ANALYSIS OF GENOMIC IMPRINTING OF *UBE3A* IN NEURONS

A Dissertation

by

PAUL RANDOLPH HILLMAN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Scott V. Dindot
Co-Chair of Committee,	Julian Leibowitz
Committee Members,	Kayla Bayless
	Geoffrey Kapler
	Rajesh Miranda
Head of Department,	Van Wilson

May 2015

Major Subject: Medical Sciences

ABSTRACT

Angelman syndrome (AS), chromosome 15q11-q13 duplication syndrome (Dup15q), and Prader-Willi syndrome (PWS) are neurodevelopmental disorders associated with dysregulated expression of imprinted genes located within the human 15q11-13 imprinted region. Angelman syndrome is caused by loss-of-function or loss-of-expression of the maternally inherited *UBE3A* allele; Dup15q syndrome is attributed to maternally inherited copy number gains of *UBE3A*; and, paternally inherited deletions of the *SNORD116* cluster cause PWS. The *UBE3A* gene is imprinted in the brain with maternal-specific expression and biallelically expressed in all other cell types. The imprint is regulated by expression of the *UBE3A* antisense transcript (*UBE3A-AS*), which is expressed only in neurons and imprinted with paternal-specific expression. The *UBE3A-AS* represents the 3` end of a long polycistronic transcript that includes the *SNORD116* and *SNORD115* gene clusters. Thus, the genes causing AS, Dup15q, and PWS are transcriptionally linked; however, the functional significance of the neuron specific imprint is largely unknown. In this dissertation, it was hypothesized that imprinting of *UBE3A* evolved as a mechanism to negatively regulate UBE3A protein levels in neurons. This hypothesis was tested by examining allelic expression patterns and associated protein levels of the mouse 7c imprinted region, the orthologous region of human 15q11-q13.

Analyses revealed that imprinted expression of *Ube3a* in the brain resulted in elevated RNA and protein levels compared to tissues where *Ube3a* was biallelically expressed.

Likewise, *Snord116, Snord115,* and *Ube3a-AS* transcripts were highly expressed in the brain. The elevated Ube3a protein levels in the brain were due to increased maternal-allelic expression during neurogenesis concurrent with paternal-allelic suppression. Analysis of *UBE3A* expression in the opossum, a metatherian mammal lacking an orthologous imprinted region, showed that the *UBE3A* imprint did not evolve to negatively regulate UBE3A protein levels in the brain. Extensive alternative splicing of *Ube3a-AS* was detected in the brain, which generated at least two transcripts containing novel open reading frames. Novel *Ube3a* alternatively spliced transcripts were also identified in the brain. Collectively, these data reject the hypothesis that the *UBE3A* imprint evolved to negatively regulate UBE3A protein levels in the brain; instead, they suggest that the *UBE3A* imprint may allow co-expression of the *UBE3A* and *SNORD* gene cluster in neurons, which may also facilitate or regulate the expression of novel brain-specific *UBE3A* transcripts.

To my patient, beautiful, and supportive wife; my encouraging family and friends; and, my compatriots in the trenches, members of the Dindot laboratory.

AND

To all people who have been touched or affected by Angelman syndrome, chromosome 15q11-q13 duplication syndrome, or Prader-Willi syndrome.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Dindot, and my committee members, Dr. Bayless, Dr. Kapler, Dr. Leibowitz, and Dr. Miranda, for their insight, guidance, and support throughout the course of my doctoral research training.

Special thanks go to Dr. Leibowitz and Dr. Wells for serving as advisors of the MD/PhD program and for their constant encouragement and motivation through the entirety of my graduate education. Their selfless support and encouragement are immeasurable and their genuine thirst for knowledge and discovery are an inspiration.

Thanks also go to my department, faculty, and staff for their perpetual assistance. Their warm and welcoming support enriched my experience. The Texas A&M Health Science Center College of Medicine Department of Molecular and Cellular Medicine and the Texas A&M University College of Veterinary Medicine & Biomedical Sciences Department of Veterinary Pathobiology were instrumental in supporting my research pursuits.

I would like to recognize: Ms. Sally Shepard for starting my education down the right path and showing me that it is ok to be one of the smart kids; Mrs. Lesha Harenberg for cementing my love of biological sciences and her constant belief in my bright future; and Dr. Suma Datta for taking in a naïve undergraduate student and starting his research career.

Thank you to Tiana Endicott-Yazdani for her friendship, conversations, and support as a partner in crime who knows and appreciates exactly what we have gone through.

Dr. Rola Mouneimne, Dr. Clifford Stephan, Dr. Roger Smith, and Ms. Mary Sobieski provided valuable technical support and education that allowed for the completion of intricate aspects of this study.

Immense thanks go to my fellow graduate students, colleagues and lab mates: Dr. Sarah Black Christian, Dr. Ryan Doan, Cole McQueen, Jade Benjamin, Erica Downey, Dr. Katia Amstalden, Kathy Rector, and Kristy Fickey. I would also like to thank all the undergraduate trainees in the Dindot laboratory: Alessandra Jimenez, Tori Wagner, Kimberly Rabee, Cammi Gobbel, and Gregory Whitaker. Thank you all for your assistance, camaraderie, late nights, laughs, and lunches. You all contributed to this work and taught me enumerable things.

Dr. Sarah Black Christian deserves more recognition and thanks than space will allow. Her contribution to this work was unfathomable, as was her friendship, education, and reality checks. She will be a co-author on the manuscript resulting from the research contained in this dissertation. Thank you.

This work would not have been possible without my advisor, Dr. Dindot. Thank you for your enthusiasm and dedication in all aspects of this project. These attributes and your patience were an inspiration. Thank you for taking me on as your graduate student and educating me beyond the realm of bench work.

Finally, words cannot express the love and support of my family during this endeavor. My amazing wife has shown unwavering support, patience, and sacrifice throughout this entire process. The rest of my family has been full of encouragement and belief from the beginning. Your support and belief has kept me going.

Thank you everybody who has touched my life and helped me through this project with your kind words and support.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

Genomic imprinting is an epigenetic phenomenon that results in the differential expression of diploid alleles in a parent of origin specific manner [1]. Imprinted genes are functionally hemizygous, often dosage sensitive and responsible for various pathological states when their expression patterns are dysregulated [2, 3]. Genomic imprinting only occurs in therian mammals and flowering plants [4, 5]. It has also been observed in some insects, but it's unclear whether this phenomenon is a true form of genomic imprinting [6]. Numerous theories exist to explain the evolution and function of imprinted genes (e.g., conflict, host-defense, maternal time-bomb, and developmental plasticity hypotheses [see 2.1.6]); however, there is currently no unifying theory that is applicable to all, much less most, imprinted genes [7]. Overall, the functional significance of imprinting is largely unknown and seemingly contradictory to the presumed evolutionary advantage of diploidy [8].

The human 15q11-q13 region contains a cluster of conserved imprinted genes that evolved approximately 110 million years ago (**Figure 1**) [9, 10]. Most of the genes in this region are exclusively expressed from the paternal chromosome in the tissues in which they are transcribed. The imprinted genes in 15q11-q13 with paternal-specific expression include: small nucleolar ribonucleoprotein polypeptide N (*SNRPN*), SNRPN upstream reading frame (*SNURF*), necdin (*NDN*), mage-like 2 (*MAGEL2*), makorin 3

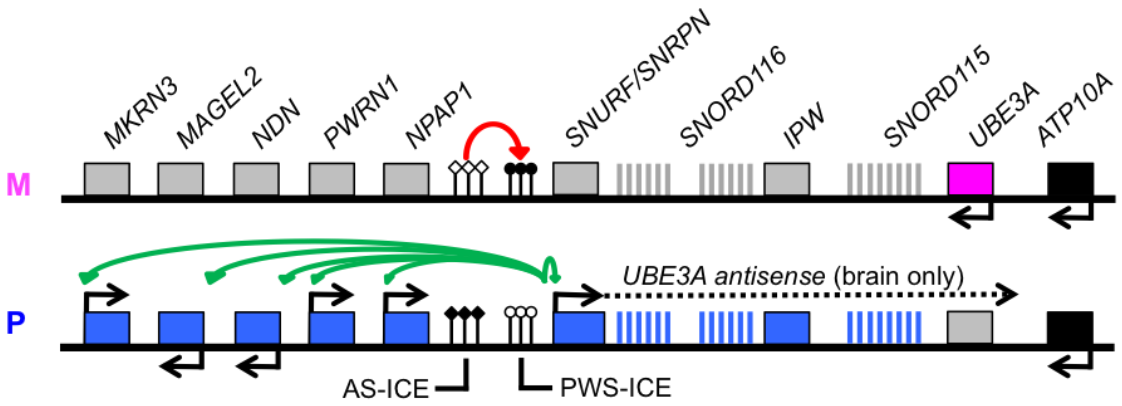


Figure 1. The 15q11-q13 imprinted region. Genomic imprinting of genes in 15q11-q13 is regulated by a partite imprinting center comprised of the two imprinting control elements: the Angelman syndrome imprinting control element (AS-ICE) and Prader-Willi syndrome imprinting control element (PWS-ICE). The AS-ICE is methylated and inactivated in sperm (filled diamond lollipops), whereas the maternal-allele is shielded from methylation by a bound germline protein complex (open diamond lollipops). Subsequently, the maternal AS-ICE negatively regulates the maternal PWS-ICE (red arrow) leading to its methylation (filled circle lollipops). The unmethylated paternal PWS-ICE (open circle lollipops) positively regulates (green arrows) the flanking paternally expressed genes (blue boxes) by an unknown mechanism. The PWS-ICE overlaps the *SNURF-SNRPN* promoter, which utilizes alternative upstream exons of *SNRPN* to gives rise to a long RNA transcript containing the *SNORD116*, *SNORD115*, *IPW*, and the *UBE3A-AS*. Expression of the *UBE3A-AS* overlaps the paternal *UBE3A* allele, which acts *in cis* to inhibit paternal-allelic expression by some unknown mechanism. Expression of the maternal allele (pink box) is unaffected by the *UBE3A-AS*. Figure legend: gray boxes indicate the inactive allele; black boxes indicate biallelically expressed genes; arrows indicate the direction of transcription (5' → 3'). Abbreviations: M = maternal allele; P = paternal allele. Schematic is not drawn to scale.

(*MKRN3*), nuclear pore associated protein 1 (*NPAP1*) genes, Prader-Willi region non-protein coding RNA 1 (*PWRN1*), and two C/D box small nucleolar RNA (snoRNA) clusters (*SNORD115* and *SNORD116) [11, 12]*. At the distal end of the region lies the ubiquitin protein ligase E3A (*UBE3A*) gene [13]. Unlike the other imprinted genes in the region, *UBE3A* is imprinted with maternal-specific expression in the brain and biallelically expressed in all other cell types [14-17]. Also unique to *UBE3A* is the presence of active post-translational histone modifications and RNA polymerase II at the inactive paternal allele in the brain [18, 19]. Consequently, the paternal allele

transcribes a 5` truncated transcript, but the function, if any, of this transcript is unknown [18, 20].

Genomic imprinting of the genes in 15q11-13 is controlled by the Prader-Willi and Angelman syndrome imprinting control elements (PWS-ICE and AS-ICE [**Figure 1**]) [21, 22]. On the maternal chromosome, the AS-ICE acts *in cis* to negatively regulate the activity of the adjacent PWS-ICE [23, 24]. On the paternal chromosome, the AS-ICE is inactive; consequently, the PWS-ICE is functional and acts *in cis* to positively regulate expression of the paternally expressed genes [11, 23, 24]. In the brain, the PWS-ICE facilitates the expression of a large polycistronic transcript containing the *SNURF/SNRPN*, *SNORD115*, *SNORD116*, and the *UBE3A* antisense (*UBE3A-AS*) genes [12, 25, 26]. The *SNORD115* and *SNORD116* gene clusters are a series of repeated snoRNA paralogs that are located within the introns of two snoRNA genes [27]. Canonical processing of the snoRNA genes releases the *SNORD115*and *SNORD116* RNAs from the host-gene transcript [12, 25, 26]. The *UBE3A-AS* represents the 3` end of the snoRNA genes and its expression across the paternal *UBE3A* allele is believed to be solely responsible for establishing the *UBE3A* imprint [18, 20, 25, 28-32]. The mechanism by which the *UBE3A-AS* silences *UBE3A* paternal-allelic expression and the functional significance of the *UBE3A* imprint in the brain are largely unknown.

Dysregulation of genes in 15q11-q13 are associated with at least three human conditions: Angelman syndrome (AS), chromosome 15q11-q13 duplication syndrome

(Dup15q), and Prader-Willi syndrome (PWS) [33-36]. Angelman syndrome is caused by mutations or epimutations that affect the expression or function of the maternally inherited *UBE3A* allele [37-39]. Angelman syndrome is characterized by developmental delay, loss of speech, ataxia, epilepsy, and a characteristic happy disposition [40]. Conversely, Dup15q syndrome is associated with copy number gains of the maternally inherited *UBE3A* allele, and it shares many of the neurological features associated with AS (e.g., autism spectrum disorder, developmental delay, intellectual disability, epilepsy, ataxia, and reduced or absent speech [41]. The genotype-phenotype correlations of AS and Dup15q suggest that *UBE3A* is a dosage sensitive gene. Prader-Willi syndrome is caused by paternally inherited deletions of 15q11-q13 that cause loss of expression of the *SNORD116* cluster [42, 43]. Although, AS, Dup15q, and PWS are clinically distinct conditions, the causative genes underlying their respective phenotypes are transcriptionally, and perhaps functionally, linked.

The *UBE3A* gene encodes an E3 ubiquitin-protein ligase that is a central member of the ubiquitin proteasome system [44-46]. The UBE3A protein possesses a conserved HECT domain (homologous to E6-AP C-terminus) that catalyzes the terminal step of the ubiquitination reaction [44]. The *UBE3A* gene expresses numerous alternatively spliced transcripts that encode at least three protein isoforms [47, 48]. The isoforms differ at the N-terminus and are each expressed in the brain; however, the precise cellular role of each isoform is unknown [44, 47, 48]. The UBE3A protein is known to ubiquitinate numerous proteins that are associated with a myriad of cellular pathways [49-54]. The

UBE3A protein also functions as a transcriptional co-activator of steroid hormone receptors, indicating numerous and diverse roles in the cell [55, 56]. The specific cellular pathways underlying the neurological deficits observed in AS and Dup15q, however, are currently unknown.

