C12 myotubes was 153 [pmol/(s*ml)] (figure 27C), while the ΔJO_2 after addition of PMGSO in C2C12 deletion mutants was only 113 [pmol/(s*ml)] (Figure 27B). The DE 2.3 WT and mutant cells also showed very low respiration from FAO and no differences between the two cell types in response to palmitoyl-CoA and carnitine possibly due to the cytotoxic effects of the LCFAs (data not shown).

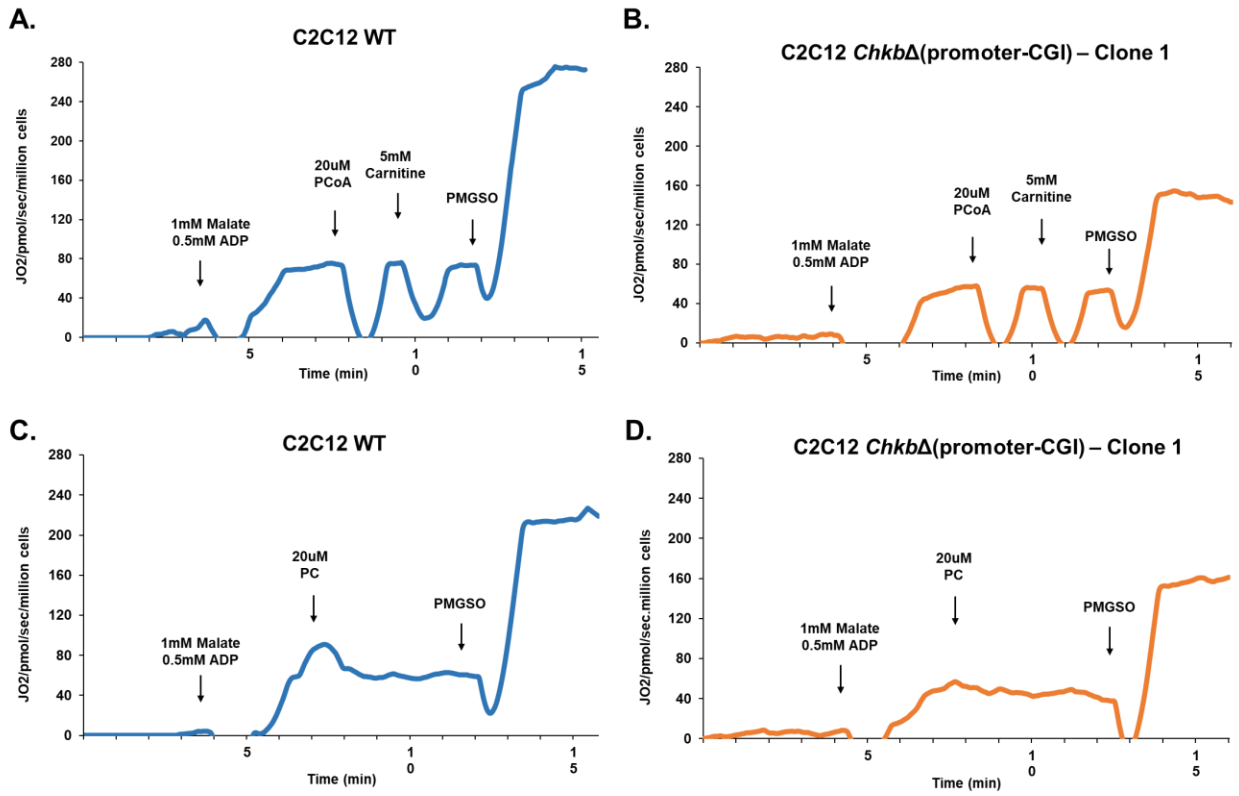

Figure 27. O2K respirometry assays show low rates of FAO in C2C12 WT and C2C12 *Chkb*Δ(promoter-CGI) clone 1 myotubes. (A-B) pre-permeabilized myotubes were assayed for oxygen consumption rate when supplied with palmitoyl-CoA and Carnitine sequentially to assess the Carnitine shuttle system and CPT1B function. PMGSO= pyruvate, malate, glutamate, succinate, octanoyl-carnitine added together to assess cell viability/ mitochondrial function. (A) C2C12 wt myotubes. (B) C2C12 *ChkbΔ*(promoter-CGI clone 1 myotubes. (C-D) Pre-permeabalized myotubes were assayed for oxygen consumption rate when supplied with palmitoyl-carnitine as a substrate to bypass the activity of CPT1B. (C) C2C12 wt myotubes. (D) C2C12 *ChkbΔ*(promoter-CGI) clone 1 myotubes.

***Chka* is expressed at similar levels to *Chkb* in WT cells and the expression remains unchanged between WT and mutant cells.**

We wanted to also look at the expression of the other isoform of Choline kinase, *Chka*, in both of our wild-type cell lines at basal conditions and also wanted to compare transcript levels of *Chka* between mutants and WT. We found that *Chka* is expressed similarly between C2C12 mutants and WT (figure 28 A) and DE mutants compared to WT (Figure 28 B). We also wanted to compare the levels of expression of *Chka* to *Chkb* in WT cells and found that C2C12 WT cells express both forms of choline kinase at similar levels (figure 28 C) and DE WT cells also express both forms of choline kinase at similar levels (figure 28 D). The lack of any severe morphological differences between mutants and WT cells from both lines can be attributed to compensatory PC biosynthesis via CHKA.

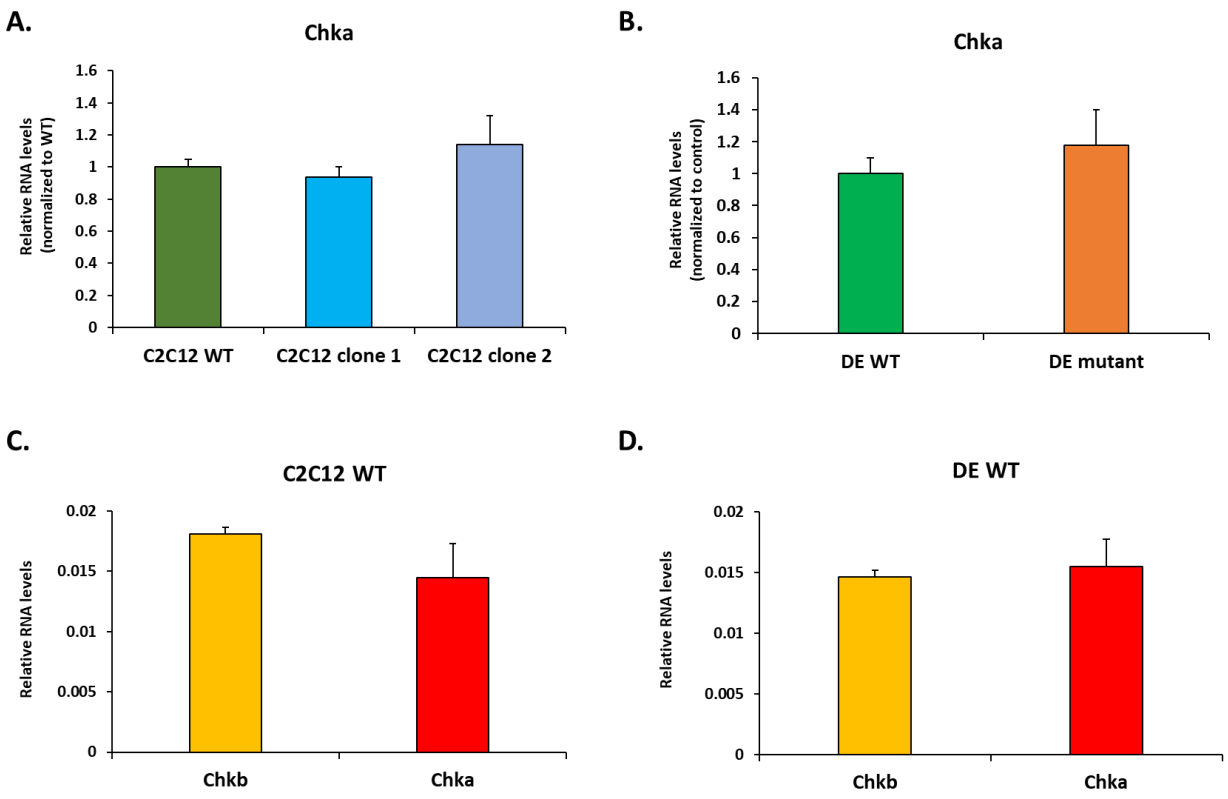

Figure 28. *Chka* is also expressed in WT and mutant cells at similar levels. (A) *Chka* transcript levels are similar between C2C12 mutants and WT cells. (B) *Chka* transcript levels are similar between DE mutant and WT cells. (C) Basal *Chka* transcript levels are only slightly less than *Chkb* transcript levels in C2C12 WT cells. (D) Basal *Chka* transcript levels are similar to *Chkb* transcript levels in DE WT cells. Internally controlled mRNA levels (vs. β2-microglobulin mRNA) were normalized to the WT controls. The data indicates the mean ± SE (n=3). Student's t-test determined no statistical significance in transcript levels between wild-type and mutant cells ($P > 0.05$). Student's t-test determined no statistical significance between *Chkb* and *Chka* transcript levels in wild-type C2C12 or DE 2.3 cells ($P > 0.05$).