The expression and dosage levels (i.e., RNA and protein) of UBE3A are clearly important for the human brain to develop and function properly. As such, it is perplexing that *UBE3A* is specifically imprinted in neurons and biallelically expressed in all other cell types. The imprint is not a random event, and it has been evolutionarily constrained for over 100 million years [10]. Taken together, these observations indicate the imprint plays a critical role in the development and/or function of the brain; however, the functional significance of the imprint has been largely ignored. The doctoral dissertation outlined here attempts to answer some of the most pressing questions regarding the imprinting of *UBE3A* in the brain: *(i) When during neurogenesis is Ube3a imprinted?; (ii) Is Ube3a only imprinted in neurons of the central nervous system?; (iii) What effect does the imprint have on the overall expression of Ube3a in brain?; (iv) Does the Ube3a-AS have a function independent of silencing Ube3a paternal-allelic expression?* Understanding the functional significance of the imprint is important because it may provide insight into the function of Ube3a, the snoRNAs, and the *Ube3a-AS* in the brain. Furthermore, it may reveal novel pathways to treat AS, PWS, and Dup15q.

2. LITERATURE REVIEW

2.1 Genomic Imprinting

2.1.1 History of Genomic Imprinting

Diploid organisms possess two copies of their genome, one copy from each of their parents. For the majority of genes, these two copies are expressed at relatively equal proportions; however, a small subset of genes express only one of the two parental alleles in a parent of origin specific matter, this is genomic imprinting [1]. The idea that genes or chromosomes had a parental identity was first described in 1960 using observations in gametogenesis and sex determination in *Sciara*, a species of gnat. During gametogenesis and sex determination, only the maternally inherited X chromosome is used [57]. The ability to track the maternally inherited X chromosome during both processes provided the basis of genomic imprinting theory. This idea was supported during the development of nuclear transfer research. Early attempts found that copies of both parental genomes were essential for proper development. Embryos bearing two maternal genomes or two paternal genomes are not viable [58, 59]. This emphasized the existence of parental differences in the genome and supported the imprinting theory. In 1991, DeChiara et al discovered the first imprinted gene. They observed that mice heterozygous for an insulin-like growth factor II (*Igf2*) null mutation only showed an abnormal phenotype when the defective allele was inherited from the paternal side and demonstrated that the gene is expressed predominately from the paternal allele [60]. Since the identification of *Igf2* as an imprinted gene, approximately

150 other imprinted genes have been identified in mammals [61]. Interestingly, although there is a high degree of correlation between species, all show variations in what genes are imprinted and where they are imprinted [62].

2.1.2 General Features of Imprinted Genes

Imprinting has only been demonstrated in therian mammals, flowering plants, and some insects. In mammals, genomic imprinting exhibits temporal, tissue, and cell specificity and regulates genes involved in placental function, fetal growth, maternal nurturing, energy homeostasis, and neurologic function [7]. Interestingly, approximately half of imprinted genes displaying tissue specificity are expressed in the placenta and most of the remaining known tissue specific imprinted genes are expressed in the brain, however some imprinted genes are imprinted ubiquitously [63]. Theories discussing these observation are discussed in section 2.1.6 "Theories on the Evolution and Function of Genomic Imprinting. Imprinted genes tend to occur in clusters throughout the genome and display parentally conserved expression patterns within these loci [64]. These clusters all contain an imprinting control element (ICE) that regulates the imprinted expression of the genes in the cluster in *cis* [65]. Furthermore, many of these clusters also contain long non-coding RNAs [66]. Inheritance and regulation of imprinted genes is mediated by epigenetic processes, that is heritable factors regulating temporal and spatial gene expression patterns independent of changes in DNA sequence [1].

2.1.3 Inheritance of Epigenetic Modifications Regulating Genomic Imprinting

DNA methylation is a heritable modification of cytosine at the 5 position most often found at cytosine-guanine (CpG) dinucleotides and associated with DNA silencing [66, 67]. Discrete differences in methylation patterns of the germlines are termed germline differentially methylated regions (DMRs) [68]. Most ICEs are DMRs and appear to be the generationally inherited epigenetic marks that regulate imprinting [69]. During embryogenesis, the genomes of primordial germ cells migrating to the fetal gonad are stripped of the previously established epigenetic marks. During gametogenesis de novo methylation occurs to reflect gamete specific methylation patterns and germline DMRs. Upon fertilization, the genome goes through another large scale change in epigenetic programming; however, imprinting control elements escape reprogramming and retain the mark of their parental lineage [61]. These actions establish the imprint in the individual and allow ICEs to persist and regulate imprinted clusters.

2.1.4 Regulation of Imprinted Genes

The mechanisms used by ICEs to mediate genomic imprinting vary between loci. The ICE's general activity is based on the ability of proteins to specifically bind methylated or unmethylated DNA and subsequently exercise an effect on the surrounding locus [65]. This is the fundamental basis for the regulation of imprinting through three primary mechanisms: DNA methylation, histone modification, and expression of non-coding RNAs.

2.1.4.1 DNA Methylation

As discussed previously, DNA methylation occurs on cytosine residues throughout the genome. The covalent modification of the 5 position of cytosine typically occurs at cytosine-guanine dinucleotides (CpG) islands in the genome and is most often associated with gene silencing. Four DNA methyltransferases (DNMTs) regulate DNA methylation in cells and its stable inheritance in daughter cells [70-72]. De novo methylation, like that occurring in embryogenesis and gametogenesis, is regulated by DNMT3A and DNMT3B [73, 74]. The activity of DNMT3A proteins at imprinted loci during gametogenesis is regulated by an accessory protein, DNMT3L, which is essential to the imprinting process [75, 76]. Maintenance of the established DNA methylation patterns during cell division is controlled by DNMT1 [77]. The presence of methylated DNA at some locations can lead to passive methylated DNA spread surrounding the region. The primary mechanism of action of methylated DNA is to recruit or block DNA binding proteins based on their ability to bind methylated DNA [78, 79]. DNA methylation blocks the ability of both RNA polymerase II (RNApolII) and the insulator protein CCCTC binding factor (CTCF) [80]. Blocking of RNApolII inhibits transcription leading to gene silencing and blocking binding of CTCF leads to local dysregulation of expression (discussed later). Conversely, methylated DNA recruits other DNA insulators including polycomb repressive complexes (PRCs) responsible for repressive histone modifications [61].

2.1.4.2 Histone Post-Translational Modifications

Histones form the core of the nucleosome or "beads on a string" structure of DNA organization, and mediate the packaging of DNA within the cell. The four core histone proteins - H2A, H2B, H3, and H4 - form an octamer made up of four core histone protein dimers and represent major contributors to chromatin structure [81, 82]. Each histone has an amino acid tail that protrudes from the nucleosome and can be modified [83]. These modifications include varying degrees of methylation (mono-, di-, tri-), acetylation, phosphorylation, ubiquitinylation, sumolation, and biotinylation and are restricted to lysine, serine, threonine, and arginine residues [84]. These covalent modifications directly and indirectly mediate chromatin structure and correlate with transcriptional activity of the genome [83, 84]. Imprinted genes are associated with four of these marks. All ICEs are associated with the repressive mark histone 4 lysine 20 tri-methylation (H4K20me3) [85]. Furthermore, repressive H3K9me3 and H3K27me3 are associated with imprinted gene promoters in overlap with active H4K4me3 [68, 85]. Deciphering this "histone code" is a continuing project further complicated by theories and observation of combinatorial effects of these modifications, but recently bolstered by the ENCODE project [84, 86].

2.1.4.3 Non-Protein Coding RNAs

The majority of imprinted clusters are associated with the expression of an oppositely imprinted long non-coding RNA (lncRNA) [65]. Despite this correlation, not all lncRNA transcripts are involved in direct regulation of the imprint [87]. Furthermore,

the posttranscriptional processing and functionality of these lncRNAs vary by locus, with some of them containing other functional RNA units [65]. The role these lncRNAs take in the imprint of their cluster is addressed at individual loci in the following section.

2.1.5 Examples of Other Imprinted Genes and Regions

2.1.5.1 Igf2/H19 Region

The ICE regulating *the Igf2/H19* locus is methylated on the paternal chromosome [88]. Methylation of the paternal chromosome prevents the genetic insulator CTCF from binding to this site and allows a distant enhancer element to positively regulate *Igf2* [80, 89]. The ICE is also in close proximity to the promoter for *H19*, an lncRNA oriented in the opposite orientation of *Igf2 [90]*. This proximity leads to methylation of the promoter by unknown means, either passive DNA methylation spread or active targeting of DNA methyltransferases [65]. Conversely, absence of methylation on the maternal allele allows CTCF binding and the formation of a chromatin boundary between *IGF2* and *H19*. This blocks the enhancer interaction with *IGF2* and causes it to be silenced. The maternal *H19* promoter is unmethylated and amenable to positive regulation by the same enhancer [80, 89]. Thus the *Igf2/H19* locus is imprinted by the formation of a chromatin boundary and promoter methylation independent of expression of the lncRNA *H19* (**Figure 2A**) *[87]*.

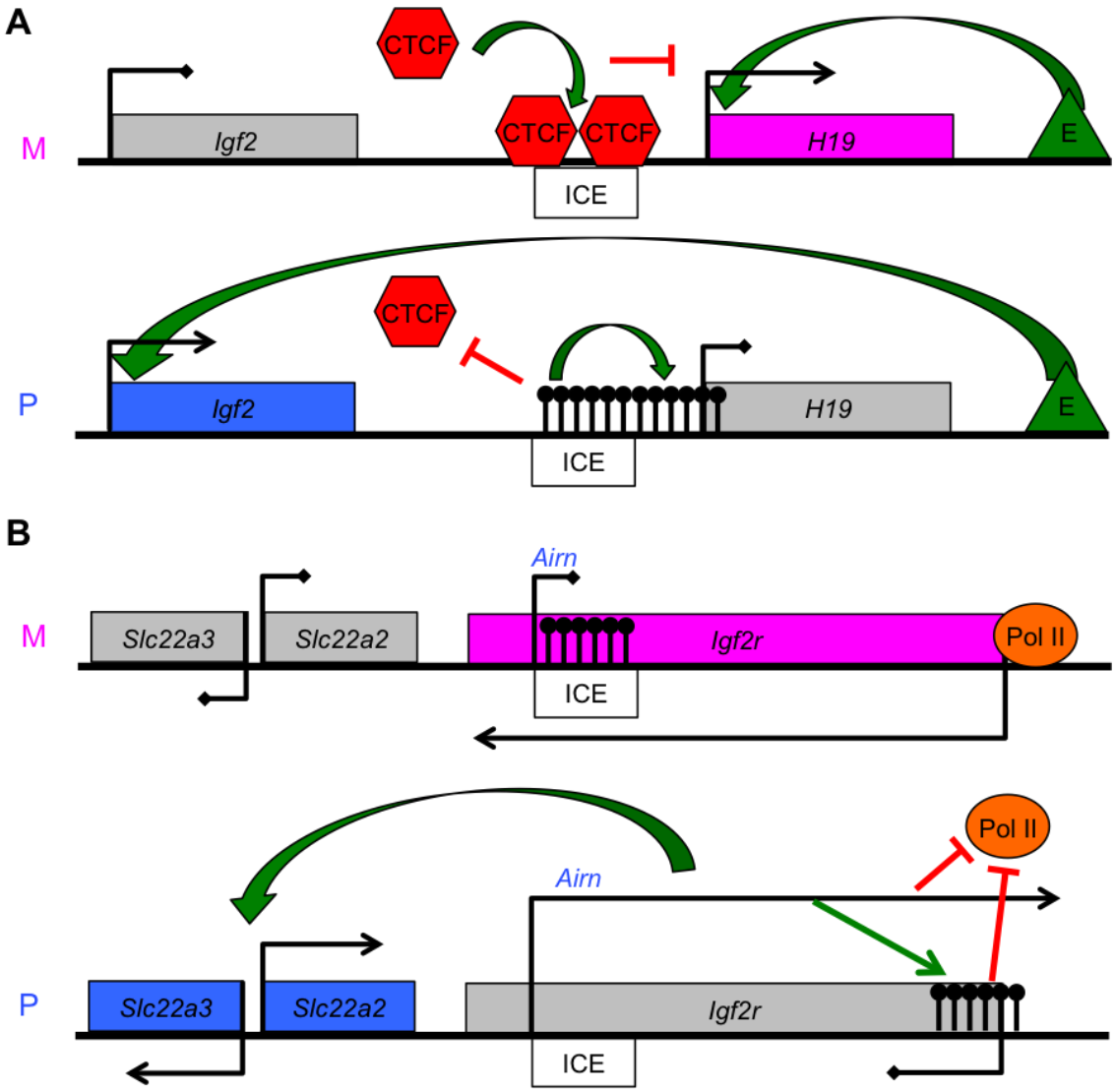


Figure 2. Mechanisms of genomic imprinting at the *Igf2/H19* and *Igf2r* locus. A) The insulin like growth factor 2 (*Igf2*) imprinting control element (ICE) is methylated on the paternal allele (filled lollipops). The presence of DNA methylation inhibits binding of the DNA insulator, CTCF, and spreads to the promoter of *H19*. Absence of CTCF allows positive regulation (green arrow) of *Igf2* from the distal enhancer (E). Absence of methylation of the maternal allele allows binding of CTCF and blocks enhancer action on *Igf2* (red bar). Lack of methylation at the *H19* promoter allows expression on the maternal allele and positive regulation by the enhancer. **B)** The insulin like growth factor 2 receptor (*Igf2r*) ICE is methylated on the maternal allele and overlaps the promoter of the lncRNA, *Airn*, blocking its expression. Absence of *Airn* expression on the maternal allele allows expression of *Igf2r*. Absence of ICE methylation on the paternal allele permits expression *Airn*. *Airn* overlaps the paternal *Igf2r* promoter, inhibiting recruitment of RNA polymerase II (Pol II) and promoting methylation. *Airn* also positively regulates the paternally expressed *Slc22a2* and *Slc22a3* genes by an unknown mechanism. Genes are colored to represent alleleic expression: pink = maternal, blue = paternal, grey = not expressed on this allele. Arrows indicate 5' → 3' transcription. "M" represents maternal allele and "P" represents paternal allele.

2.1.5.2 Insulin-like growth factor 2 receptor (IGF2r) Locus

The *IGF2R* locus is comprised of three maternally expressed transcripts and one paternally expressed lncRNA. The ICE lies in an intron of *Igf2r* and is methylated on the maternal allele. This ICE corresponds to the promoter of the cluster's lncRNA, *Airn [91]*. Expression of *Airn* on the paternal allele leads to repression of the three maternally expressed genes in cis [92]. The *Airn* transcript must be transcribed through the *Igf2r* promoter to induce methylation and silencing, but once established this methylation is sufficient to maintain the imprint in the absence of *Airn* [93, 94]. How AIRN expression induces *IGF2R* methylation is unknown. *Airn* regulates silencing of one of the other genes by recruiting repressive histone methyltransferases without transcriptional overlap [95]. On the methylated maternal allele, inhibition of *Airn* expression allows the expression of the three sequential maternally expressed genes. Thus, expression of the lncRNA at this locus is necessary for establishment and maintenance of the imprint (**Figure 2B**) [93, 94].