Conclusions

Our findings using Crispr/Cas9 to delete the *Chkb* 5' flank and CpG island in transformed murine C2C12 skeletal muscle and DE 2.3 brown fat cell lines showed significant changes in chromatin modifications and transcriptional regulation. We can confidently state that the *Chkb* CpG island is crucial in maintaining an open chromatin domain that encompasses both the *Chkb* and *Cpt1b* genes in WT cells while deletion of the CpG island results in a severe loss of acetylation and trimethylation across the entire locus as well as diminished RNA Pol occupancy in both mutant cell lines. Furthermore, transcriptional initiation predominantly occurs at the *Chkb* CpG island, and the RNA polymerase is able to either transcribe the entire locus or be delivered to the *Cpt1b* promoter region via a looping mechanism to transcribe the *Cpt1b* gene as a separate product from *Chkb*. Interestingly, we did discover independent recruitment of RNAP II at the Cpt1b promoter region in DE 2.3 brown fat cells that seems to be functional, albeit to a much smaller degree compared to the *Chkb* 5' region. We find that in DE 2.3 this independent transcription may be contributing to the transcript pool, yet we still observe an almost 75-80% reduction in transcript levels compared to WT cells. Further evidence for a functional Cpt1b promoter is given by ChIP assays using PRDM16 and PGC1a antibodies that show while the dominant region for binding is localized at the *Chkb* CpG island, there is a separate site of binding localized to the beginning of *Cpt1b*. Another important piece of evidence for a transcriptional connection between these two genes is the presence of a detectable readthrough fusion transcript that contains exons from both genes which likely performs a regulatory role in maintaining and open and active chromatin domain. While transcription of *Cpt1b* in skeletal muscle cells seems to be almost completely dependent on the *Chkb* 5' CpG island, 20-25% of the *Cpt1b* transcript pool in brown fat cells is made up from independent *Cpt1b*-promoter mediated transcription while 75-80% is driven by *Chkb 5'*- mediated *Cpt1b* transcription.

One of the limitations associated with using transformed cell lines as our model as opposed to an animal model is that they may not being functionally relevant or similar to the actual physiological system. We observed no drastic differences in morphology between mutants and WT cells from both cell lines. As changes in PC levels would be the likely culprit for any alterations in cell membrane structure and integrity, we believed that knocking out *Chkb* would be the reason for any observed differences. The lack of differences could be attributed to the consistent expression of *Chka* that remained unchanged between WT and mutants and is expressed at a similar level compared to *Chkb* based on qRTPCR transcript data. Unfortunately, we found that functional assays to test for mitochondrial respiration using Oroboros O2K respirometer showed no significant differences in response to PCoA/PC between mutants and WT cells. In fact, even the WT cells did not respond significantly to palmitoyl-coA and carnitine or palmitoyl-carnitine as substrates. However we did observe dramatic differences in the ΔJO_2 after addition of PMGSO between WT and mutant C2C12 myotubes, with mutants showing a diminished mitochondrial response compared to WT. We can therefore say that the drop in mitochondrial function may likely be due to loss of *Chkb* and *Cpt1b*. While it was disheartening to see that neither WT nor mutant cells utilized palmitoylCoA or palmitoylcarnitine, we acknowledge that this could likely be a result of a shift in substrate preference for transformed cell lines that have been passaged continuously and maintained in high glucose media for extended periods of time. We also observed possible cell apoptosis resulting from palmitoylCoA addition in several O2K assays in both C2C12 and DE 2.3 cells (data not shown). LCFAs can have a detergent-like effect on permeabilized cells and lead to cytotoxicity; therefore, we believe that a more sensitive assay is needed using intact cells. Furthermore, we also discovered an increase in *Cpt1a* expression in mutant cells compared to WT indicating the ability for these cells to invoke a compensatory action for the loss of CPT1B. We propose that this amplified expression of *Cpt1a* may be driven by the culture media conditions in which these cells proliferate and differentiate. Even when we compare the basal expression of *Cpt1a* vs.

Cpt1b in WT cultured DE 2.3 and C2C12 cells, we find that *Cpt1a* is already more highly expressed than *Cpt1b*. Contrarily, transcript analyses of freshly excised murine cardiac, skeletal muscle, and brown fat tissue indicate a much higher expression of *Cpt1b* compared to *Cpt1a*, while *Cpt1a* expression is higher in liver compared to *Cpt1b*. These findings are expected based on what we know about *Cpt1a* and *Cpt1b* tissue specificity, so the question is, why is there such a big shift in preference for Cpt1a over Cpt1b in cultured cells? Additionally, qRT-PCR analysis of primary cells cultured from murine skeletal muscle show the same shift in preference for *Cpt1a* over *Cpt1b* after being cultured and maintained in high glucose DMEM media (supplemented with 10% FBS). Transcript levels of *Chka*, the other isoform of Choline kinase, are unchanged between the C2C12 and DE mutants and WT populations, however it is important to note that basal expression of *Chka* is comparable to *Chkb* in WT DE 2.3 cells suggesting that there isn't a strong tissue specific preference for *Chkb* in the C2C12 or DE cell lines. We also observed a reduction of Myoblast determination protein 1 (*MyoD*) transcript levels in both C2C12 *ChkbΔ*(promoter-CGI) deletion clones compared to WT, with C2C12 clone 1 showing a more significant drop in transcript levels compared to clone 2 (data not shown). Whether this is a direct outcome of ablated *Chkb* or *Cpt1b* expression is not certain and needs to be characterized further.

It is possible that respirometry assays using O2K may not be sensitive enough to detect the low level of FAO occurring in C2C12 and DE 2.3 cells, and therefore a more sensitive assay is required. Using radiolabeled [1-^{14}C]-palmitic acid as a substrate supplied directly in basal glucose-free media to intact cells, and measuring CO_2 production would be a more sensitive detection method and would also allow for longer incubation time. To determine if CPT1A is in fact able to compensate for loss of CPT1B in mutant DE 2.3 and C2C12 cells, future studies will need to be done to known down *Cpt1a* and assess mitochondrial respiration and ATP production from LCFA substrates. Furthermore, it will be more informing and physiologically

relevant to generate a transgenic mouse model with the same *ChkbΔ*(promoter-CGI) deletion. This would allow for body temperature measurements and whether the mice would have difficulties dealing with cold exposure. Additionally, a transgenic mouse model would also make it possible to test how the mice respond to cold challenges, and what effects the catecholamine stimulation has on brown adipocyte induction in mice with a loss of *Chkb* and *Cpt1b*. Furthermore, it would also be possible to test whether the mutant mice are more prone to obesity and other metabolic disorders when provided with high-fat diet compared to WT mice.

These findings are crucial in establishing a regulatory and transcriptional connection between the Chkb and Cpt1b genes in both skeletal muscle and brown fat cells. We also were able to elaborate on the importance of CpG islands in maintaining an open chromatin domain and being the major site for transcriptional initiation. The loss of the *Chkb* CpG island therefore resulted in a complete loss of *Chkb* and almost a complete loss of Cpt1b transcripts, providing further evidence for the dependence of *Cpt1b* on *Chkb* after the condensation of the intergenic space not only brought the genes in close proximity but also removed the *Cpt1b*- specific CpG island in eutherians.

Materials and Methods

CRISPR Design and Planning

SgRNA-specifying oligo sequences were chosen with the smallest off target cleavage scores according to the CHOPCHOP algorithm available online. The selected sgRNAs targeted either the proximal promoter sequence upstream of the *Chkb* transcription start site or exonic sequences within the coding sequence of the *Chkb* gene according to refseq genomic sequence. A 4-nucleotide “CACC” sequence was added to the 5’ end of the sgRNA- specifying sequence and “AAAC” was added to the 5’ end of the reverse complement of the sgRNA-specifying oligo sequence for cloning using the BbsI restriction enzyme. Additionally, considering that the expression of the sgRNA is driven by the U6 RNA polymerase III promoter in these constructs, a Guanine nucleotide “G” was added immediately following the “CACC” sequence due to its preference for a guanine nucleotide as the first base of the transcript. Subsequently, a “C” was added to the reverse complement sgRNA following the “AAAC” sequence. The type II CRISPR system is one of the most well characterized systems that uses the Cas9 nuclease to perform double stranded genomic cuts. One stipulation, however, for Cas9 is the requirement of a protospacer adjacent motif (PAM) recognition sequence. Therefore, the target DNA sequence must immediately be followed by a 5’-NGG-3’ PAM site for Cas9 to induce a double strand break, alternatively; the gRNA for the reverse complement strand should be followed by a 5’-NCC-3’ motif. The DNA break occurs exactly 3 base pairs upstream of the PAM recognition site on both strands.

We utilized the CHOPCHOP web tool (https://chopchop.cbu.uib.no/) to pick the optimal gRNAs that targeted cut sites at or near our desired regions. All chosen gRNAs were within the top 10 list based on Cas9-specific efficiency, off-target effects and mismatches, existence of self-complementarity regions, and GC content. sgRNA 1-3 sequences are listed in table 3. pSpCas9(BB)-2A-Puro CRISPR Cas9 (PX459) V2.0 was a gift from Feng Zhang (Addgene

plasmid # 62988 ; http://n2t.net/addgene:62988 ; RRID:Addgene_62988) that already contained the tracrRNA binding scaffold sequence and only required the insertion of the “25-mer” sgRNA annealed oligos. Lyophilized gRNA oligos were resuspended in nuclease-free molecular grade dH2O to a final concentration of 100 µM. 300 pmols of the top and bottom single stranded oligos were first phosphorylated separately using T4 Polynucleotide Kinase (10 units) (New England Biolabs) at 37ºC for 1 hour. The enzyme was heat-inactivated at 65ºC for 30 minutes. 30 pmol of top and bottom phosphorylated ssDNAs were subsequently annealed in annealing buffer with the following temperature conditions: 95ºC for 5 minutes; and then removing the heat block and letting it gradually cool down to 22ºC on the benchtop. 1µg of circular PsPCas9(BB) plasmid was then digested using BbsI restriction enzyme (20 units) (New England Biolabs) in NEB 2.1 buffer (1x) at 37ºC overnight. 100ng of the digested plasmid was run on a 1% agarose TAE gel to confirm complete digestion. The presence of a single band at 9.175kb confirms complete digestion of the plasmid into linear DNA. The annealed sgRNA oligos were subsequently cloned into pSpCas9(BB) using a 3:1 insert to vector molar ratio. 100ng of digested PspCas9 (BB) (0.017 pmol) and 820 ng of annealed oligos (0.053 pmol) in 1x DNA ligase buffer and 1 µL of T7 DNA ligase enzyme (3000 units) (New England Biolabs). The reaction was placed in an ice bath at approximately 12ºC and allowed to gradually warm to room temperature (22ºC-25ºC) overnight.

Bacterial transformation

Chemically competent JM109 E Coli were transformed with the sgRNA:Cas9 ligated plasmids using the heat-shock method. 100uL of JM109 cells were pipetted into appropriately labeled round bottom tubes with 100ng of respective gRNA: Cas9 plasmids and incubated on ice for 15 minutes. Each tube was then placed into a 42ºC water bath for 45 seconds followed by a 2-minute incubation on ice. 1 mL of LB media was added to each tube and the tubes were

incubated at 37^{o}C and shaken at 225 RPM for 1 hour. After an hour, bacterial growth was observed indicated by turbid media. 100uL of the bacterial culture was spread onto LB agar plates containing 150 μg $^{-1}$ ampicillin and incubated at 37^{o}C overnight. The next day, plates were inspected for colony growth and 5-10 colonies, per transformation, were picked to check for proper insertion of the sgRNA into the Cas9 plasmid. These colonies were grown in 5 mL of 150 μg $^{-1}$ ampicillin containing LB media at 37^{o}C with 225 RPM overnight, after which point plasmid DNA was extracted and precipitated from 1.5mL of bacteria culture using the alkaline lysis protocol described in the Sambrook and Russell Molecular cloning laboratory manual.

SgRNA:Cas9 Plasmid verification

Plasmids were checked for proper gRNA insertion using BbsI digestion, where plasmids with the insertion would be protected from enzyme cutting and therefore would be distinguishable from digested vectors using gel electrophoresis. After overnight digestion, DNA samples were loaded and ran on a 1% agarose gel in TAE buffer. Clones containing plasmids with gRNA insertion were identified by the presence of a DNA smear and multiple bands representing circular, and supercoiled plasmid conformations. 1 mL of bacteria culture medium from positive colonies was transferred to 500mL of LB media containing 150 μg^{-1} ampicillin and incubated overnight in 37^{o}C with vigorous shaking (250 RPM). To obtain a large quantity of pure transfection grade sgRNA:CAS9 plasmids, large-scale plasmid preps were performed using the Qiagen Maxiprep kit according to their protocol. Plasmids were dissolved in nuclease-free dH_2O, and concentrations were checked using NanoDrop® ND-1000 UV-Vis Spectrophotometer (Thermo-Scientific). To confirm that these plasmids had been successfully ligated with the sgRNA insert, 0.5 μg of plasmid DNA from sgRNA:Cas9 insert plasmids as well as negative control PspCas9(BB) vectors were digested overnight with 5 units of BbsI (NEB) according to the manufacturer's protocol. Gel electrophoresis of the digestion products showed the presence of circular and supercoiled plasmid bands for the sgRNA:Cas9 plasmids as opposed to a single

band of linear digested control vectors at 9.2kb, suggesting that the sgRNAs had been ligated into the vector and resulting in the loss of BbsI cut site. To determine if the inserted sgRNA was ligated correctly, samples were sent off to get sequenced via Sanger sequencing from the U6 promoter using the U6 promoter Forward primer.

Transfections

To set up for transfections, C2C12 myoblast cells and DE 2.3 pre-adipocyte cells were plated from frozen stock and placed in an incubator. Once the plates reached 80-90% confluence, cells were split into 3 wells of a 6 well plate per transfection, so that each well would be 80-90% confluent by the time of transfection. Since transfections were performed using Lipofectamine 2000, the cells were plated using antibiotic-free DMEM media with 10% FBS to prevent antibiotic-induced cytotoxicity during plasmid uptake. Transfections were performed using lipofectamine 2000 according to the manufacturers protocol (Invitrogen). For a single well of cells in a 6-well plate, the total amount of DNA to be used for transfection is 2500ng. We therefore added equal molar amounts of 5' -cas9:sgRNA1 and 3' -Cas9:sgRNA2/3 constructs to induce two simultaneous double stranded breaks and subsequent deletion of targeted sequences. Furthermore, 1/10 of a molar volume of pMaxGFP plasmid was added to serve as a fluorescent marker. 1200ng of Cas9:sgRNA1 plasmid and 1200ng of Cas9:SgRNA2 were mixed equaling to a total molar concentration of 0.414pmol to which 90ng of pMaxGFP was added which equated to 0.041pmols into 150uL of Opti-MEM reduced serum media (Gibco) (per well). We also transfected one well from each cell line with 2400ng PspCas9(bb) empty Cas9 vector (0.414 pmol) with no gRNA along with 1/10 molar volume of GFP (0.041 pmol) to serve as a negative control. A 3:1 ratio of lipofectamine 2000 volume per microgram of total DNA was empirically confirmed to provide the best transfection efficiency therefore, 7.5uL of Lipofectamine 2000 reagent was diluted in a separate tube containing 150uL of Opti-MEM media. The contents of the two tubes were combined and incubated at room temperature for 15

minutes before slowly pipetting the mixture drop by drop into the respective wells. Cells were placed back into the 37°C incubator overnight. After 24 hours, cells were checked for GFP fluorescence using light microscopy. Transfected C2C12 wells indicated around a 30% transfections efficiency (based on 1/10 molar volume of GFP), while De 2.3 adipocytes had a 15% transfection efficiency.

Fluorescence activated cell sorting.

We elected to use pMaxGFP plasmid (Lonza) as a visual reporter for transfection efficiency and plasmid uptake. Cells were first lifted with light trypsinization and counted using a hemocytometer; a small aliquot of the cell suspension was mixed with 0.4% Trypan Blue at a 1:4 volume ratio to assess cell viability. Cells were then centrifuged, and the supernatant was aspirated; based on the cell count, enough Fluorescence activated cell sorting (FACS) buffer was added to the cell pellet to generate a single cell suspension of 1 million cells/mL. The FACs buffer composition was 25mM of HEPES buffer containing 5mg/mL BSA adjusted to a pH of 7.5. The BD FACSAria fusion flow cytometer was used to sort single GFP-fluorescent cells into 96 well plates. We avoided high GFP-fluorescing cells as there is a good chance a high amount of GFP can induce cell death via ER-mediated stress. Instead, we choose to sort cells with medium to low GFP intensity to avoid this problem. All wells contained 100uL of 25% conditioned media to promote single cell division and proliferation. Briefly, conditioned media was made by first growing wild type cells to 60-70% confluence in regular high glucose DMEM media with 10% serum and antibiotics. While the cells were still in log growth phase, the media was carefully removed, sterile filtered through a 0.2 µm filter, and transferred into sterile 15mL tubes. 25% conditioned media was made by mixing the conditioned media with fresh regular DMEM supplemented with 20% FBS and antibiotics at a volume ratio of 1:3. Plates were placed into the 37°C incubator for 10-14 days to allow for wells to get confluent while regularly checking the plates every 2 days and replenishing with fresh growth media as needed.

Colony growth and expansion

Once wells in 96-well plates were at least 90% confluent, cells were trypsinized and 80% of the cell solution was plated into a replica 96-well plate while the remaining 20% of the cells replated in the same plate to continue proliferating for expansion. The cells were expanded from 96 wells into 24 wells and finally into 6 well plates. After reaching 80-90% confluency in 6 well plates, cells were frozen down using complete media containing 20% FBS and 8% DMSO and kept in -80 °C. Once the replica 96-well plates wells were close to 100% confluency, the cells were lifted and collected in 1.5mL Eppendorf tubes for DNA extraction.

Genomic DNA extraction

Cell pellets were suspended in 100uL ice cold 1x PBS containing 40ug of Proteinase-K and 30ug of RNase A after which point, 100uL of cell lysis buffer (10mM Tris pH 8.0, 100mM NaCl, 10mM EDTA pH 8.0, 0.5% SDS) was added to the cell suspension and mixed by inverting repeatedly. Cells were lysed at 56 °C with occasional vortexing. DNA was extracted with an equal volume of buffer-saturated phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated using 2.5x volume of 100% ethanol. DNA pellets were washed with 70% ethanol and suspended in nuclease-free water. The concentration and purity were measured using NanoDrop.

Genomic Endpoint PCR screening

Polymerase chain reaction (PCR) was performed using different sets of primer pairs flanking the deletion, listed in table 5, using the OneTaq 2x master mix with standard buffer. In some cases, a primer pair encompassed within the *Chkb*(promoter-CGI) deletion, primer pair E3 listed in table 5, were also used as an additional verification step for deletion. PCR was performed using the following cycling conditions: 95 °C for 3 min; 35 cycles of 95 °C for 30 s, 60 °C for 45 s, and 68°C for 1 min/kb; 68 °C for 5 minutes. PCR products were separated by gel

electrophoresis with ethidium bromide staining and visualized using UV fluorescence on the ChemiDoc™ Imaging System (BioRad). The presence of WT Nondeletion band would indicate that the deletion has not taken place, the presence of a single deletion band would indicate a bi-allelic deletion, and the presence of both deletion and nondeletion bands would represent a heterozygous deletion.

Genomic qPCR screening

A second level of screening was also done using qPCR primer pairs that amplify 100-150bp regions of the DNA spanning the entire *Chkb-Cpt1b* locus. CRISPR screening Amplicons are mapped in figure 3B and 5B for C2C12 and DE2.3 mutants, respectively. Genomic qPCR was performed using and the following cycling conditions: 95 °C for 5 min; 35 cycles of 95 °C for 30 s, 60 °C for 45 s, and 68°C for 1 min/kb; 68 °C for 5 minutes using Power SYBR Green PCR Master Mix (Thermo Fisher Scientific) in accordance with manufacturer's instructions using a C1000 Touch™ Thermal Cycler (Bio-Rad, Hercules, CA).