2.1.5.3 Kcnq1 Locus

The *Kcnq1* is centrally located in an imprinted cluster and flanked by a group placenta specific imprinted genes and a group of ubiquitously imprinted genes. Again, the ICE element is in an intron of *Kcnq1* and overlaps the promoter of the region's lncRNA, *Kcnq1*ot*1*. Although *Kcnq1*ot*1* is very long and overlaps the promoter of *Kcnq1*, this overlap is not required for imprinting of *Kcnq1*. A 5' critical region of *Kcnq1*ot*1* mediates chromatin folding to bring the *Kcnq1* promter into proximity with the ICE and

recruit histone transcriptional silencers [96] . The requirement for *Kcnq1*ot*1* expression to maintain the imprint is debated [96, 97]. Therefore, unlike the IGF2R locus, direct transcriptional overlap is not required for *Kcnq1*ot*1* function. Conversely, the lncRNA and transcriptional elongation is required for silencing of the other imprinted genes in the region by an undefined mechanism (**Figure 3**) [98].

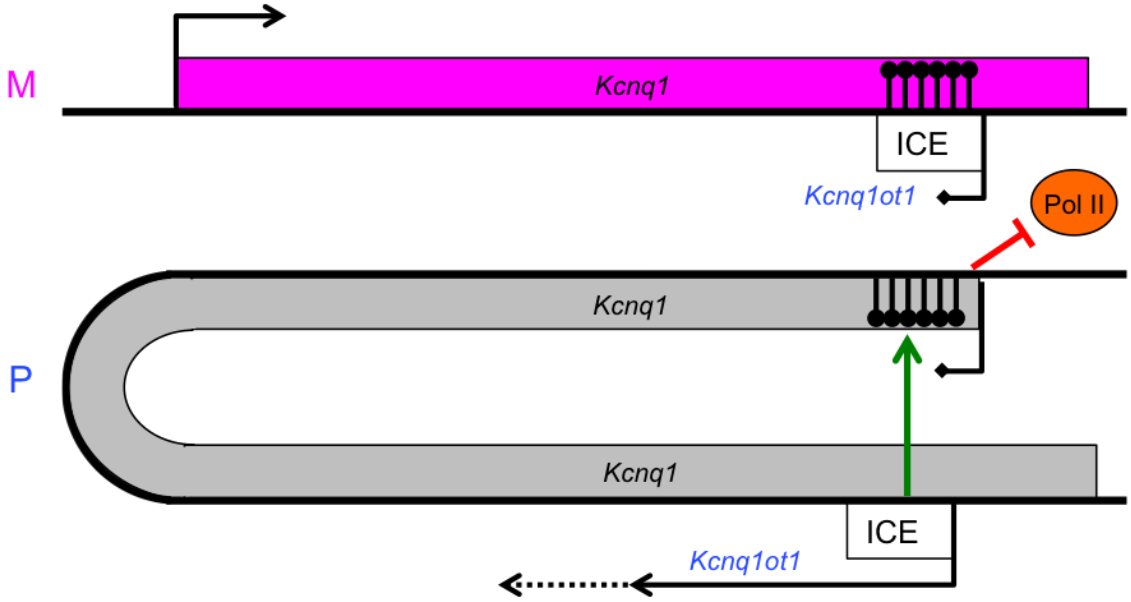


Figure 3. Mechanism of the *Kcnq1* imprint. The *Kcnq1* imprinting control element ICE overlaps the promoter of the lncRNA antisense transcript *KCNQ1ot1*. Methylation of the ICE (filled lollipops) on the maternal allele silences *KCNQ1ot1*allowing maternal expression of *KCNQ1*. The paternal ICE is unmethylated promoting expression of *KCNQ1ot1* for an undetermined distance (dashed arrow). The expression of *KCNQ1ot1* mediates folding of the paternal allele to bring the paternal *KCNQ1* promoter and ICE into proximity. This juxtaposition promotes methylation of the *KCNQ1* promoter (filled lollipops), blocking Pol II binding and silencing paternal *KCNQ1*. The other genes of this locus (not pictured) are also imprinted, but regulated by an unknown mechanism.

2.1.5.4 Gnas Locus

The *Gnas* ICE overlaps the maternally methylated Nespas-DMR over the promoter for the regions lncRNA, *Nespas* [99]. Paternal expression of *Nespas* from its unmethylated promoter leads to antisense transcriptional overlap with *Nesp*. This action results in loss of the histone 3 lysine 4 trimethylation and subsequent DNA methylation over *Nesp* [100]. The Nespas DMR overlaps the promoter of *GnasXL* leading to its paternal specific expression [101]. A second maternally methylated DMR overlaps *Exon1a* and is responsible for the tissue specific imprint of Gnas [102, 103]. These collective elements – the *Nespas* transcript, Nespas ICE, and Exon1a DMR – work together to regulate usage of alternative 5' exons that all share common parts of a group of 11 downstream exons in *Gnas*. On the maternal allele absence of the paternal transcripts allows both *Nesp* and *Gnas* to be expressed [104]. Therefore the imprint of this locus leads not to allele specific expression of a single gene, rather allele specific expression of alternative transcripts of single locus (**Figure 4**).

2.1.5.5 Imprinted Loci without lncRNAs

The *Copg2/Mest* region codes for two protein-coding genes and no identified lncRNA. These genes are arranged in antisense directions to each other, with overlapping 3'UTRs. Thus, expression of either gene precludes the expression of its counter-part. *Mest* is imprinted early in embryogenesis and hosts the only DMR in the region at its promoter. In eutharians, *Mest* expresses a transcript during neuro-development that utilizes an alternative polyadenylation site within *Copg2*. This transcriptional overlap

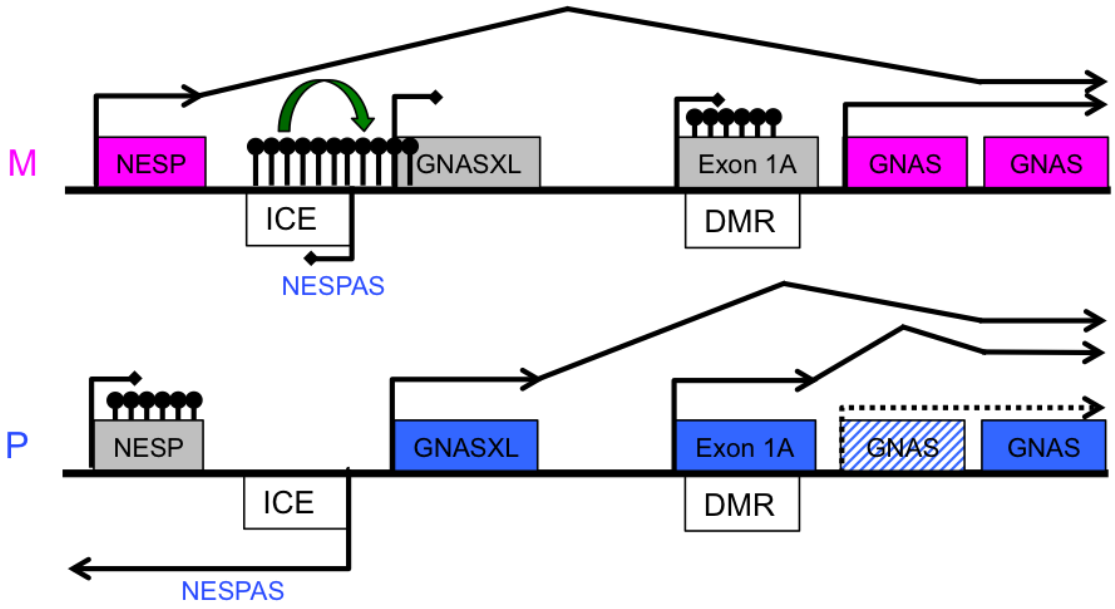


Figure 4. Mechanisms of the *Gnas* locus imprint. The *Gnas* locus contains an ICE and a differentially methylated region (DMR). The ICE overlaps the promoter of the lncRNA *Nespas*. The ICE is methylated on maternal allele and promotes methylation spread over the promoter of *GnasXL* (filled lollipops). This permits expression of *Nesp* from the maternal allele. Absence of methylation at the paternal ICE promotes expression of *Nespas* and *GnasXL*. *Nespas* expression inhibits paternal *Nesp* expression and leads to its methylation (filled lollipops). In tissues where the *Exon 1A* transcript is expressed, the maternal DMR is methylated (filled lollipops). This leads to imprinted expression of *Exon 1A* from the paternal allele which subsequently inhibits expression of paternal *Gnas* (stripped box) leading to predominately maternal expression of *Gnas* in these tissues. All of the coding transcripts at this locus – *Nesp, GnasXL, Exon 1A, and Gnas* – are functionally distinct but share the same 3' exons demarcated by *Gnas* [*].

leads to the imprint of *Copg2* not seen in representative metatherians that biallelically express *Copg2* (**Figure 5**) [105, 106]. Finally, a mouse specific imprint of the Murr1/Commd1 gene provides evidence of another innocent bystander imprint. In mouse an intronless zinc-finger gene, *Zrsr1*, lies in an antisense orientation in the first exon of *Commd1* and contains a DMR at its promoter. Thus, similarly to Mest/Copg2, the imprint of *Commd1* appears to be incidental and a result of DMR regulation of *Zrsr1* specifically in mice [107].

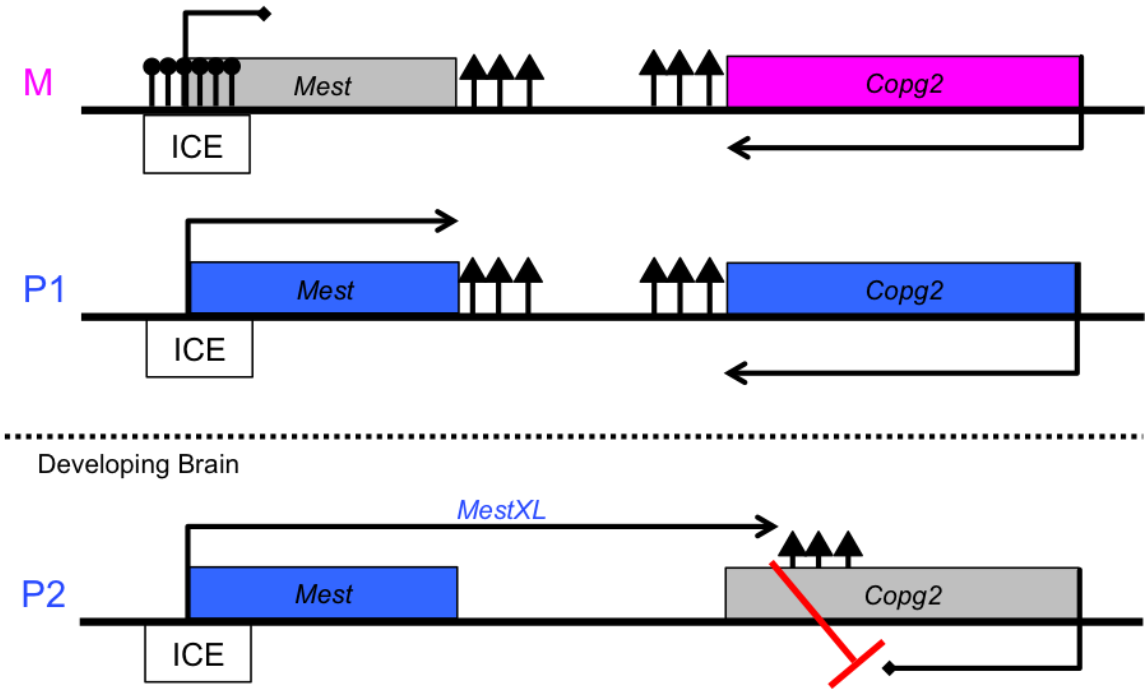


Figure 5. Differential imprinting of the Mest/Copg2 locus. In all tissues (M and P1). methylation of the maternal ICE (filled circular lollipops) leads to paternal specific expression of *Mest.* In these tissues both *Mest* and *Copg2* terminate at proximally located poly-adenylation sites (filled triangle lollipops). In the developing brain (P2, dashed line) paternally expressed *MestXL* uses an alternate poly-adenylation sequence that lies within *Copg2*. The overlap of *MestXL* and *Copg2* lead to inhibition of paternal expression of *Copg2* (red bar) and maternal allelic expression of *Copg2* in brain.

2.1.6 Theories on the Evolution and Function of Genomic Imprinting

There are many theories addressing the purpose and evolution of genomic imprinting, however, none of them fully encompass the myriad of cases that have been presented in the literature [7, 108]. When discussing the origins of genomic imprinting it is important to appreciate the impact it has on the organism. Genomic imprinting renders the organism functionally hemizygous for the allele, a historically poor indicator of fitness, allowing any deleterious changes in the active allele to be expressed [2, 3]. Therefore, the imprinting of a locus must have provided improved fitness in the face of possible

deleterious effects to be evolutionarily maintained [7, 109]. The host defense theory suggests that genomic imprinting derived from previously constructed cellular defenses to combat transposable elements [110]. Supporters point to the three key observations: 1) the relative expansion of eukaryotic genomes over their evolutionary course, noting particularly the increase of repetitive elements hallmarking transposons in therians compared to prototherians who lack imprinting; 2) the proximity of transposable elements to many imprinted clusters; and 3) the utilization of DNA methylation in silencing imprinted genes and transposable elements [7].

Another theory of the purpose of genomic imprinting is the kinship model or "parental conflict hypothesis." This idea proposes that parents have differing priorities for their offspring. Males strive to have large and strong progeny who will pass on their traits with little consideration of this effect on the mother, who carries and cares for most offspring, or their potential half-sib littermates. Conversely females would temper the males goals by having smaller progeny that would be less taxing of maternal resources during gestation and perinatally, ensuring survivability of larger litters [111]. This theory is strongly bolstered by the discovery of *Igf2*, an indicator of gestational size, as the first imprinted gene displaying paternal specific expression. The observation that many imprinted genes are maternally expressed and found in the placenta and extraembyronic tissues gives further support particularly when considering that placental mammals, females especially, show much higher levels of investment in their progeny compared to other animals lacking imprinting [108].