RNA extraction and qRT-PCR analysis

All primers for quantitative reverse transcription and PCR (qRT-PCR; Table 1) and genomic quantitative PCR (qPCR; Table 2) were designed and validated as previously described to ensure that the requirements for quantification by the comparative threshold cycle (C_t) method were met [151]. RNA purification, cDNA synthesis, and qPCR were performed as previously validated and described in detail [151]. Briefly, cells were directly lysed in 6 well plates using trizol reagent and RNA was purified according to the manufacturer's instructions. RNA was reverse transcribed using the Verso cDNA synthesis kit (Thermo Fisher Scientific, Waltham, MA). QPCR was performed with Power SYBR Green PCR Master Mix (Thermo Fisher Scientific) in accordance with manufacturer's instructions using a C1000 Touch™ Thermal Cycler (Bio-Rad, Hercules, CA). The specified primer pairs are shown in Table 1.

Transcript levels were measured by the comparative C_t method using beta-2 microglobulin (B2M) as an internal control.

Puromycin selection

Puromycin antibiotic kill curves were performed to empirically derive the lowest concentration of puromycin that would result in 100% cell death for wild type C2C12 and DE 2.3 cells. Cells were equally split and grown in 24-well plates until 80-90% confluent. Media was replenished with varying concentrations of puromycin ranging from 0.5µg/mL to 10µg/mL and plates were incubated 24 hours at 37 ^{o}C. After 24 hours, cell survival percentages were noted for each well. All wells were replenished with fresh media containing puromycin and incubated for another 24 hours. After 48 total hours of puromycin selection, the lowest concentration of puromycin that resulted in 100% cell death was noted and used for puromycin selection of CRISPR mutants. Ideal puromycin concentration for C2C12 cells was discovered to be 2µg/mL while DE2.3 cells required a concentration of 3µg/mL for 100% cell death within 48 hours. Transfections were carried out as described earlier for *Chkb*Δ(promoter-CpG) deletions as well as negative control transfections for both cell lines using only the empty pspCas9(BB) vector and GFP. After 24 hours, the media was changed with fresh complete DMEM (10%FBS 1% Pen/strep) supplemented with the predetermined concentrations of puromycin. Cells were placed into 37 ^{o}C incubator for 24 hours; after 24 hours the percentage of surviving cells were noted, and media was replenished with puromycin, and cells were placed back into the incubator for an additional 24 hours. After a total of 48 hours of puromycin selection, media was removed, and wells were washed with 1x PBS and replenished with regular complete DMEM growth media (10% FBS and 1% pen/strep) and cells were allowed to recuperate overnight before limiting dilutions. Surviving cells were plated at a concentration of 1 cell per well in 96 well plates to isolate monoclonal cell populations. Cells were passaged as described previously and genomic DNA was extracted for PCR screening.

FAO Respiration (Oroboros O2K)

Mitochondria respiration assay was performed in a high-resolution respirometer (O2K, Oroboros) to determine the steady-state rate of oxygen consumption (JO2, i.e., e– flow through electron transport system, ETS). Intact cells at empirically determined concentrations (0.5 or 1 million cells per ml) in buffer Z were loaded into the O2K chamber (1ml) with temperature setting at 37°C and continuous stirring. After loading to O2K chamber (0.5 ml, 37°C, continuous stirring). After baseline respiration was established, cells were energized with various mitochondrial substrate combinations along with ADP. Malate (1mM) was added to maintain the tricarboxylic acid cycle, followed by ADP (0.5mM) to stimulate respiration. Substrates were added in the following concentrations: palmitoyl-CoA (20 uM), l-carnitine HCl (5mM), palmitoyl-carnitine (20μM). PMGSO: pyruvate (5mM), malate (1mM), glutamate (5mM), succinate (5mM), octanoyl-carnitine (250μM).

Table 3. Primer sequences for CRISPR genotyping analysis by qPCR.

	Top strand	Bottom strand
SgRNA-1	AAGCTACGAGTTTCGGTGTA	TACACCGAAACTCGTAGCTT
SgRNA-2	TAGGCCAATCCCGAGCGCCG	CGGCGCTCGGGATTGGCCTA
SgRNA-3	CGTGGTTCGGTAGTGAGCAT	ATGCTCACTACCGAACCACG

Table 4. Primer sequences for CRISPR genotyping analysis by qPCR.

Amplicon	Primer Sequences	
g1	CCCACCTGCACGTTTATCTG	CACTGAGAGCTGGAGAGAGG
g2	TCGGATCCCATTACATACAGACT	GCACTGATCCCTCTTCCAGA
g3	CTCTTCAGCCCCATCCCTAG	GAAAGGGCCTCAAAGCTACG
g4	TTAACGTCTCCCTCTTCCCG	TAAGAATCAGTCCGGCGTGT
g5	CCACAGGCTCGGCTAATCTA	ATGGACGATGGTAGTTTGCG
g6	TAAACCTAACGGAGCCCACG	CATCCTGCAAACCGTCCTTG
g7	AGGTGCTGCTACGACTCTAC	TTGCTCCAGATGTCCGAACT
g8	ATTCCAATCTCCCCAGTCCC	CCTTCTGAACCTCTGCCAGA
g9	CTCCTGGTGACCTTTTCCCT	CACAAGGTTGCTGGAAGGTC
g10	GGCACACTAAGCACCTTCTG	CTCAGAGCCTCCCGACTAAG
g11	CTCAAGTCATGGTGGGCAAC	TCTTCCCACCAGTCACTCAC

Table 5. Primer sequences for Endpoint PCR.

Amplicon	Primer Sequences	
E1	CTCTTCAGCCCCATCCCTAG	CATCCTGCAAACCGTCCTTG
E2	TAACCGCTTAGCCACCTCTT	ACCAAGGAGTCTACACCCTG
E3	TTAACGTCTCCCTCTTCCCG	CATCCTGCAAACCGTCCTTG
E4	CCCACCTGCACGTTTATCTG	ACCAAGGAGTCTACACCCTG

CHAPTER IV: SIGNIFICANCE AND FUTURE DIRECTIONS

This project provided more needed information to decipher the molecular pathway required for BAT emergence in placental mammals. The peculiar, conjoined nature of *Chkb* and *Cpt1b* may have been an important arrangement that allowed for elevated expression of *Cpt1b* through a more constitutively active *Chkb* promoter. The *Chkb-Cpt1b* apposition is unique and conserved in eutherians. Furthermore, the transcriptional and regulatory linkage allowed for co-expression of these genes, which are functionally connected, at higher levels in cells that utilize higher rates of FAO (skeletal muscle, heart, and BAT). These findings support a model in which the genomic convergence of syntenic *CHKB* and *CPT1B* genes through an intergenic deletion resulted in their coordinated expression as a conjoined gene in eutherians and provided for the novel attainment of elevated mitochondrial FAO in an extant cold-induced multilocular adipocyte lineage, contributing to the evolution of thermogenic BAT in eutherian mammals. We also found evidence for *Chkb-Cp1b* synteny being constrained by a functional UCP1 (figure 8) and loss of UCP1 leading to relaxation of constraints on *Chkb-Cpt1b* synteny that are evident in the Saurian clade (figure 8b).

It's important to also mention the existence of a skeletal muscle based non-shivering thermogenesis mechanism that seems to predate BAT-based NST in mammalian evolution. Briefly, calcium slippage from the sarcoendoplasmic reticulum ATPase (SERCA) results in heat production through continued ATP hydrolysis, meanwhile myofibril contraction does not take place. This mechanism can occur naturally as described and may be present to some extent in some fish and reptiles. However, a mammalian-specific muscle NST may also be present that involves sarcolipin (SLN) acting as a regulator and uncoupling the SERCA pump to cause futile calcium cycling and heat production from ATP hydrolysis. Anecdotal evidence in hibernating echidnas, of the monotreme clade, indicated that the organism rewarmed in its hibernaculum

from 13^{o}C to 18^{o}C without any noticeable shivering or muscular movements [92]. This means that perhaps muscle NST may have predated classical UCP1-dependent BAT thermogenesis and may have been the predominant NST program in metatherians such as marsupials and monotremes. The question arises though, if muscle NST was existent in mammals, why was there a need for BAT NST? A few proposed explanations suggest that muscle NST may not have been sufficient to help defend body temperature from cold exposure at birth, or for rapid arousal from torpor and furthermore, a possible disturbance in locomotion from muscle NST which would've been a costly trade-off for survival against predation [91, 93-95].

Additionally, this work highlights the ability of CpG islands to act as promoters, even being located within the 5' coding region of the gene, without any functional TATA boxes. We found significant enrichment of transcriptional initiation along with H3K4me3 and AcH3 chromatin modifications specifically localized to the *Chkb* CpG island in C2C12 and DE cells, primary myocytes cultured from humans and mice, as well as in excised murine skeletal muscle, heart, and BAT tissue. The expression of *Cpt1b* is higher in mouse skeletal muscle, heart, and BAT tissue compared to cultured primary and transformed cells, however transcriptional initiation is still predominantly localized to the *Chkb* CpG island suggesting that the dependence of *Cpt1b* expression being coupled to *Chkb* is not merely a feature of culture conditions rather it is present at every level of the physiological system.

Although we have presented the scenario where both genes are transcribed in concert, transcription of *Chkb* is not necessarily required to transcribe *Cpt1b*, instead, just the mere maintenance of an open chromatin domain encompassing both genes may be the only requirement for increased *Cpt1b* expression. The presence of a detectable read through transcript may also play a regulatory role (lncRNA) in maintaining this open chromatin state to allow for RNAPol II processing through the entire *Chkb- Cpt1b* locus and RNA Pol pausing in the intergenic region to transcribe *Cpt1b* as a separate mRNA product.