The natural selection model posits that genetic imprinting is a much more fluid phenomenon used to link environmental factors to the fitness of the animal. It suggests that environmental conditions during gestation play a role in the imprinting status of genes to promote fitness in the current climate the organism is being brought into. This theory relies on the preponderance of imprinted genes in the brain, particularly neurons, indicating the ability to have long term behavioral effects on the organism. Although environmental factors have been shown to mediate some epigenetic regulation, the evolutionary constraint, conservation, and stability of most imprinted genes does not support this as a prevailing theory for its origins [7, 108].

The ovarian time bomb/parthenogenic prevention model uses the observation that parthenogenesis does not occur in mammals, and when it does is disease based to the mother. This proposes that imprinting evolved to prevent parthenogenesis in mammals through requiring both parental genomes to be present in a developing embryo [112]. As none of these theories fully encompass all the examples of known imprinting, it is likely a combination of all of these that led to the genomic imprinting phenomenon [7, 108].

2.2 15q11-q13 Imprinted Region

2.2.1 Overview

The human 15q11-q13 region is syntenic to 7c in mouse and is highly conserved in eutherians. This region encompasses numerous imprinted and non-imprinted protein-

coding and non-protein coding transcripts. The imprinted cluster is primarily paternally expressed: *SNURF, SNRPN, MAGEL2, NDN, MKRN2,* and the *SNORD115* and *SNORD116* clusters) [11, 12]. UBE3A is the single maternally expressed imprinted gene in the region [13]. The imprinted cluster in this region, the PWS/AS imprinted cluster, is associated with three genetic diseases affecting cognition and development) [33-36]. The mechanisms regulating the imprinting of this cluster and their functional significance are still being investigated.

2.2.2 Evolution

The imprinted cluster within 15q11-q13 arose 105-180 million years ago. A duplication event of *Snrpb,* part of a gene family that encodes a small ribonucleoprotein-associated proteins, created a brain specific homolog *Snrpn* 180-210 million years ago at the divergence of therians and prototherians [10, 113]. Subsequently, *Snrpn* was translocated into proximity and antisense to the highly conserved gene *UBE3A* after the divergence of eutherian and metatherians 105-180 million years ago [10]. Independent retrotransposition events gave rise to *MAGEL2, MRKN3,* and *NDN* in proximity to *SNRPN [10, 114]*. The source and evolution of the SNORD clusters in the SNRPN-UBE3A region is unknown [10].

2.2.3 Imprinted Regulation

This region is regulated by a bipartite ICE comprised of two DMRs, the Prader-Willi syndrome imprinting center (PWS-IC) and the Angelman syndrome imprint center (AS-

IC) [21, 22]. The AS-IC is a germline DMR that is methylated in sperm and unmethylated in oocytes, but bound by a complex of approximately 7 proteins [24]. Alternately, the PWS-IC is unmethylated in both germlines. Upon fertilization, the maternal AS-IC is protected from methylation by its bound protein complex during epigenetic reprograming of the embryo. At the same time, it is believed that this same protein complex interacts with the maternal PWS-IC, epigenetically flagging it for methylation by an unknown mechanism. At the peri-implantation stage, the maternal PWS-IC epigenetic flag recruits DNA methyltransferases to gradually methylate the PWS-IC. Subsequently, the protein complex dissociates from the maternal AS-IC allowing for passive methylation. The location of AS-IC is not known in mouse. The end result is that in developing and adult cells the PWS-IC, which overlaps the promoter and exon1 of *SNRPN*, is only unmethylated, and thus functional, on the paternal allele, allowing it to positively regulate paternal specific transcripts in the region [24]. The PWS-IC is believed to directly regulate paternally expressed genes as paternal deletion of the PWS-IC leads to loss of paternally expressed genes [115]. In addition to regulating regional protein coding genes, the PWS-IC overlaps the promoter of *SNURF/SNRPN*, the *SNORD115* and *SNORD116* RNA clusters, and the regions lncRNA [21, 25, 29]. This lncRNA is antisense to and overlaps *UBE3A*, resulting in neuron specific imprinting of *UBE3A* in the brain [20, 25, 28-32]. The precise mechanism by which this lncRNA mediates the imprint is as yet unknown. Unlike other imprinted loci there are no repressive epigenetic marks surrounding UBE3A and the silenced paternal allele has active epigenetic marks (namely H3K4me3) and is bound by RNA

polymerase II (RNApolII) [31]. Furthermore, little investigation has been made in investigating the functional significance of this imprint or its effect on Ube3a expression.

2.2.4 The 15q11-q13 lncRNA

The 15q11-q13 lncRNA is known as the UBE3A-ATS, UBE3A-AS, or LNCAT and is necessary to imprint UBE3A [18, 31, 32]. This lncRNA is at least two times larger than any other imprint associated lncRNA [25, 65]. This lncRNA spans the UBE3A-SNRPN intergenic region that houses multiple imprinted non-protein coding transcripts and exhibits splicing across the two resident clusters of small nucleolar RNAs (snoRNAs), SNORD115 and SNORD116 [12, 26]. The independence of the transcripts containing the SNORDs and the portion of the lncRNA overlapping UBE3A is currently debated [116, 117].

Small nucleolar RNAs are short (27-35 nucleotide) RNAs that are processed from spliced out elements of lncRNAs and leave behind an lncRNA host gene transcript. Small nucleolar RNAs regulate nucleolar ribosomes through mediating their methylation, but the function of the host genes is unknown [117]. SNORD115 and SNORD116 are not predicted to have this function. SNORD115 regulates splicing of a serotonin receptor. There are approximately 23 predicted targets of SNORD116, but no empiric evidence exists for any of them [117]. The host genes for these snoRNAs, SNORD115-HG and SNORD116-HG respectively, may also have a function as they are retained at their site of transcription forming RNA clouds, physical clusters of RNA

[116, 117]. The significance of these snoRNAs and SNORD-HGs require further investigation.

2.2.5 Prader-Willi Syndrome

Prader-Willi syndrome (PWS) was first described by Prader, Labhart, and Willi in 1956 and occurs at an incidence of 1 in 10-30,000. This disease is hallmarked by mild to moderate mental deficiencies, hypotonia leading to poor nursing in the newborn, short stature, hypogonadism, and hyperphagia – often leading to obesity – in children and adults [118]. This disease most often arises from paternal inheritance of large deletions of 15q11-q13. Recently the disease was linked to specific defects in the SNORD116 cluster of this imprinted region, but the pathophysiologic contribution of this defect to PWS is unknown [42, 43].

2.2.6 Angelman Syndrome

2.2.6.1 Clinical Presentation

Angelman syndrome (AS), first characterized by pediatrician Harry Angelman as "puppet children" in 1965, has a worldwide prevalence of 1 in 12-20,000 [119, 120]. AS is a severe neurodevelopmental disorder characterized by the hallmark findings of severely impaired speech, gait disturbance and/or tremulous movement of the limbs, developmental delay without regression, and the characteristic disposition of frequent laughter, happy demeanor, and "marionette-like" hand flapping [40, 120, 121]. These findings are present in all cases of AS and provide the basis of strong clinical suspicion

for the disease. Furthermore, delayed head growth leading to microcephaly, seizures, characteristic abnormal electroencephalograms are found in more than 80% of cases and contribute to clinical suspicion of disease [40, 122]. There is a constellation of other symptoms with associations to AS, many of which relate to feeding and gastrointestinal related problems, changes in sleep/wake cycles, and ambulation. It is important to note that most AS patients have a normal pre-natal and perinatal history. The AS phenotype does not present before 6 months of age when developmental delay becomes evident and children lack abnormalities in typical laboratory profiles or gross changes in brain structure. Furthermore it is important to recognize a multitude of diseases that may mimic AS in the infant [123].

2.2.6.2 Genetic Causes

Angelman syndrome is caused by maternally inherited deficiencies in UBE3A. The majority of cases (65%) result from large maternally inherited deletions of 15q11-q13. Furthermore, imprinting defects (ID) leading to a maternally active PWS-IC or complete paternal uniparental disomy (UPD) accounted for approximately 3% and 7% of cases respectively. Point mutations in UBE3A are only responsible for approximately 15% of all cases. These various genetic causes of AS do lead to subtle variations in the AS phenotype. Cases of AS resulting from large deletions are the most severely affected group in all aspects of the disease with higher incidence of the less frequent associated AS phenotypes such as hypopigmentation. Children with UPD and ID causes have the mildest phenotype with decreased incidence of seizures and improved growth and

movement. Finally, children with non-deletion imprinting defects have the best expressive language skills with vocabularies of 50-60 words and the ability to formulate simple sentences [40, 121].

2.2.6.3 Diagnosis

Genetic testing is the standard for diagnosing suspected AS cases. DNA methylation testing of the *SNRPN* locus is the testing standard and can identify large maternal deletions, UPD, and imprinting defects. DNA fluorescent in-situ hybridization (FISH) remains the standard for diagnosis of smaller deletions and chromosomal rearrangement. Children not diagnosed by these tests can be investigated for mutations in *UBE3A*. Despite several related testing measures, 15-20% of clinically diagnosed AS patients fail to have any genetic defect detected [122].

There are currently no treatments for Angelman syndrome; however, various drugs and therapies are used to treat symptoms – seizures, irritability, etc. – to varying degrees of effect. Identifying treatment possibilities addressing the root cause of AS is an active field of research Daily et al demonstrated rescue of associative learning deficiencies in a mouse model through introduction of a viral vector expressing *UBE3A* [124]. Another group demonstrated that the imprinted paternal allele could be unsilenced using drugs, specifically topoisomerase inhibitors and subsequently showed that this occurred because of disruption of the paternally expressed antisense transcript [32]. Subsequently a mouse model that prematurely terminates the antisense transcript revealed that

reactivation of paternal *Ube3a* was able to rescue motor, learning, and neuronal functional deficiencies in AS mice to varying degrees, but had little effect on the AS behavioral phenotype [18]. These studies demonstrate the treatment potential for Angelman syndrome and a need for further understanding of its genetic regulation.

2.2.7 UBE3A

Ubiquitin protein ligase E3A (UBE3A) was first identified as E6-associated protein (E6-AP) through its interaction with the E6-protein of human papillomavirus (HPV) leading to ubiquitination of p53 in the highly oncogenic strains of HPV, 16 and 18 [44, 46]. *UBE3A* was mapped to the 15q11-q13 region in 1994 [13]. Subsequently, five transcriptional variants were identified, coding for three protein isoforms [47, 48]. Four transcriptional isoforms coding for the one unique and two shared protein isoforms have been identified in mouse. The functional significance of the transcript variants and protein isoforms remains unclear, with the exception of mouse isoform 1 and human isoform 3 lacking the functional HECT domain.

Although UBE3A was identified as the causative gene of Angelman syndrome, the exact pathophysiology of the disease remains to be elucidated. The discovery of UBE3A characterized a unique family of E3 ligases sharing a conserved catalytic C-terminal domain, HECT (homologous to E6-AP C-terminus) [51, 125]. The 28 identified HECT E3 ligases are distinguished from other E3 ligases (RING-finger and U-box E3 ligases) because of their ability to play a direct catalytic role in the ubiquitination of their targets.

These HECT ligases are conserved from yeast to humans but vary greatly in size and can be divided into three subclasses based on the binding domains used for substrate recognition at the N-terminus. Although many of the HECT E3 ligases have been associated with a variety of diseases, UBE3A is the only one shown to be the causative gene for a disease [45].

Twenty-four proteins have been described as targets of or associate with UBE3A independently of interactions with E6 protein. The functions of these proteins includes: cell cycle regulation, chromosomal segregation during cell division, transcriptional regulation, histone modification, proliferation and differentiation, cell migration, protein aggregates associated with other movement and neurodegenerative disorders, and excitatory synapse formation [49-54, 126, 127]. UBE3A is also a target of itself [128]. These pathways provide possible pathophysiologic pathways for the disease, but provide little conclusive evidence. There are, however, two targets of UBE3A involved in the formation of excitatory synapse formation that suggest the most likely explanation for the phenotypes seen in AS. Excitatory synapses are important neural connections in; the hippocampus, responsible for learning and memory; the cerebellum, responsible for motor coordination; and seizure initiation [129, 130]. Ephexin5 and Arc are both negative regulators of excitatory synapse formation and plasticity that are normally targeted for degradation by UBE3A. Ephexin5 binds the post-synaptic receptor EphB and increases RhoA, inhibiting synapse maturation. When EphB is bound by ligands on the presynaptic exon, it target Ephexin5 for degradation and the associated dendritic

spine and synapse can mature [131, 132]. Arc is a synaptic protein that mediates endocytosis of AMPA-glutamate receptors at the synapse following strong neural stimulation. It was initially believed that this same action induced Ube3a upregulation to inhibit uncheck progression of Arc activity, but it now appears that Ube3a may negatively regulate *Arc* transcription [133, 134]. In either instance, loss of Ube3a leads to reduction in neural plasticity through loss of AMPA receptors. Furthermore, excitatory synapse formation, particularly in response to experience dependent learning, upregulates UBE3A to promote inhibition of these negative regulators [129]. Thus UBE3A deficiency could directly inhibit appropriate synapse formation. This is further supported by observations that mice deficient for maternal *Ube3a* show aberrant dendritic spine morphology [135]. It is also important to mention that UBE3A, demonstrates steroid-receptor coactivator activity independent of its ligase activity, but that the role of this function if any in AS is unknown [56].

3. MATERIAL AND METHODS

3.1 Mouse RNA-Sequencing Analysis

RNA sequencing reads were obtained from the Mouse Genomes Project at the Wellcome Trust Sanger Institute under accession number ERP000591. The RNA sequencing reads represent total RNA derived from tissues of female F1 progeny (C57BL/6J females x DBA/2J males [n = 6 heart and 6 brain]). Mapping, normalized gene expression values, and SNV calling were computed using the CLC Genomics Workbench suite. Reads were aligned to the mm9 genome assembly with the RefSeq annotation database using "RNA-Seq mapper". The "Experiment tool" was used to calculate and normalize expression values between replicates and tissues. The "Probabilistic variant calling algorithm" was used to identify SNVs. These SNVs were then filtered to only include those previously identified in the parental lines, quality bases greater than 20, and called using at least 10 uniquely mapped reads. Independent mapping of the samples using Tophat was conducted for use in DESEQ analysis. Samples were mapped using default Tophat settings and –r set to 118. This resulted in an average of 54 million reads per sample. For DESEQ analysis gene bed files were generated from the RefSeq track at UCSC (mm9) and counts per gene generated using bedtools multicov. DESEQ normalization values were calculated from a restricted list of gene counts representing all genes with no known overlapping genes, alternative isoforms, or regional duplications in RefSeq. The Cufflinks (v.0.0.6) tool in Galaxy (galaxyproject.org) was used to generate transcript predictions from merged brain CLC mappings using default

parameters (min isoform fraction = 0.05 and Pre mRNA fraction = 0.02). Predicted annotations of the mouse Snord RNAs and Ube3a-AS were taken from the human RefSeq database (hg19) and PolyA-sequencing tracks were taken from the UCSC genome browser (http://genome.ucsc.edu/). Visualization, Sashimi plot generation, and images of RNA-seq data were generated in IgV 2.3. Sashimi plots were visualized to only display junctions with a depth of 10 reads in line with the threshold for SNVs identifications and represents 20% of average *Ube3a* intronic read coverage.