We have also discovered the presence of a CpG island within the intergenic sequence in the distal promoter region of the opossum *cpt1b* gene which serves as an independent site of transcriptional initiation and histone modifications for *cpt1b* unlinked from the *CHKB* activity. Furthermore, expression data for the *CHKB* and *CPT1B* genes in various opossum tissues suggest that the CHKB gene is associated with a WAT specific enhancer while *CPT1B* is more highly expressed in skeletal muscle. The intergenic deletion and specifically the deletion of the *Cpt1b*-specific CpG island may have resulted in novel expression of *CPT1B* in the newly acquired BAT that has the characteristics of WAT and skeletal muscle.

Using C2C12 skeletal muscle and De 2.3 brown fat murine transformed cell lines, respectively, had obvious functional limitations, however we discovered that CRISPR-mediated deletion of the *Chkb* promoter and CpG island resulted in a drastic reduction in not only *Chkb* but also *Cpt1b* transcript levels. Furthermore, the deletion also resulted in a loss of histone modifications and RNA Pol II initiation and elongation that spans the entire *Chkb-Cpt1b* chromatin domain. These cell models were informative on the transcriptional and regulatory mechanism of *Chkb-Cpt1b* especially considering we are only in the beginning stage of this project. While we did not discover strong morphological differences that would be expected from the loss of *Chkb*, we can attribute this to the compensatory action of *Chka* which is still highly expressed in mutants and WT to contribute to the PC pool. Additionally, we even detected an increased expression of Cpt1a with the knockout of Cpt1b. It will be more physiologically significant to induce the same CRISPR-mediated deletion in a transgenic mouse model because O2K respiration assays indicated that the rate of FAO was already very low in C2C12 and DE 2.3 WT cells which is not surprising considering these cells have been maintained in high glucose (25mM) media suggesting a shift in substrate preference for glucose over fatty acyl CoAs. While neither group of cells (mutants or WT) showed significant change in oxygen consumption when supplied with palmitoylCoA or palmitoylcarnitine, we did observe a

diminished change in oxygen flux in C2C12 mutant 1 compared to wildtype in response to PMGSO addition, indicating that overall mitochondrial function is perturbed in the mutants with the *Chkb*Δ(promoter-CGI) deletion.

Though still in its infancy, we hope that further characterization of the *Chkb-Cpt1b* gene locus could be used as a therapeutic target to combat obesity and metabolic disorders. Of course, as obesity results from increased triglyceride storage leading to lipotoxicity and leakage into the plasma, the formation of more thermogenic adipocytes to ameliorate the positive lipid balance via increased *Chkb-Cpt1b* expression would be a way to amplify and take advantage of the natural framework already present in the body.

This dissertation successfully addresses the emergence of BAT, a formative event in the evolution of placental mammals with a novel focus on the evolution of the *CHKB-CPT1B* gene locus beyond the occurrence of functional UCP1. The elevated capacity for homeothermy from BAT provided significant advantages including exploitation of new ecological niches, increased maternal energy investment in reproduction, and support for smaller-bodied mammalian survival. Aside from the BAT based homeothermy, the apposition of *CHKB* and *CPT1B* resulted in an elevated basal metabolic rate which includes skeletal muscle FAO. Additionally, this project explored the conjoined gene phenomenon and addressed how changes in regulatory sequences can influence the evolution of phenotype. Deletion of the intergenic region which likely included a *CPT1B*-specific CpG island enhancer resulted in syntenic genes and a novel transcription regulatory network dominated by the upstream *CHKB* gene.

REFERENCES

[1] Kumar S, Hedges SB. A molecular timescale for vertebrate evolution. Nature. 1998;392:917-20.

[2] Mezentseva NV, Kumaratilake JS, Newman SA. The brown adipocyte differentiation pathway in birds: an evolutionary road not taken. BMC Biol. 2008;6:17.

[3] Sorensen TIA, Martinez AR, Jorgensen TSH. Epidemiology of Obesity. Handb Exp Pharmacol. 2022;274:3-27.

[4] Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract. 2022;183:109119.

[5] van Weeghel M, Abdurrachim D, Nederlof R, Argmann CA, Houtkooper RH, Hagen J, et al. Increased cardiac fatty acid oxidation in a mouse model with decreased malonyl-CoA sensitivity of CPT1B. Cardiovasc Res. 2018;114:1324-34.

[6] Lopez-Vinas E, Bentebibel A, Gurunathan C, Morillas M, de Arriaga D, Serra D, et al. Definition by functional and structural analysis of two malonyl-CoA sites in carnitine palmitoyltransferase 1A. J Biol Chem. 2007;282:18212-24.

[7] McGarry JD, Brown NF. The mitochondrial carnitine palmitoyltransferase system. From concept to molecular analysis. Eur J Biochem. 1997;244:1-14.

[8] McGarry JD, Sen A, Esser V, Woeltje KF, Weis B, Foster DW. New insights into the mitochondrial carnitine palmitoyltransferase enzyme system. Biochimie. 1991;73:77-84.

[9] Steffen ML, Harrison WR, Elder FF, Cook GA, Park EA. Expression of the rat liver carnitine palmitoyltransferase I (CPT-Ialpha) gene is regulated by Sp1 and nuclear factor Y: chromosomal localization and promoter characterization. Biochem J. 1999;340 (Pt 2):425-32.

[10] Bruce CR, Hoy AJ, Turner N, Watt MJ, Allen TL, Carpenter K, et al. Overexpression of carnitine palmitoyltransferase-1 in skeletal muscle is sufficient to enhance fatty acid oxidation and improve high-fat diet-induced insulin resistance. Diabetes. 2009;58:550-8.

[11] Kim T, Moore JF, Sharer JD, Yang K, Wood PA, Yang Q. Carnitine Palmitoyltransferase 1b Deficient Mice Develop Severe Insulin Resistance After Prolonged High Fat Diet Feeding. J Diabetes Metab. 2014;5.

[12] Ji S, You Y, Kerner J, Hoppel CL, Schoeb TR, Chick WS, et al. Homozygous carnitine palmitoyltransferase 1b (muscle isoform) deficiency is lethal in the mouse. Mol Genet Metab. 2008;93:314-22.

[13] Ricci C, Marzocchi C, Riolo G, Ciuoli C, Benenati N, Bufano A, et al. The impact of CPT1B rs470117, LEPR rs1137101 and BDNF rs6265 polymorphisms on the risk of developing obesity in an Italian population. Obes Res Clin Pract. 2021;15:327-33.

[14] Maples JM, Brault JJ, Shewchuk BM, Witczak CA, Zou K, Rowland N, et al. Lipid exposure elicits differential responses in gene expression and DNA methylation in primary human skeletal muscle cells from severely obese women. Physiol Genomics. 2015;47:139-46.

[15] Maples JM, Brault JJ, Witczak CA, Park S, Hubal MJ, Weber TM, et al. Differential epigenetic and transcriptional response of the skeletal muscle carnitine palmitoyltransferase 1B (CPT1B) gene to lipid exposure with obesity. Am J Physiol Endocrinol Metab. 2015;309:E345-56.

[16] Aoyama C, Liao H, Ishidate K. Structure and function of choline kinase isoforms in mammalian cells. Prog Lipid Res. 2004;43:266-81.

[17] Aoyama C, Yamazaki N, Terada H, Ishidate K. Structure and characterization of the genes for murine choline/ethanolamine kinase isozymes alpha and beta. J Lipid Res. 2000;41:452-64.

[18] Patel BV, Yao F, Howenstine A, Takenaka R, Hyatt JA, Sears KE, et al. Emergent Coordination of the CHKB and CPT1B Genes in Eutherian Mammals: Implications for the Origin of Brown Adipose Tissue. J Mol Biol. 2020;432:6127-45.

[19] Wittenberg J, Kornberg A. Choline phosphokinase. J Biol Chem. 1953;202:431-44.

[20] Ishidate K, Nakagomi K, Nakazawa Y. Complete purification of choline kinase from rat kidney and preparation of rabbit antibody against rat kidney choline kinase. J Biol Chem. 1984;259:14706-10.

[21] Hosaka K, Kodaki T, Yamashita S. Cloning and characterization of the yeast CKI gene encoding choline kinase and its expression in Escherichia coli. J Biol Chem. 1989;264:2053-9.

[22] Porter TJ, Kent C. Purification and characterization of choline/ethanolamine kinase from rat liver. J Biol Chem. 1990;265:414-22.

[23] Hosaka K, Nikawa J, Kodaki T, Yamashita S. A dominant mutation that alters the regulation of INO1 expression in Saccharomyces cerevisiae. J Biochem. 1992;111:352-8.

[24] Monks DE, Goode JH, Dewey RE. Characterization of soybean choline kinase cDNAs and their expression in yeast and Escherichia coli. Plant Physiol. 1996;110:1197-205.

[25] Aoyama C, Nakashima K, Ishidate K. Molecular cloning of mouse choline kinase and choline/ethanolamine kinase: their sequence comparison to the respective rat homologs. Biochim Biophys Acta. 1998;1393:179-85.

[26] Aoyama C, Nakashima K, Matsui M, Ishidate K. Complementary DNA sequence for a 42 kDa rat kidney choline/ethanolamine kinase. Biochim Biophys Acta. 1998;1390:1-7.

[27] Kim KH, Carman GM. Phosphorylation and regulation of choline kinase from Saccharomyces cerevisiae by protein kinase A. J Biol Chem. 1999;274:9531-8.

[28] Wu G, Sher RB, Cox GA, Vance DE. Differential expression of choline kinase isoforms in skeletal muscle explains the phenotypic variability in the rostrocaudal muscular dystrophy mouse. Biochim Biophys Acta. 2010;1801:446-54.