3.2 Mouse Strains

Ube3aYFP mice are a laboratory maintained strain [135]. Ube3a m-/p+ mice were generated in the laboratory of A. Beaudet and locally maintained on a C57BL/6 background [136]. C57BL/6 mice (Jackson Laboratories) were used for breeding.

3.3 Real-Time Polymerase Chain Reaction

RNA was extracted from tissue samples using the Purelink RNA Mini Kit (Life Technologies). The Superscript III First Strand Synthesis kit and oligo-dT primers (Life Technologies) were used for reverse transcription of processed RNA, and the High Capacity RNA to cDNA kit (Life Technologies) was used for reverse transcription of whole RNA. Real-time PCR was performed using Taqman Gene Expression Master Mix and Taqman Gene Expression Assays per manufacturer's protocol (Life Technologies). Beta-2 microglobulin (Taqman Assay #Mm00437762_m1) was used as a ubiquitiously expressed endogenous control. Taqman Assay #Mm00839910_m1 was

used to assess *Ube3a* levels. This primer and probe set targets an amplicon of 121 base pairs which spans exons 6 and 7 of Isoforms 1 and 3 and exons 8 and 9 of Isoform 2 of mouse *Ube3a*. The reactions were run on an ABI 7900HT real-time PCR machine, and results were analyzed using the $\Delta\Delta C^t$ method.

3.4 Reverse Transcription PCR, Complementary DNA Cloning, and Sequencing

RNA was extracted and reverse transcribed as described above. cDNA was amplified using primers (**APPENDIX Q**) and GoTaq DNA polymerase (Promega, Madison, WI). Amplification mix consisted of: 8.5 μl water, 5.0 μl GoTaq buffer, 5.0 μl betaine, 2.5 μl $MgCl_2$, 0.75 μl dNTPs, 0.5 μl of each primer, 0.25 μl polymerase, and 2 μl of 1:5 diluted cDNA. Reactions were amplified by denaturing at 95C for 5 minutes followed by 35 cycles – 95C for 45 seconds, 57C for 45 seconds, 72C for 3.5 minutes – followed by 1 cycle of 72C for 7 minutes. Products were resolved on 1% agarose gel. The resulting PCR products were gel purified using the Zymoclean Gel DNA Recover Kit (Genesee Scientific, San Diego, CA). Incubation of the purified products with GoTaq polymerase and dNTPs at 72 C for 15 minutes added T overhangs for T/A cloning. The TOPO TA Cloning Kit for Subcloning (Life Technologies) was used according to manufacturer's protocol. Plasmids containing the cloned products were then sequenced. Some PCR products were sequenced from direct PCR products. The resulting sequences were trimmed and aligned to the mouse mm9 using UCSC BLAT.

3.5 Western Blot Analysis

Tissues and cells were disrupted for western blot analysis using a 1% Nonidet P40/ 0.01% SDS lysis buffer with protease inhibitors (Roche, Indianapolis, IN). The resulting lysates were mixed 1:1 with Laemmli Loading Buffer (Bio-Rad, Hercules, CA). The samples were run on SDS-PAGE gels (7.5%) at 25V overnight, and transferred to nitrocellulose blots at 100V for 2 hours. To control for loading, the blots were dyed with Ponceau stain (Sigma-Aldrich) and digitally photographed. The blots were then blocked in 5% milk in Tris-buffered saline plus Tween-20 (T-TBS) for one hour at room temperature. Primary antibody (**APPENDIX R**) was diluted in 2.5% milk/T-TBS and incubated on the blot for one hour at room temperature. After three 15 minute washes in T-TBS, the secondary antibody (**APPENDIX R**) was diluted 1:2000 in 2.5% milk/T-TBS and incubated on the blot for one hour at room temperature. Three 15 minute washes in T-TBS were performed before developing with Clarity Western ECL Substrate (Bio-Rad), according to protocol. Blots were imaged using the FluoroChem system. Analysiss of western blot imaging was conducted in ImageJ as outlined in Gassman et al, using the Ponceau stain image analysis for blot normalization [137].

3.6 Tissue Perfusion and Immunofluorescence

Mice were anesthetized with 0.5-1.0 mLs of 20mg/mL Avertin (Sigma-Aldrich, St. Louis, MO) via intraperitoneal injection. Mice were perfused with ice-cold phosphate buffered saline (PBS) and 4% paraformaldehyde in PBS. Dissected brains and colons were post-fixed in 4% paraformaldehyde solution overnight and then cryoprotected in

30% sucrose solution. 50 µm sections were cut on a cryostat and stored in PBS. Sections were washed in PBS and blocked in 0.3% Triton-X100 in PBS (T-PBS) plus 5% normal goat or donkey serum for 1-2 hours at room temperature with gentle agitation. Primary antibodies were incubated with sections for 48 hours at 4˚C with gentle agitation (**APPENDIX R**). Sections were washed 3 times in 0.1% Tween 20 1x PBS for 15 minutes each and then incubated with fluorescently labeled secondary antibodies (**APPENDIX R**) for 24 hours at 4˚C in the dark. Sections were washed 4 times in 0.1% Tween 20 1x PBS for 15 minutes each. Nuclei were labeled by adding TO-PRO-3 (Invitrogen) at 1:1000 dilution in the third wash. Sections were mounted on glass slides with Vectashield (Vector Laboratories, Burlingame, CA) mounting reagent. Confocal images were obtained using a LSM 510 META NLO multiphoton microscope (Zeiss, Oberkochen, Germany). Confocal microscopy was performed in the Texas A&M University College of Veterinary Medicine & Biomedical Sciences Image Analysis Laboratory, supported by NIH-NCRR (1 S10 RR22532-01). See supplementary imaging methods for pertinent immunofluorescent imaging collection settings.

3.7 Neural Stem Cell Cultures and Differentiation

Methods for neural stem cell culture and differentiation were adapted from Shetty et al [138]. The hippocampal formation (HF) was removed from E17.5 mice. The HF was enzymatically digested using a 10X Trypsin-EDTA solution, triturated into a single cell suspension and then seeded in neural stem cell medium. The neural stem cell media consists of DMEM-F12 (Invitrogen, Carlsbad, CA), B-27 supplement (Invitrogen),

progesterone, putresceine (Sigma-Aldrich), epidermal growth factor (Sigma-Aldrich), glucose, Penicillin/Streptomycin (Invitrogen), Insulin-transferrin-sodium selenite (Sigma-Aldrich), HEPES, and heparin. The neurospheres were passaged every 3-4 days by spinning down and incubating in Tryp-LE (Invitrogen) for 20 minutes before resuspending and dissociating in NSC media. To differentiate the NSCs into neurons, the neurospheres were dissociated as above and plated into neuron growth media (Neurobasal-A [Invitrogen], B-27 supplement [Invitrogen], and Glutamax [Invitrogen]) and on glass coverslips coated with poly-ornithine [Sigma-Aldrich] and laminin [Invitrogen] at a density of 380,000 cells per well in a 12-well cell-culture plate. NSCs and neurons were maintained in humidified incubators at 37˚C and 5% CO_2. For fixation, cells were washed twice with PBS, fixed in 4% paraformaldehyde in PBS for 15 minutes with gentle agitation, and washed three times in 1x PBS. Immunofluorescence staining and imaging of the neurospheres and neurons were performed as detailed above for the brain tissues, with blocking and incubation with antibodies lasting one hour each.

3.8 Primary Neuronal Cell Cultures

Culturing of primary neurons was conducted as in Hilgenberg *et al.* with only slight departures from physical handling and culturing media [139]. Briefly, P1 pups were sacrificed by decapitation and brains removed to dissection solution. Meninges were removed and brains were coronally sectioned by hand using #20 surgical scalpels. Cortical rind was dissected from coronal sections, minced to ~1mm pieces and digested

in the presence of papain (Worthington Biochemical). Tissues were washed multiple times to stop enzyme activity and clean tissues. Brain pieces were triturated 8-10 times in neuron media (see above) using a 1mL pipette. Cells were counted and plated at a density of 6400 cells/well in optical bottom 96 well plates (Thermo Scientific) pretreated with poly-d-lysine (Sigma Aldrich) or 40,000 cells/well in 24 well plates with 12mm coverslip coated in poly-d-lysine and maintained in humidified incubators at 37˚C and 5% CO_2. Cells were supplemented with an equal volume of neuron media conditioned on previously plated cerebellar astrocyte cultures after 24 hours. Fifty percent of media was changed at 4 and 7 days *in vitro* (DIV) using fresh conditioned neuron media. For treatment of cells, topotecan hydrochloride (Sigma Aldrich) was added to replacement media at a 600nM concentration to bring final concentration on cells to 300nM. Similarly DMSO was added to replacement media at a concentration of 0.4% to bring final concentration to 0.2% on cells. Treatments were carried out with media change at 7 DIV. Staining and imaging of 24 well coverslips was carried out as for differentiated neurons above. Staining of 96 well plates was carried at as previously stated for differentiated neurons after 11 day *in vitro* using YFP, NeuN, and TO-PRO-3 staining as in *in vivo* studies. Imaging was performed on a GE In Cell Analyzer 6000 at Texas A&M University Health Science Center Institute for Biotechnology, Center for Translational Cancer Research. Wells were tiled using a 10x objective with 0% overlap.

3.9 Astrocyte Cultures

Cerebellum from similarly P1 pups were collected as above, meninges removed, and loosely dissociated by pipetting with 1mL pipette in 10x trypsin and incubated at 37 ˚C and shaken at 750 rpm for 1 minute before being allowed to settle and resuspend in 10% Fetal bovine serum in DMEM-F12 (Invitrogen, Carlsbad, CA) supplemented with glucose and Penicillin/Streptomycin (Invitrogen) for another 1 minute shaking incubation. Tissues were allowed to settle and then resuspended and triturated into single cell suspension in 10% FBS culturing media before being plated into 5mls of media in 25cm^2 flasks and maintained at 37˚C and 5% CO_2.

3.10 Analysis of Immunofluorescent Images

Image J (NIH) was used in image preparation and to measure gray scale values of individual cells in images of non-high throughput cell culture and *in vivo* studies. Neurons were delineated from other cells by positive Tubb3, Map2, or NeuN staining. The fluorescence of each cell was expressed as YFP intensities overlapping neuronal TO-PRO staining.

Analysis of high throughput imaging was carried out with the In Cell Developer 6.0. Briefly, individual track masks were generated for both TO-PRO-3 nuclear stain and NeuN staining channels by optimizing inclusion and exclusion parameters based on size and intensity on randomly selected images. A third mask was generated of the overlapping area between objects in the TO-PRO-3 and NeuN masks and YFP intensity

was measured for these objects. Measurements and statistics were examined as an average of measured YFP intensities per well.

3.11 Statistical Analyses

All statistical analysis was carried out in R (3.0.1 -- "Good Sport"). The Shapiro-Wilk normality test (shapiro.test) was run on all data sets. Variance in all comparisons was tested using the F test to compare two variances (var.test). Significance was tested using R defaults when controlling for normality and variance. A Two Sample t-test was used to compare normally distributed data sets with equal variance. A Welch Two Sample t-test was used to compare normally distributed data sets with unequal variance. Wilcoxon rank sum test was used to compare data sets when at least one was not normally distributed but both had equal variance. Wilcoxon rank sum test with continuity correction was used to compare data sets when at least one was not normally distributed and there was unequal variance. Ashman's D was chosen as a measure of severability of two sub-populations and calculated by hand.

For comparisons of total Ube3a:YFP expression between peripheral and central neurons and DIV 1 and DIV 16 NSC differentiated neurons random sums were used. Briefly, data sets for each allele were randomized and paired (n= 3) and summed. This generated 3 lists of paired sums for each sample (DIV 1 neurons, DIV 16 neurons, peripheral neurons, central neurons). All pairwise comparisons were tested as above, and the lowest *P* value was reported.

3.12 Graphs

All graphs were generated in GraphPad Prism 6™.

4. RESULTS

4.1 *Ube3a* Allelic Expression Does Not Correlate with Its RNA and Protein Levels

The transcriptional profiles, steady-state total RNA (toRNA) levels, and single nucleotide variants of RNA transcripts (RNA-genes, intergenic and intragenic RNAs of unknown function) and protein-coding genes within and flanking the mouse 7C imprinted locus were examined using RNA-sequencing data derived from brain and heart tissues of F1 hybrid mice (**Figure 6A-B**). *Ube3a* was imprinted with preferential maternal-specific expression in the brain (M:P ratio = 83:17) and biallelically expressed in heart (43:57) (**Table 1 and Appendix A-C**). Numerous RNA transcripts identified in the *Ube3a-Snrpn* region were expressed only in brain and exclusively from the paternal chromosome. These transcripts were consistent with the expression of the long polycistronic RNA transcript involving *Snurf*/*Snrpn*, *Snord116/Ipw*, *Snord115*, and the *Ube3a-AS* (**Table 1 and Appendix A-C**). Overall, the normalized toRNA levels of protein-coding genes, with the exception of *Nipa2* and *Peg12*, were higher in brain than in the heart ($P < 0.05$). The *Ube3a* toRNA levels in brain were 1.6-fold higher than in heart, and toRNA levels of the parental alleles diverged between the tissues (maternal allele: brain>heart [$P < 0.001$]; paternal allele: brain<heart [$P = 0.016$], [**Table 1 and Appendix A-C**]). Additional assays were used to validate the *Ube3a* RNA-seq data, determine *Ube3a* messenger RNA (mRNA) steady-state levels, and examine additional brain regions and tissues. Accordingly, *Ube3a* toRNA and mRNA levels were 2- to 18-fold higher in each brain region relative to the other tissues ($P = 0.002$) (**Figure 7A**) and

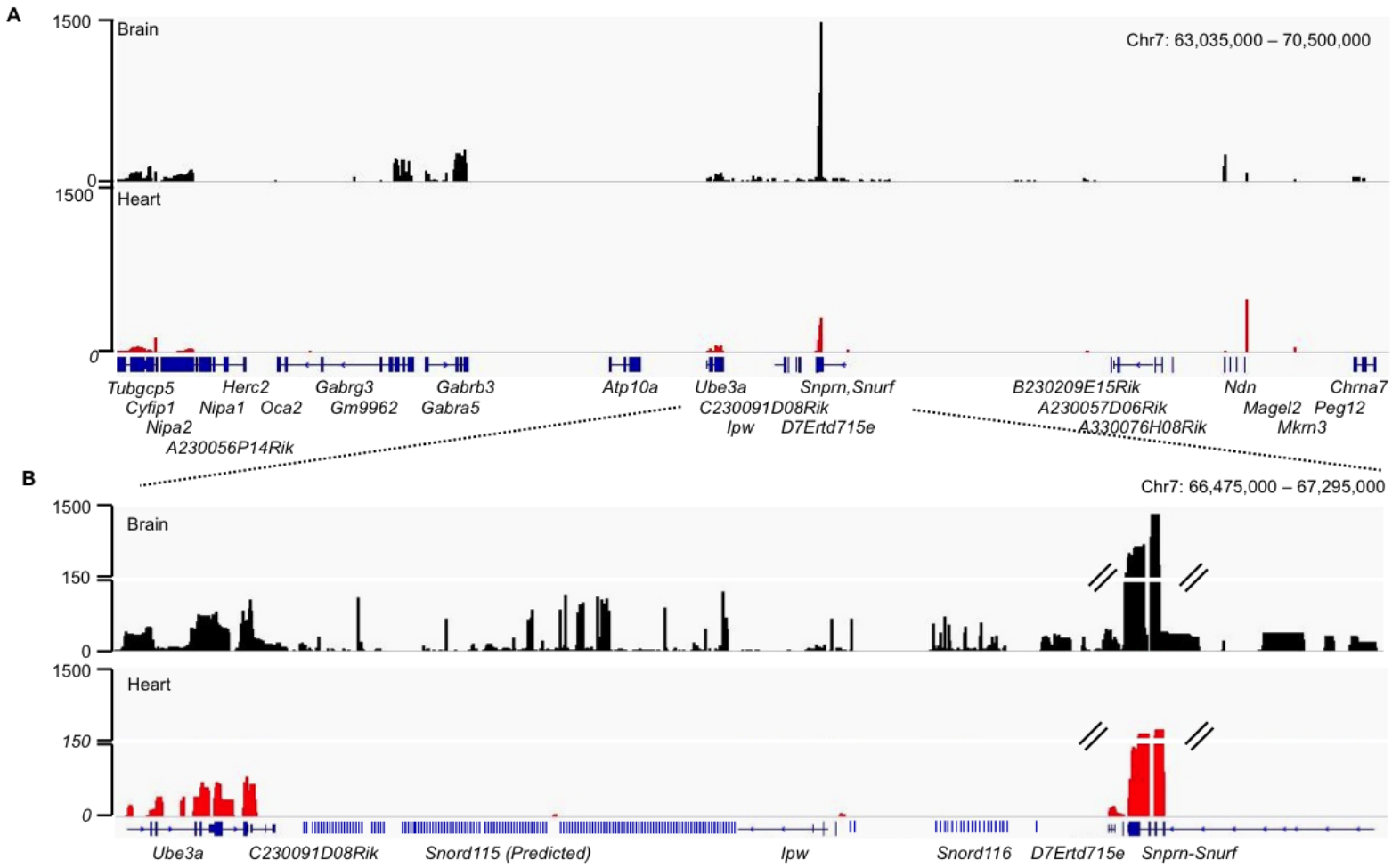


Figure 6. RNA-seq mapping by tissue for C57BL/6J x DBA2/J hybrid mice. A) Representative sample of total read counts across mouse 7C imprinted region summarized as a 10bp window BigWig file. Black, brain; Red, heart. **B**) Representative sample of total read counts across *Ube3a-Snprn* intergenic region containing proposed *Ube3a-AS* expression summarized as a 10bp window BigWig

Table 1. **Summary of relative expression levels and allelic ratios of genes in the mouse 7c imprinted region as determined by RNA-seq analysis.**

Gene	Heart Mean Expression	Hippocampus Mean Expression	Fold Increase in Brain	Heart Allelic Ratio M:P	Hippocampus Allelic Ratio M:P
Tubgcp5	2.94	8.39	2.86***	40 : 60	48 : 52
Cyfip1	4.35	7.86	1.81**	39 : 61	59 : 41
Nipa2	4.12	4.68	1.13	ND	ND
A230056P14Rik	0.53	4.84	9.16***	41 : 59	54 : 46
Nipa1	2.38	12.93	5.44***	ND	ND
Herc2	3.70	23.51	6.35***	43 : 57	52 : 48
Oca2	0.00	0.12	∞	ND	ND
Gabrg3	0.01	2.72	249.05***	ND	50 : 50
Gm9962	0.00	0.12	∞	ND	ND
Gabra5	0.00	69.96	∞***	ND	ND
Gabrb3	0.02	84.92	4959.63***	ND	51 : 49
Atp10a	0.53	1.62	3.04***	36 : 64	53 : 47
Ube3a	8.05	13.29	1.65*	43 : 57	83 : 17
C230091D08Rik	8.53	14.09	1.65*	45 : 55	41 : 58
Ipw	0.00	4.75	∞***	ND	ND
D7Ertd715e	4.54	12.98	2.86**	0 : 100	0 : 100
Snrpn-Snurf	26.63	152.96	5.74***	0 : 100	0 : 100
B230209E15Rik	0.00	9.87	∞**	ND	0 : 100
A230057D06Rik	0.00	0.42	∞***	ND	0 : 100
A330076H08Rik	0.00	1.41	321.53***	ND	0 : 100
Ndn	6.12	105.14	17.18***	0 : 100	0 : 100
Magel2	0.00	0.25	∞**	ND	ND
Mkrn3	0.10	0.20	2.04	ND	0 : 100
Peg12	0.18	0.04	0.25	ND	ND
Chrna7	0.02	12.99	662.37**	ND	56 : 44

ND = No allelic data; * = p<0.05, ** = p<0.01, *** = p<0.001

Ube3a protein levels were 2.8- to 3.5-fold higher in the brain relative to the other tissues ($P < 0.001$) (**Figure 7B**), indicating a nonlinear relationship between the number of expressed *Ube3a* alleles and RNA or protein levels.

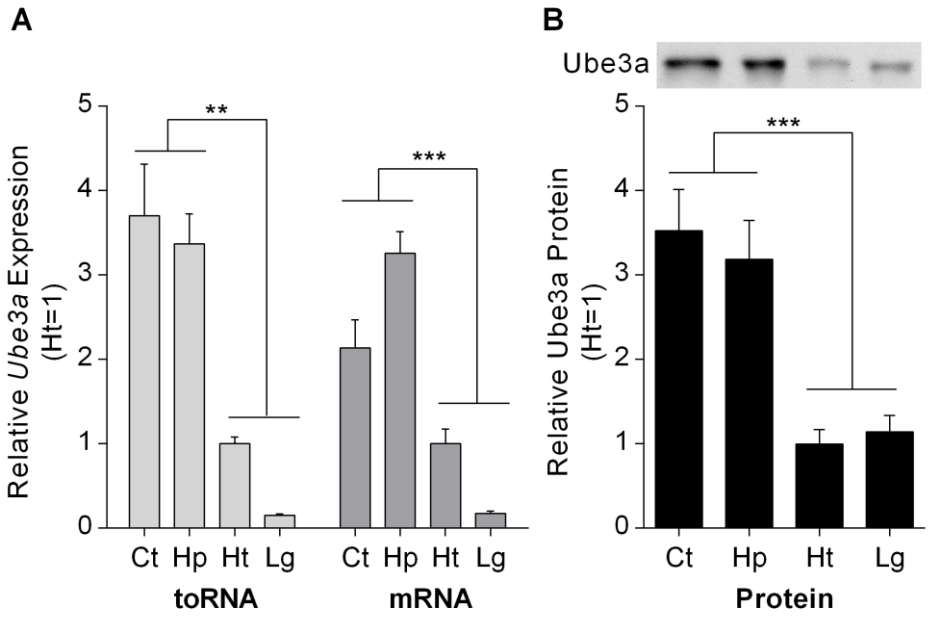


Figure 7. The mouse 7C imprinted domain is highly expressed in brain compared to other tissues. **A)** Normalized *Ube3a* levels from toRNA and mRNA isolated from wild-type mouse tissues. Levels are shown as the ratio of expression in tissues relative to heart, normalized to the housekeeping gene Beta-2 microglobulin. **B)** Normalized Ube3a protein levels isolated from wild-type mouse tissues. Levels are shown as the ratio of expression in tissues relative to heart, normalized by Ponceau total protein stain. n = 4. *** P< 0.001. Ct, cortex; Hp, hippocampus; Ht, heart; Lg, lung.

4.2 *Ube3a* Is Not Imprinted in Neurons of the Peripheral Nervous System

The *Ube3a* gene is imprinted in neurons of the central nervous system (CNS), but its imprinting status in neurons of the peripheral nervous system (PNS) is unknown [19,

135]. Therefore, we used quantitative immunofluorescence imaging of the Ube3a:YFP reporter protein in *Ube3a*$^{YFP/+}$ and *Ube3a* $^{+/YFP}$ mice to infer *Ube3a* parental-allelic expression patterns in myenteric neurons of the colon. We detected both paternal- and maternal-Ube3a:YFP in myenteric neurons; however, the level of protein generated by each parental allele was skewed toward the maternal allele (58:42; P = 0.033) (**Figure 8A-B**). Comparisons between myenteric and hippocampal granular neurons of the same animal revealed similar levels of total Ube3a:YFP (P = 0.113) between the two neuronal cell types, suggesting *Ube3a* allelic expression patterns do not correlate with total Ube3a protein levels in neurons (**Figure 8A-B**).

4.3 The *Ube3a* Imprint Is Developmentally Regulated

Our observation of elevated Ube3a protein in the brain and restriction of the imprint to CNS neurons, prompted us to investigate the developmental timing of the imprint in two neurogenic niches of the adult brain: the subventricular zone (SVZ) of the lateral ventricles and the subgranular zone (SGZ) of the dentate gyrus (DG). In the SVZ, paternal- and maternal- Ube3a:YFP were detected in neural stem cells and precursor cells lining the lateral ventricles and in immature neurons throughout the rostral migratory stream (RMS) (**Figure 9A-D**). In the olfactory bulb (OB), maternal-Ube3a:YFP levels were visually similar to adjacent neurons in the cortex (**Figure 9C**).

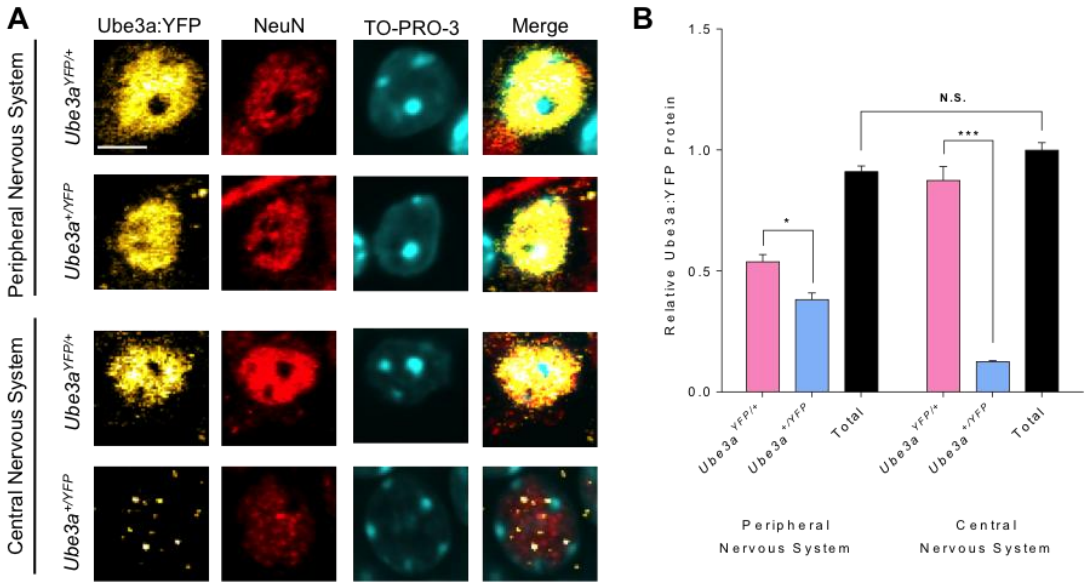


Figure 8. ***Ube3a*** **is biallelically expressed in mature neurons of the peripheral nervous system**. **A)** Immunofluorescence of mature neurons in the central (dentate gyrus) and peripheral (myenteric ganglia) nervous system of *Ube3a*$^{+/YFP}$ and *Ube3a*$^{YFP/+}$ mice. YFP, Ube3a:YFP; NeuN, neuronal nuclei RBFOX3; TO-PRO-3 Iodide nuclear stain. Scale Bar, 10μm. **B)** Relative Ube3a:YFP expression as fold change over wild-type controls and total expression as a sum of allelic expression. n = 26 central and 15 peripheral *Ube3a*$^{+/YFP}$, 20 central and 30 peripheral *Ube3a*$^{YFP/+}$, 12 central and 5 peripheral *Ube3a*$^{+/+}$. * P < 0.05, ** P < 0.01, *** P < 0.001, N.S. not significant.

Paternal-Ube3a:YFP was only weakly detected in the OB, but more abundant than observed in cortical neurons, which may have resulted from increased cellular density in the OB (**Figure 9D**). A sharp decline in paternal-Ube3a:YFP was observed at the boundary between the RMS and OB, which mirrored the expression pattern of doublecortin-labeled immature neurons in the RMS. In the DG, maternal-Ube3a:YFP was detected in the neurogenic niche and granular cortical layer neurons (**Figure 10A**), whereas paternal-Ube3a:YFP was only detected in the neurogenic niche (**Figure 10B**).