[29] Wu G, Aoyama C, Young SG, Vance DE. Early embryonic lethality caused by disruption of the gene for choline kinase alpha, the first enzyme in phosphatidylcholine biosynthesis. J Biol Chem. 2008;283:1456-62.

[30] Chan SH, Ho RS, Khong PL, Chung BH, Tsang MH, Yu MH, et al. Megaconial congenital muscular dystrophy: Same novel homozygous mutation in CHKB gene in two unrelated Chinese patients. Neuromuscul Disord. 2020;30:47-53.

[31] Sher RB, Aoyama C, Huebsch KA, Ji S, Kerner J, Yang Y, et al. A rostrocaudal muscular dystrophy caused by a defect in choline kinase beta, the first enzyme in phosphatidylcholine biosynthesis. J Biol Chem. 2006;281:4938-48.

[32] Mitsuhashi S, Hatakeyama H, Karahashi M, Koumura T, Nonaka I, Hayashi YK, et al. Muscle choline kinase beta defect causes mitochondrial dysfunction and increased mitophagy. Hum Mol Genet. 2011;20:3841-51.

[33] Sayed-Zahid AA, Sher RB, Sukoff Rizzo SJ, Anderson LC, Patenaude KE, Cox GA. Functional rescue in a mouse model of congenital muscular dystrophy with megaconial myopathy. Hum Mol Genet. 2019;28:2635-47.

[34] Nishino I, Kobayashi O, Goto Y, Kurihara M, Kumagai K, Fujita T, et al. A new congenital muscular dystrophy with mitochondrial structural abnormalities. Muscle Nerve. 1998;21:40-7.

[35] Vance JE, Vance DE. Phospholipid biosynthesis in mammalian cells. Biochem Cell Biol. 2004;82:113-28.

[36] Mitsuhashi S, Ohkuma A, Talim B, Karahashi M, Koumura T, Aoyama C, et al. A congenital muscular dystrophy with mitochondrial structural abnormalities caused by defective de novo phosphatidylcholine biosynthesis. Am J Hum Genet. 2011;88:845-51.

[37] Quinlivan R, Mitsuahashi S, Sewry C, Cirak S, Aoyama C, Mooore D, et al. Muscular dystrophy with large mitochondria associated with mutations in the CHKB gene in three British patients: extending the clinical and pathological phenotype. Neuromuscul Disord. 2013;23:549-56.

[38] Mitsuhashi S, Nishino I. Megaconial congenital muscular dystrophy due to loss-of-function mutations in choline kinase beta. Curr Opin Neurol. 2013;26:536-43.

[39] Castro-Gago M, Dacruz-Alvarez D, Pintos-Martinez E, Beiras-Iglesias A, Delmiro A, Arenas J, et al. Exome sequencing identifies a CHKB mutation in Spanish patient with megaconial congenital muscular dystrophy and mtDNA depletion. Eur J Paediatr Neurol. 2014;18:796-800.

[40] Cabrera-Serrano M, Junckerstorff RC, Atkinson V, Sivadorai P, Allcock RJ, Lamont P, et al. Novel CHKB mutation expands the megaconial muscular dystrophy phenotype. Muscle Nerve. 2015;51:140-3.

[41] Marchet S, Invernizzi F, Blasevich F, Bruno V, Dusi S, Venco P, et al. Alteration of mitochondrial membrane inner potential in three Italian patients with megaconial congenital muscular dystrophy carrying new mutations in CHKB gene. Mitochondrion. 2019;47:24-9.

[42] Bardhan M, Polavarapu K, Bevinahalli NN, Preethish-Kumar V, Anjanappa RM, Arunachal G, et al. Correction: Megaconial congenital muscular dystrophy secondary to novel CHKB mutations resemble atypical Rett syndrome. J Hum Genet. 2021;66:841.

[43] Tavasoli M, Lahire S, Sokolenko S, Novorolsky R, Reid SA, Lefsay A, et al. Mechanism of action and therapeutic route for a muscular dystrophy caused by a genetic defect in lipid metabolism. Nat Commun. 2022;13:1559.

[44] Magri F, Antognozzi S, Ripolone M, Zanotti S, Napoli L, Ciscato P, et al. Megaconial congenital muscular dystrophy due to novel CHKB variants: a case report and literature review. Skelet Muscle. 2022;12:23.

[45] Prakash T, Sharma VK, Adati N, Ozawa R, Kumar N, Nishida Y, et al. Expression of conjoined genes: another mechanism for gene regulation in eukaryotes. PLoS One. 2010;5:e13284.

[46] Kim DS, Kim DW, Kim MY, Nam SH, Choi SH, Kim RN, et al. CACG: a database for comparative analysis of conjoined genes. Genomics. 2012;100:14-7.

[47] Kim RN, Kim A, Choi SH, Kim DS, Nam SH, Kim DW, et al. Novel mechanism of conjoined gene formation in the human genome. Funct Integr Genomics. 2012;12:45-61.

[48] Miyagawa T, Honda M, Kawashima M, Shimada M, Tanaka S, Honda Y, et al. Polymorphism located between CPT1B and CHKB, and HLA-DRB1*1501-DQB1*0602 haplotype confer susceptibility to CNS hypersomnias (essential hypersomnia). PLoS One. 2009;4:e5394.

[49] Miyagawa T, Kawashima M, Nishida N, Ohashi J, Kimura R, Fujimoto A, et al. Variant between CPT1B and CHKB associated with susceptibility to narcolepsy. Nat Genet. 2008;40:1324-8.

[50] Lidell ME. Brown Adipose Tissue in Human Infants. Handb Exp Pharmacol. 2019;251:107-23.

[51] Hany TF, Gharehpapagh E, Kamel EM, Buck A, Himms-Hagen J, von Schulthess GK. Brown adipose tissue: a factor to consider in symmetrical tracer uptake in the neck and upper chest region. Eur J Nucl Med Mol Imaging. 2002;29:1393-8.

[52] Cohade C, Osman M, Pannu HK, Wahl RL. Uptake in supraclavicular area fat ("USA-Fat"): description on 18F-FDG PET/CT. J Nucl Med. 2003;44:170-6.

[53] van Marken Lichtenbelt WD, Vanhommerig JW, Smulders NM, Drossaerts JM, Kemerink GJ, Bouvy ND, et al. Cold-activated brown adipose tissue in healthy men. N Engl J Med. 2009;360:1500-8.

[54] Saito M, Okamatsu-Ogura Y, Matsushita M, Watanabe K, Yoneshiro T, Nio-Kobayashi J, et al. High incidence of metabolically active brown adipose tissue in healthy adult humans: effects of cold exposure and adiposity. Diabetes. 2009;58:1526-31.

[55] Stanford KI, Middelbeek RJ, Townsend KL, An D, Nygaard EB, Hitchcox KM, et al. Brown adipose tissue regulates glucose homeostasis and insulin sensitivity. J Clin Invest. 2013;123:215-23.

[56] Gunawardana SC, Piston DW. Insulin-independent reversal of type 1 diabetes in nonobese diabetic mice with brown adipose tissue transplant. Am J Physiol Endocrinol Metab. 2015;308:E1043-55.

[57] Gunawardana SC, Piston DW. Reversal of type 1 diabetes in mice by brown adipose tissue transplant. Diabetes. 2012;61:674-82.

[58] Seale P, Bjork B, Yang W, Kajimura S, Chin S, Kuang S, et al. PRDM16 controls a brown fat/skeletal muscle switch. Nature. 2008;454:961-7.

[59] Shinde AB, Song A, Wang QA. Brown Adipose Tissue Heterogeneity, Energy Metabolism, and Beyond. Front Endocrinol (Lausanne). 2021;12:651763.

[60] Sanchez-Gurmaches J, Guertin DA. Adipocytes arise from multiple lineages that are heterogeneously and dynamically distributed.

[61] Torres N, Vargas-Castillo AE, Tovar AR. Adipose Tissue: White Adipose Tissue Structure and Function. In: Caballero B, Finglas PM, Toldrá F, editors. Encyclopedia of Food and Health. Oxford: Academic Press; 2016. p. 35-42.

[62] Vargas-Castillo A, Torres N, Tovar AR. Chapter 12 - Endocrine Regulation of Brown and Beige Adipose Tissue. In: Ulloa-Aguirre A, Tao Y-X, editors. Cellular Endocrinology in Health and Disease (Second Edition). Boston: Academic Press; 2021. p. 247-59.

[63] Harms MJ, Ishibashi J, Wang W, Lim HW, Goyama S, Sato T, et al. Prdm16 is required for the maintenance of brown adipocyte identity and function in adult mice. Cell Metab. 2014;19:593-604.

[64] Harms MJ, Lim HW, Ho Y, Shapira SN, Ishibashi J, Rajakumari S, et al. PRDM16 binds MED1 and controls chromatin architecture to determine a brown fat transcriptional program. Genes Dev. 2015;29:298-307.

[65] Feldmann HM, Golozoubova V, Cannon B, Nedergaard J. UCP1 ablation induces obesity and abolishes diet-induced thermogenesis in mice exempt from thermal stress by living at thermoneutrality. Cell Metab. 2009;9:203-9.

[66] Tseng YH, Cypess AM, Kahn CR. Cellular bioenergetics as a target for obesity therapy. Nat Rev Drug Discov. 2010;9:465-82.

[67] Lowell BB, Spiegelman BM. Towards a molecular understanding of adaptive thermogenesis. Nature. 2000;404:652-60.

[68] Lu Y, Fujioka H, Joshi D, Li Q, Sangwung P, Hsieh P, et al. Mitophagy is required for brown adipose tissue mitochondrial homeostasis during cold challenge. Sci Rep. 2018;8:8251.

[69] Krahmer N, Guo Y, Wilfling F, Hilger M, Lingrell S, Heger K, et al. Phosphatidylcholine synthesis for lipid droplet expansion is mediated by localized activation of CTP:phosphocholine cytidylyltransferase. Cell Metab. 2011;14:504-15.