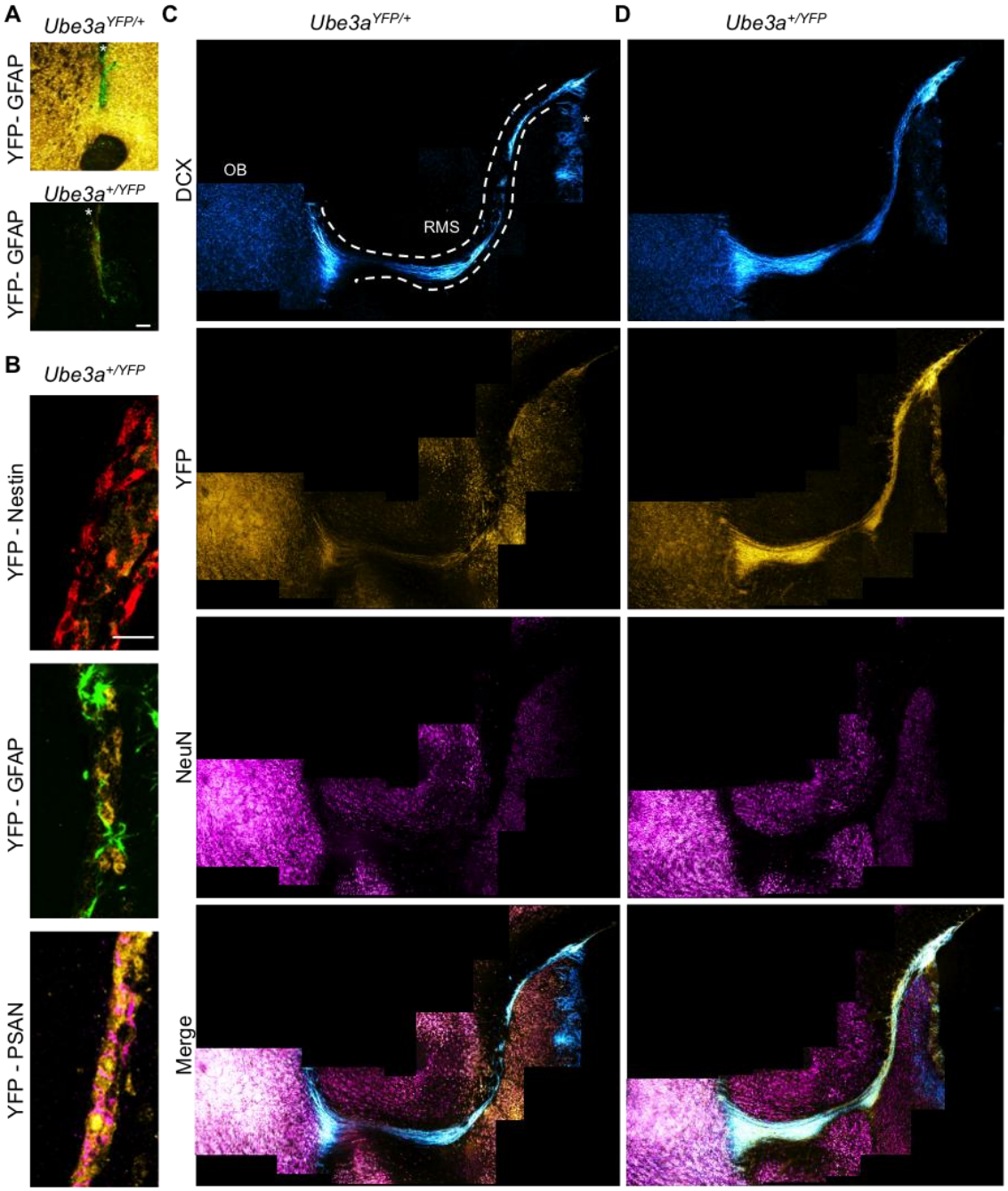


Figure 9. ***Ube3a*** **is biallelically expressed in the subventricular zone neural stem cell niche of the adult central nervous system (CNS)** ***in vivo*****. A)** Subventricular zone of *Ube3a*$^{+/YFP}$ and *Ube3a*$^{YFP/+}$ mice showing biallelic expression of Ube3a:YFP. 10x Magnification. Scale Bar = 100µm**. B)** Subventricular zone of *Ube3a*$^{+/YFP}$ mice demonstrating paternal Ube3a expression in GFAP expressing neural stem cells, Nestin expressing neural progenitor cells, and PSAN expressing neural progenitor cells. 43x Magnification. Scale Bar = 20µm**. C-D)** Immunofluorescence of neural differentiation markers and allelic Ube3a^{YFP} in the rostral migratory stream of adult mice (6-8 weeks), **C** = *Ube3a*$^{+/YFP}$ and **D** = *Ube3a*$^{YFP/+}$. Scale bar, 500µm. YFP, Ube3a:YFP; GFAP, glial fibrillary acid; PSAN, polysialylated neuronal cell adhesion molecule, DCX, doublecortin – immature neurons ; RMS, rostral migratory stream – demarcated by dashed lines ; OB, olfactory bulb ; *, lateral ventricles

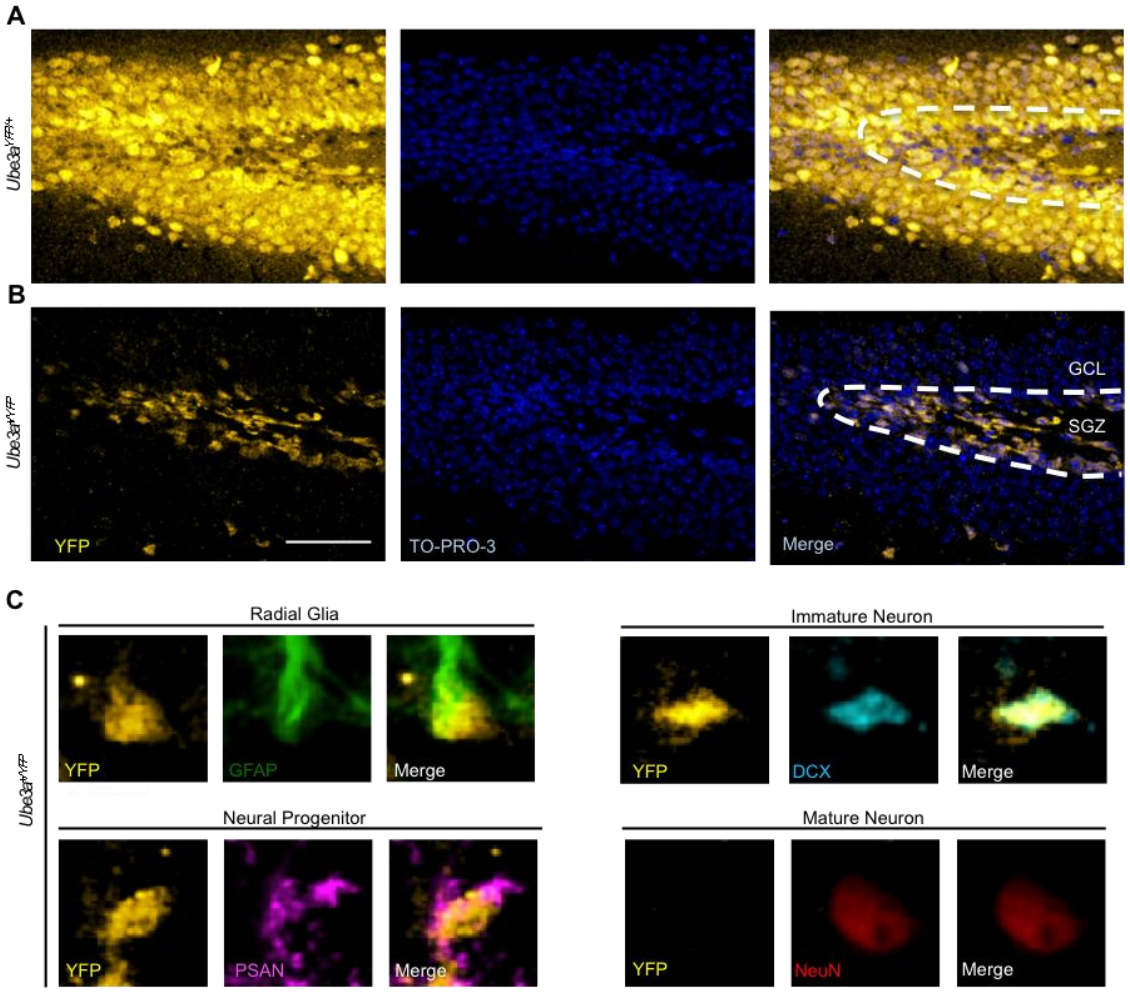


Figure 10. ***Ube3a*** **is biallelically expressed in the neural stem cell niche in the subgranular zone of the dentate gyrus of the adult central nervous system (CNS)** ***in vivo***. **A-B)** Dentate gyrus of adult (6-8 weeks) *Ube3a*$^{YFP/+}$ (**A**) and *Ube3a*$^{+/YFP}$ (**B**) mice. YFP, Ube3a:YFP; Topro, TO-PRO-3 Iodide nuclear stain; GCL, granular cortical layer; SGZ, subgranular zone. 10x Magnification. Scale Bar, 50μm. **C)** Single cell immunofluorescence neural maturation from SGZ of adult (6-8 weeks) *Ube3a*$^{+/YFP}$ mouse. 10x Magnification. Scale Bar = 20μm. YFP, Ube3a:YFP; GFAP, glial fibrillary acid; PSAN, polysialylated neuronal cell adhesion molecule; DCX, doublecortin; NeuN, neuronal nuclei RBFOX3; SGZ, subgranular zone; GCL, granular cortical layer.

Colocalization studies showed that paternal-Ube3a:YFP was present in neural stem cell radial glia, neural precursor cells, and immature neurons; however, it was minimally present in mature granular neurons (**Figure 10C**). Taken together, these data suggest

Ube3a transitions to an imprinted state as immature neurons differentiate into developmentally mature neurons.

4.4 *Ube3a* Maternal-allelic Expression Increases During Establishment of the Imprint

We next quantified steady-state paternal- and maternal-Ube3a:YFP levels in neural stem cell derived neurons and astrocytes during the early stages of cellular differentiation. Primary neural stem cell/progenitor cell (NSC) cultures were established from the hippocampal formation of *Ube3a*$^{YFP/+}$ and *Ube3a*$^{+/YFP}$ mice and differentiated into astrocytes and neurons. The allelic ratios were equivalent (~50:50) in the NSC cultures after multiple passages in culture (**Figure 11A-C**). Similarly, paternal- and maternal-

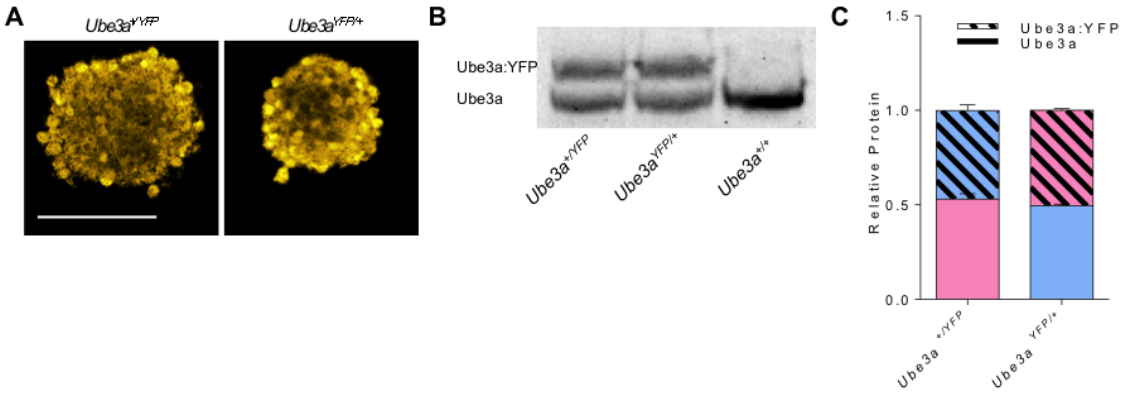


Figure 11. Biallelic expression of *Ube3a* is maintained in hippocampal stem/progenitor cell cultures. **A)** Immunofluorescence of Ube3a:YFP expression in neurospheres derived from hippocampus of *Ube3a*$^{+/YFP}$ and *Ube3a*$^{YFP/+}$. Scale Bar, 100µm. **B)** Western blot of hippocampal NSCs from *Ube3a*$^{+/+}$, *Ube3a*$^{+/YFP}$, and *Ube3a*$^{YFP/+}$ mice probed for Ube3a. **C)** Quantification of relative allelic expression of *Ube3a* in NSCs by western blot analysis.

Ube3a:YFP were equal (51:49; P = 0.749) in NSC-derived astrocytes at 16 days *in vitro* (DIV) (**Figure 12 A-B**). In NSC-derived neurons, paternal- and maternal-Ube3a:YFP

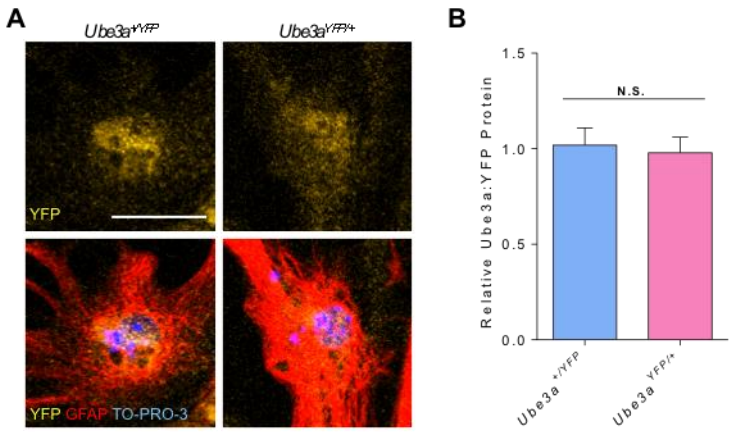


Figure 12. Biallelic expression of *Ube3a* in neural stem cell derived astrocytes. A) Immunofluorescence of Ube3a:YFP expression in astrocytes at 16 days post differentiation from NSCs derived from hippocampus of *Ube3a*$^{+/YFP}$ and *Ube3a*$^{YFP/+}$ mice. YFP, Ube3a:YFP; GFAP, glial fibrillary acid; TO-PRO-3 Iodide nuclear stain. Scale Bar, 100μm. **B)** Quantification of relative expression in differentiated astrocytes. n = 15 *Ube3a*$^{+/YFP}$ and 14 *Ube3a*$^{YFP/+}$. N.S., not significant.

were equal (54:46; P = 0.277) at 1 DIV; however, at subsequent time-points (4 DIV [61:39; P = 0.019]; 8 DIV [67:33; P < 0.001]; 16 DIV [86:14; P < 0.001]), maternal- and paternal-Ube3a:YFP levels diverged at similar rates (**Figure 13A-C**). During the course of the experiment, paternal-Ube3a:YFP levels decreased by 0.7-fold (P < 0.001); conversely, maternal-Ube3a:YFP levels increased by 1.7-fold (P < 0.001). Summation ofmaternal-Ube3a:YFP and paternal-Ube3a:YFP at each time point revealed that the total level of Ube3a:YFP remained relatively constant over the course of the experiment (P = 0.313 [**Figure 13C**]). Furthermore, at 16 DIV, paternal-Ube3a:YFP was still

detectable in the neurons, suggesting the imprint did not completely silence paternal-allelic expression; instead, it reduced its expression, consistent with prior observations [135]. Collectively, these data suggest that maternal-allelic expression increases during the acquisition of the imprint to maintain a constant level of total Ube3a protein in neurons of the central nervous system.