[70] Smith RM, Rill RL. Mobile histone tails in nucleosomes. Assignments of mobile segments and investigations of their role in chromatin folding. J Biol Chem. 1989;264:10574-81.

[71] Arya G, Schlick T. Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. Proc Natl Acad Sci U S A. 2006;103:16236-41.

[72] Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21:381-95.

[73] Vaughan RM, Kupai A, Rothbart SB. Chromatin Regulation through Ubiquitin and Ubiquitin-like Histone Modifications. Trends Biochem Sci. 2021;46:258-69.

[74] Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38:23-38.

[75] Yen RW, Vertino PM, Nelkin BD, Yu JJ, el-Deiry W, Cumaraswamy A, et al. Isolation and characterization of the cDNA encoding human DNA methyltransferase. Nucleic Acids Res. 1992;20:2287-91.

[76] Xie S, Wang Z, Okano M, Nogami M, Li Y, He WW, et al. Cloning, expression and chromosome locations of the human DNMT3 gene family. Gene. 1999;236:87-95.

[77] Goto K, Numata M, Komura JI, Ono T, Bestor TH, Kondo H. Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. Differentiation. 1994;56:39-44.

[78] Leonhardt H, Page AW, Weier HU, Bestor TH. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell. 1992;71:865-73.

[79] Hermann A, Goyal R, Jeltsch A. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. J Biol Chem. 2004;279:48350-9.

[80] Janitz K, Janitz M. Chapter 12 - Assessing Epigenetic Information. In: Tollefsbol T, editor. Handbook of Epigenetics. San Diego: Academic Press; 2011. p. 173-81.

[81] Duncan BK, Miller JH. Mutagenic deamination of cytosine residues in DNA. Nature. 1980;287:560-1.

[82] Latchman DS. CHAPTER 5 - ACTIVATION OF GENE EXPRESSION BY TRANSCRIPTION FACTORS. In: Latchman DS, editor. Eukaryotic Transcription Factors (Fifth edition). New York: Academic Press; 2008. p. 161-228.

[83] Hayward JS, Lisson PA. Evolution of brown fat: its absence in marsupials and monotremes. Canadian Journal of Zoology. 1992;70:171-9.

[84] Jastroch M, Wuertz S, Kloas W, Klingenspor M. Uncoupling protein 1 in fish uncovers an ancient evolutionary history of mammalian nonshivering thermogenesis. Physiol Genomics. 2005;22:150-6.

[85] Jastroch M, Buckingham JA, Helwig M, Klingenspor M, Brand MD. Functional characterisation of UCP1 in the common carp: uncoupling activity in liver mitochondria and cold-induced expression in the brain. J Comp Physiol B. 2007;177:743-52.

[86] Jastroch M, Withers KW, Taudien S, Frappell PB, Helwig M, Fromme T, et al. Marsupial uncoupling protein 1 sheds light on the evolution of mammalian nonshivering thermogenesis. Physiol Genomics. 2008;32:161-9.

[87] Hughes DA, Jastroch M, Stoneking M, Klingenspor M. Molecular evolution of UCP1 and the evolutionary history of mammalian non-shivering thermogenesis. BMC Evol Biol. 2009;9:4.

[88] Jastroch M, Oelkrug R, Keipert S. Insights into brown adipose tissue evolution and function from non-model organisms. J Exp Biol. 2018;221.

[89] Ricquier D, Kader JC. Mitochondrial protein alteration in active brown fat: a soidum dodecyl sulfate-polyacrylamide gel electrophoretic study. Biochem Biophys Res Commun. 1976;73:577-83.

[90] Heaton GM, Wagenvoord RJ, Kemp A, Jr., Nicholls DG. Brown-adipose-tissue mitochondria: photoaffinity labelling of the regulatory site of energy dissipation. Eur J Biochem. 1978;82:515-21.

[91] Cannon B, Nedergaard J. Brown adipose tissue: function and physiological significance. Physiol Rev. 2004;84:277-359.

[92] Grigg G, Augee ML, Beard LA. Thermal Relations Of Free-Living Echidnas During Activity And In Hibernation In A Cold Climate. 1992.

[93] Oelkrug R, Heldmaier G, Meyer CW. Torpor patterns, arousal rates, and temporal organization of torpor entry in wildtype and UCP1-ablated mice. J Comp Physiol B. 2011;181:137-45.

[94] Oelkrug R, Goetze N, Exner C, Lee Y, Ganjam GK, Kutschke M, et al. Brown fat in a protoendothermic mammal fuels eutherian evolution. Nat Commun. 2013;4:2140.

[95] Rowland LA, Bal NC, Periasamy M. The role of skeletal-muscle-based thermogenic mechanisms in vertebrate endothermy. Biol Rev Camb Philos Soc. 2015;90:1279-97.

[96] Nicol SC, Andersen NA, Arnold W, Ruf T. Rewarming rates of two large hibernators: Comparison of a monotreme and a eutherian. Journal of Thermal Biology. 2009;34:155-9.

[97] Brandt JM, Djouadi F, Kelly DP. Fatty acids activate transcription of the muscle carnitine palmitoyltransferase I gene in cardiac myocytes via the peroxisome proliferator-activated receptor alpha. J Biol Chem. 1998;273:23786-92.

[98] Chamouton J, Latruffe N. PPARalpha/HNF4alpha interplay on diversified responsive elements. Relevance in the regulation of liver peroxisomal fatty acid catabolism. Curr Drug Metab. 2012;13:1436-53.

[99] Gilde AJ, van der Lee KA, Willemsen PH, Chinetti G, van der Leij FR, van der Vusse GJ, et al. Peroxisome proliferator-activated receptor (PPAR) alpha and PPARbeta/delta, but not PPARgamma, modulate the expression of genes involved in cardiac lipid metabolism. Circ Res. 2003;92:518-24.

[100] Martinez-Jimenez CP, Kyrmizi I, Cardot P, Gonzalez FJ, Talianidis I. Hepatocyte nuclear factor 4alpha coordinates a transcription factor network regulating hepatic fatty acid metabolism. Mol Cell Biol. 2010;30:565-77.

[101] Muoio DM, MacLean PS, Lang DB, Li S, Houmard JA, Way JM, et al. Fatty acid homeostasis and induction of lipid regulatory genes in skeletal muscles of peroxisome proliferator-activated receptor (PPAR) alpha knock-out mice. Evidence for compensatory regulation by PPAR delta. J Biol Chem. 2002;277:26089-97.

[102] Yu GS, Lu YC, Gulick T. Co-regulation of tissue-specific alternative human carnitine palmitoyltransferase Ibeta gene promoters by fatty acid enzyme substrate. J Biol Chem. 1998;273:32901-9.

[103] Moore ML, Park EA, McMillin JB. Upstream stimulatory factor represses the induction of carnitine palmitoyltransferase-Ibeta expression by PGC-1. J Biol Chem. 2003;278:17263-8.

[104] Sharma V, Dhillon P, Parsons H, Allard MF, McNeill JH. Metoprolol represses PGC1alpha-mediated carnitine palmitoyltransferase-1B expression in the diabetic heart. Eur J Pharmacol. 2009;607:156-66.

[105] Baldan A, Relat J, Marrero PF, Haro D. Functional interaction between peroxisome proliferator-activated receptors-alpha and Mef-2C on human carnitine palmitoyltransferase 1beta (CPT1beta) gene activation. Nucleic Acids Res. 2004;32:4742-9.

[106] Yuan H, Niu Y, Liu X, Fu L. Exercise increases the binding of MEF2A to the Cpt1b promoter in mouse skeletal muscle. Acta Physiol (Oxf). 2014;212:283-92.

[107] van der Leij FR, Cox KB, Jackson VN, Huijkman NCA, Bartelds B, Kuipers JRG, et al. Structural and Functional Genomics of the CPT1B Gene for Muscle-type Carnitine Palmitoyltransferase I in Mammals. Journal of Biological Chemistry. 2002;277:26994-7005.

[108] Moore ML, Wang GL, Belaguli NS, Schwartz RJ, McMillin JB. GATA-4 and serum response factor regulate transcription of the muscle-specific carnitine palmitoyltransferase I beta in rat heart. J Biol Chem. 2001;276:1026-33.

[109] Wu G, Vance DE. Choline kinase and its function. Biochem Cell Biol. 2010;88:559-64.

[110] Yamazaki N, Yamanaka Y, Hashimoto Y, Hiramatsu T, Shinohara Y, Terada H. The gene encoding muscle-type carnitine palmitoyltransferase I: comparison of the 5'-upstream region of human and rodent genes. J Biochem. 2003;133:523-32.

[111] Yamazaki N, Shinohara Y, Kajimoto K, Shindo M, Terada H. Novel expression of equivocal messages containing both regions of choline/ethanolamine kinase and muscle type carnitine palmitoyltransferase I. J Biol Chem. 2000;275:31739-46.

[112] Kuan CS, Yee YH, See Too WC, Few LL. Ets and GATA transcription factors play a critical role in PMA-mediated repression of the ckbeta promoter via the protein kinase C signaling pathway. PLoS One. 2014;9:e113485.

[113] Auinger A, Rubin D, Sabandal M, Helwig U, Ruther A, Schreiber S, et al. A common haplotype of carnitine palmitoyltransferase 1b is associated with the metabolic syndrome. Br J Nutr. 2013;109:810-5.

[114] Koves TR, Ussher JR, Noland RC, Slentz D, Mosedale M, Ilkayeva O, et al. Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance. Cell Metab. 2008;7:45-56.

[115] Cohen P, Spiegelman BM. Cell biology of fat storage. Mol Biol Cell. 2016;27:2523-7.

[116] Cohen P, Spiegelman BM. Brown and Beige Fat: Molecular Parts of a Thermogenic Machine. Diabetes. 2015;64:2346-51.

[117] Oelkrug R, Polymeropoulos ET, Jastroch M. Brown adipose tissue: physiological function and evolutionary significance. J Comp Physiol B. 2015;185:587-606.

[118] Cedikova M, Kripnerova M, Dvorakova J, Pitule P, Grundmanova M, Babuska V, et al. Mitochondria in White, Brown, and Beige Adipocytes. Stem Cells Int. 2016;2016:6067349.

[119] Fuller-Jackson JP, Henry BA. Adipose and skeletal muscle thermogenesis: studies from large animals. J Endocrinol. 2018;237:R99-R115.

[120] Periasamy M, Herrera JL, Reis FCG. Skeletal Muscle Thermogenesis and Its Role in Whole Body Energy Metabolism. Diabetes Metab J. 2017;41:327-36.

[121] Nowack J, Giroud S, Arnold W, Ruf T. Muscle Non-shivering Thermogenesis and Its Role in the Evolution of Endothermy. Frontiers in Physiology. 2017;8.

[122] Bal NC, Singh S, Reis FCG, Maurya SK, Pani S, Rowland LA, et al. Both brown adipose tissue and skeletal muscle thermogenesis processes are activated during mild to severe cold adaptation in mice. J Biol Chem. 2017;292:16616-25.

[123] Wu J, Bostrom P, Sparks LM, Ye L, Choi JH, Giang AH, et al. Beige adipocytes are a distinct type of thermogenic fat cell in mouse and human. Cell. 2012;150:366-76.

[124] Seale P, Conroe HM, Estall J, Kajimura S, Frontini A, Ishibashi J, et al. Prdm16 determines the thermogenic program of subcutaneous white adipose tissue in mice. J Clin Invest. 2011;121:96-105.

[125] Mzilikazi N, Jastroch M, Meyer CW, Klingenspor M. The molecular and biochemical basis of nonshivering thermogenesis in an African endemic mammal, Elephantulus myurus. Am J Physiol Regul Integr Comp Physiol. 2007;293:R2120-7.

[126] Jastroch M, Withers KW, Stoehr S, Klingenspor M. Mitochondrial proton conductance in skeletal muscle of a cold-exposed marsupial, Antechinus flavipes, is unlikely to be involved in adaptive nonshivering thermogenesis but displays increased sensitivity toward carbon-centered radicals. Physiol Biochem Zool. 2009;82:447-54.

[127] Klingenspor M, Fromme T, Hughes DA, Jr., Manzke L, Polymeropoulos E, Riemann T, et al. An ancient look at UCP1. Biochim Biophys Acta. 2008;1777:637-41.

[128] Wu G, Sher RB, Cox GA, Vance DE. Understanding the muscular dystrophy caused by deletion of choline kinase beta in mice. Biochim Biophys Acta. 2009;1791:347-56.

[129] Gutierrez Rios P, Kalra AA, Wilson JD, Tanji K, Akman HO, Area Gomez E, et al. Congenital megaconial myopathy due to a novel defect in the choline kinase Beta gene. Arch Neurol. 2012;69:657-61.

[130] Mitsuhashi S, Nishino I. Phospholipid synthetic defect and mitophagy in muscle disease. Autophagy. 2014;7:1559-61.

[131] Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002;12:996-1006.

[132] Raney BJ, Dreszer TR, Barber GP, Clawson H, Fujita PA, Wang T, et al. Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser. Bioinformatics. 2014;30:1003-5.

[133] Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes Dev. 2009;23:781-3.

[134] Kimura H. Histone modifications for human epigenome analysis. J Hum Genet. 2013;58:439-45.

[135] Lee AK, Cockburn A. Evolutionary ecology of marsupials. Cambridge Cambridgeshire ; New York: Cambridge University Press; 1985.

[136] McNab BK. Extreme measures : the ecological energetics of birds and mammals. Chicago: The University of Chicago Press; 2012.

[137] McNab BK. Behavioral and ecological factors account for variation in the mass-independent energy expenditures of endotherms. J Comp Physiol B. 2015;185:1-13.

[138] McNab BK. An analysis of the factors that influence the level and scaling of mammalian BMR. Comp Biochem Physiol A Mol Integr Physiol. 2008;151:5-28.

[139] Newman SA. Thermogenesis, muscle hyperplasia, and the origin of birds. Bioessays. 2011;33:653-6.

[140] Newman SA, Mezentseva NV, Badyaev AV. Gene loss, thermogenesis, and the origin of birds. Ann N Y Acad Sci. 2013;1289:36-47.

[141] Gaudry MJ, Jastroch M, Treberg JR, Hofreiter M, Paijmans JLA, Starrett J, et al. Inactivation of thermogenic UCP1 as a historical contingency in multiple placental mammal clades. Science Advances. 2017;3:e1602878.

[142] Barre H, Cohen-Adad F, Duchamp C, Rouanet JL. Multilocular adipocytes from muscovy ducklings differentiated in response to cold acclimation. J Physiol. 1986;375:27-38.

[143] Bartz R, Li WH, Venables B, Zehmer JK, Roth MR, Welti R, et al. Lipidomics reveals that adiposomes store ether lipids and mediate phospholipid traffic. J Lipid Res. 2007;48:837-47.

[144] Kajimura S, Spiegelman BM, Seale P. Brown and Beige Fat: Physiological Roles beyond Heat Generation. Cell Metab. 2015;22:546-59.

[145] Pan D, Fujimoto M, Lopes A, Wang YX. Twist-1 is a PPARdelta-inducible, negative-feedback regulator of PGC-1alpha in brown fat metabolism. Cell. 2009;137:73-86.

[146] Hindi L, McMillan JD, Afroze D, Hindi SM, Kumar A. Isolation, Culturing, and Differentiation of Primary Myoblasts from Skeletal Muscle of Adult Mice. Bio Protoc. 2017;7.

[147] Shahini A, Vydiam K, Choudhury D, Rajabian N, Nguyen T, Lei P, et al. Efficient and high yield isolation of myoblasts from skeletal muscle. Stem Cell Res. 2018;30:122-9.

[148] Klein J, Fasshauer M, Ito M, Lowell BB, Benito M, Kahn CR. beta(3)-adrenergic stimulation differentially inhibits insulin signaling and decreases insulin-induced glucose uptake in brown adipocytes. J Biol Chem. 1999;274:34795-802.

[149] Keyte AL, Smith KK. Basic Maintenance and Breeding of the Opossum Monodelphis domestica. CSH Protoc. 2008;2008:pdb prot5073.

[150] Rousmaniere H, Silverman R, White RA, Sasaki MM, Wilson SD, Morrison JT, et al. Husbandry of Monodelphis domestica in the study of mammalian embryogenesis. Lab Anim (NY). 2010;39:219-26.

[151] Hogan KA, Jefferson HS, Karschner VA, Shewchuk BM. Expression of Pit-1 in nonsomatotrope cell lines induces human growth hormone locus control region histone modification and hGH-N transcription. J Mol Biol. 2009;390:26-44.

[152] Nakonechnaya AO, Jefferson HS, Chen X, Shewchuk BM. Differential effects of exogenous and autocrine growth hormone on LNCaP prostate cancer cell proliferation and survival. J Cell Biochem. 2013;114:1322-35.

[153] Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8:2281-308.

[154] Pugh BF, Tjian R. Transcription from a TATA-less promoter requires a multisubunit TFIID complex.

[155] Eberl HC, Spruijt CG, Kelstrup CD, Vermeulen M, Mann M. A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. Mol Cell. 2013;49:368-78.

[156] Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell. 2007;131:58-69.

[157] van Ingen H, van Schaik FM, Wienk H, Ballering J, Rehmann H, Dechesne AC, et al. Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. Structure. 2008;16:1245-56.

[158] Patel AB, Louder RK, Greber BJ, Grunberg S, Luo J, Fang J, et al. Structure of human TFIID and mechanism of TBP loading onto promoter DNA. Science. 2018;362.

[159] van Nuland R, Schram AW, van Schaik FM, Jansen PW, Vermeulen M, Marc Timmers HT. Multivalent engagement of TFIID to nucleosomes. PLoS One. 2013;8:e73495.

[160] Buchman AR, Berg P. Comparison of intron-dependent and intron-independent gene expression. Mol Cell Biol. 1988;8:4395-405.

[161] Palmiter RD, Sandgren EP, Avarbock MR, Allen DD, Brinster RL. Heterologous introns can enhance expression of transgenes in mice. Proc Natl Acad Sci U S A. 1991;88:478-82.

[162] Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A. Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans. Genetics. 1993;135:385-404.

[163] Koziel MG, Carozzi NB, Desai N. Optimizing expression of transgenes with an emphasis on post-transcriptional events. Plant Mol Biol. 1996;32:393-405.

[164] Simpson GG, Filipowicz W. Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organisation of the spliceosomal machinery. Plant Mol Biol. 1996;32:1-41.

[165] Duncker BP, Davies PL, Walker VK. Introns boost transgene expression in Drosophila melanogaster. Mol Gen Genet. 1997;254:291-6.

[166] Juneau K, Miranda M, Hillenmeyer ME, Nislow C, Davis RW. Introns regulate RNA and protein abundance in yeast. Genetics. 2006;174:511-8.

[167] Bieberstein NI, Carrillo Oesterreich F, Straube K, Neugebauer KM. First exon length controls active chromatin signatures and transcription. Cell Rep. 2012;2:62-8.

[168] Fiszbein A, Krick KS, Begg BE, Burge CB. Exon-Mediated Activation of Transcription Starts. Cell. 2019;179:1551-65 e17.

[169] Rose AB. Intron-mediated regulation of gene expression. Curr Top Microbiol Immunol. 2008;326:277-90.

[170] Harms MJ, Ishibashi J, Wang W, Lim HW, Goyama S, Sato T, et al. Prdm16 is required for the maintenance of brown adipocyte identity and function in adult mice.