4.5 Pharmacological Inhibition of the *Ube3a-AS* Affects Both *Ube3a* Maternal- and Paternal-allelic Expression

Our observation of increased maternal-allelic expression during neurogenesis suggested the existence of a dosage compensating mechanism that regulated *Ube3a* expression. To examine whether such a mechanism existed, we reactivated paternal-allelic expression in primary cortical neurons by inhibiting expression of the *Ube3a-AS* via topotecan. Primary cortical neurons derived from $Ube3a^{YFP/+}$ and $Ube3a^{+/YFP}$ mice (P1) were cultured for 7 DIV and treated with topotecan. Paternal-Ube3a:YFP and maternal-Ube3a:YFP levels were then examined 4 days afterwards (11 DIV) by quantitative immunofluorescence imaging. The topotecan treatment reactivated the paternal allele; however, the level of reactivation was only 83% of the maternal allele in control neurons. Unexpectedly, topotecan treatment also increased maternal-Ube3a:YFP (1.2-fold relative to untreated maternal-Ube3a:YFP [P = 0.018]). The increased protein levels from each allele yielded a net 2-fold increase of Ube3a:YFP (**Figure 14A-B**). Further investigation of the data, however, revealed two populations of neurons in the untreated

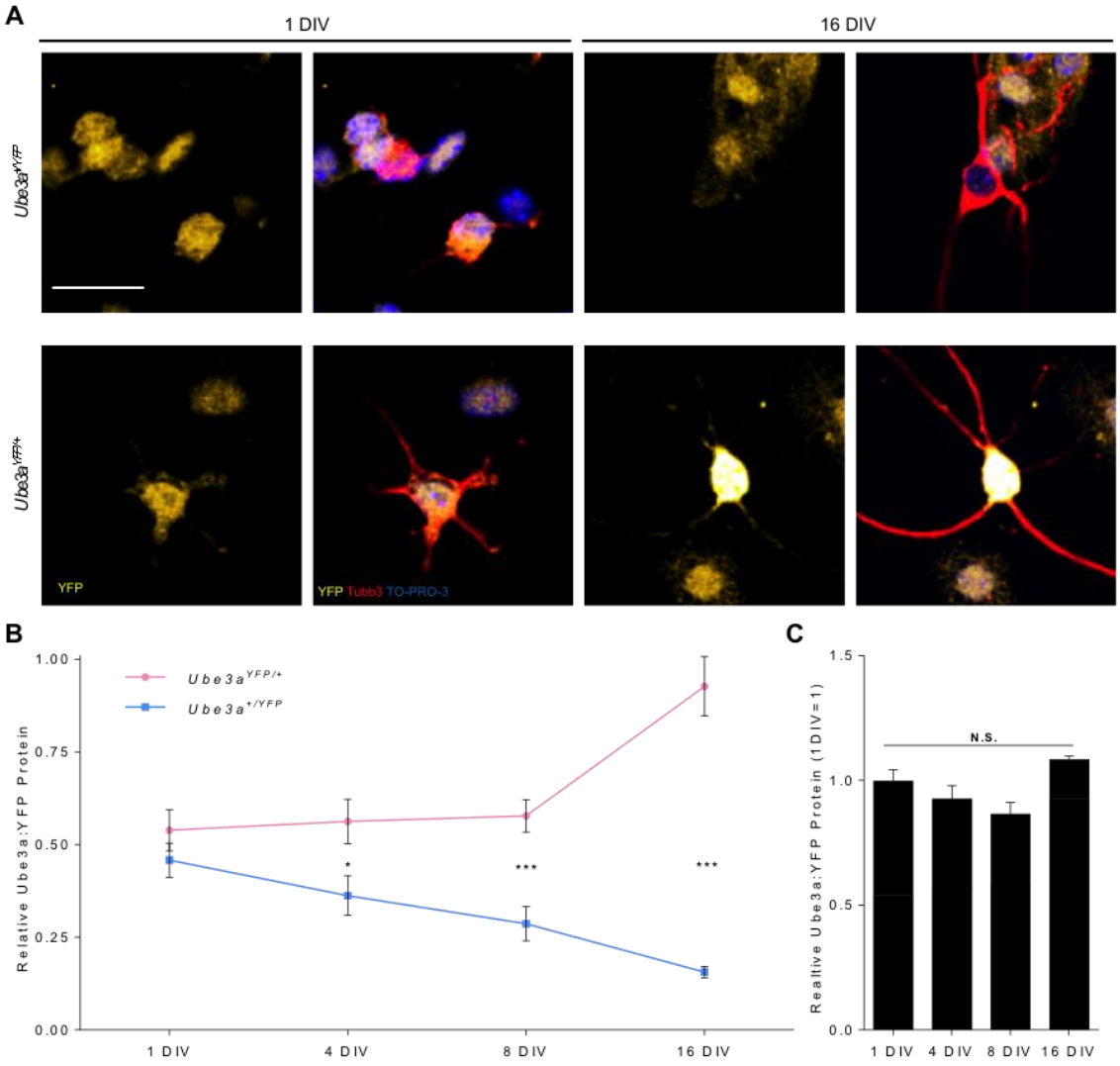


Figure 13. *Ube3a* maternal-allelic expression increases during neurogenesis. A) Immunofluorescence of allelic Ube3a^{YFP} in neurons differentiated from neural stem cell (NSC) cultures at 1 and 16 day(s) *in vitro* (DIV). YFP, Ube3a:YFP; Tubb3, β tubulin III; Topro, TO-PRO-3 Iodide nuclear stain. Scale bar, 25μm. **B)** Quantification of allelic Ube3a:YFP in neurons differentiated from NSCs at 1 DIV (n = 13 *Ube3a*$^{+/YFP}$, 14 *Ube3a*$^{YFP/+}$), 4 DIV (n = 14 *Ube3a*$^{+/YFP}$, 13 *Ube3a*$^{YFP/+}$), 8 DIV (n = 15 *Ube3a*$^{+/YFP}$, 15 *Ube3a*$^{YFP/+}$), and 16 DIV (n = 14 *Ube3a*$^{+/YFP}$, 13 *Ube3a*$^{YFP/+}$), relative to total Ube3a:YFP at 1 DIV. ** P < 0.01, *** P < 0.001 between alleles. **C)** Summation of total Ube3a:YFP in neurons differentiated from NSCs in **A** and **B**. N.S. not significant.

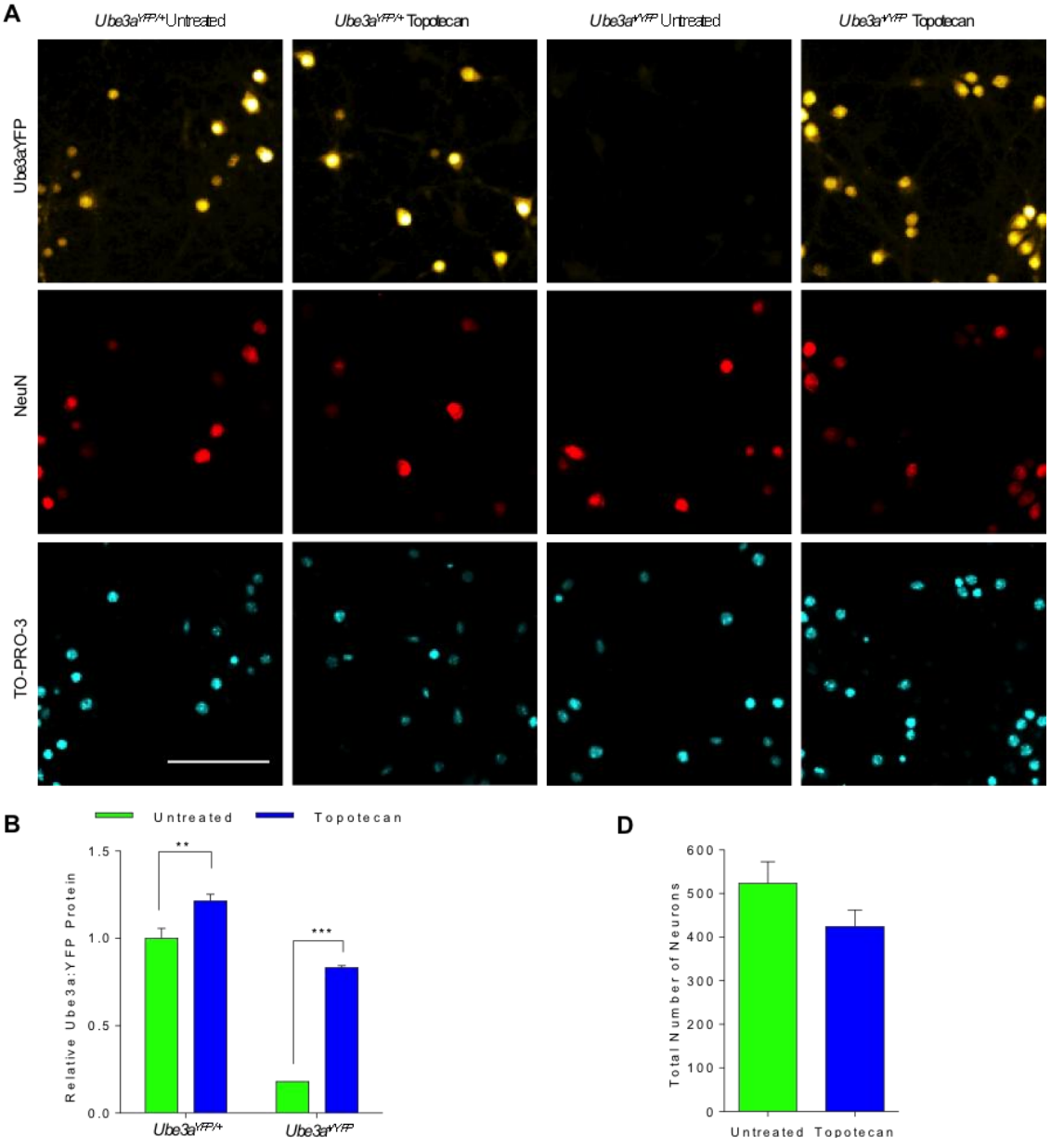


Figure 14. Pharmacological inhibition of the *Ube3a-AS* affects *Ube3a* maternal- and paternal-allelic expression. A) Immunofluorescent images of untreated topotecan treated primary neurons from *Ube3a*$^{+/YFP}$ and *Ube3a*$^{YFP/+}$ mice. NeuN, neuronal nuclei RBFOX3; TO-PRO-3 Iodide nuclear stain. 10x magnification. Scale bar, 100μm. **B)** Average relative Ube3a:YFP expression by well in 96 well plate. n = 10 wells *Ube3a*$^{YFP/+}$ untreated, 6 *Ube3a*$^{YFP/+}$ topotecan, 8 *Ube3a*$^{+/YFP}$ untreated, 8 *Ube3a*$^{+/YFP}$ topotecan. ** $P < 0.01$, *** $P < 0.001$, N.S. not significant. **C)** Frequency distribution plot of Ube3a:YFP intensities of individual cells as a percent of total neurons for *Ube3a*$^{YFP/+}$ untreated, *Ube3a*$^{YFP/+}$ topotecan, and *Ube3a*$^{+/YFP}$ topotecan. **D)** Total neuron counts of *Ube3a*$^{YFP/+}$ untreated, *Ube3a*$^{YFP/+}$ topotecan. Error bars = standard error of the mean.

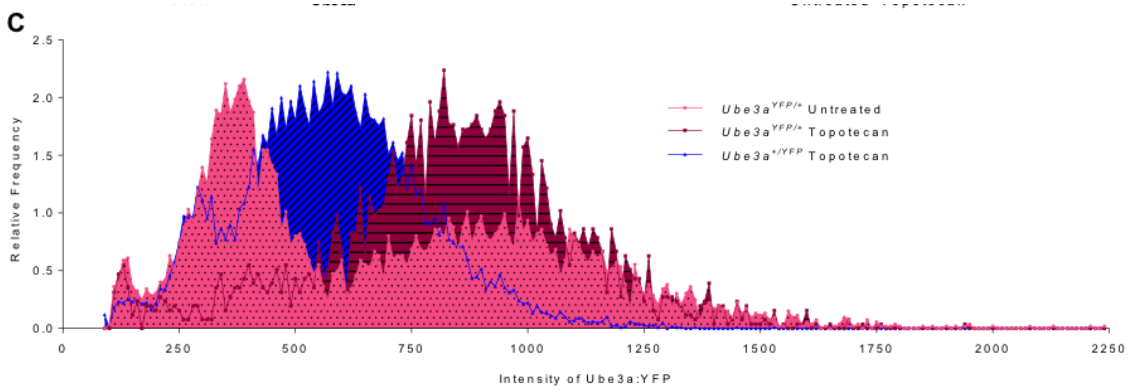


Figure 14. Continued

cultures that were distinguished by low and high maternal-Ube3a:YFP levels (Ashman's D>2 [**Figure 14C**]). The low maternal-Ube3a:YFP subpopulation was not present in the topotecan treated *Ube3a*$^{YFP/+}$ neuronal cultures, rather only the high maternal-Ube3a:YFP was present. Therefore, the increased maternal-Ube3a:YFP levels observed in the topotecan treated neurons was in fact due to loss of the low expressing subpopulation. Likewise, only one population (i.e., based on Ube3a:YFP level) of neurons was observed in the topotecan treated *Ube3a*$^{+/YFP}$ neuronal cultures; however, the neurons represented an intermediate subpopulation (Figure 15C). The effect of topotecan on the two subpopulations was not the result of neuronal cell death (P = 0.182 [**Figure 14D]**), indicating that the treatment only increased maternal-Ube3a:YFP in the low expressing subpopulation and not the high expressing subpopulation. To account for potential artifacts stemming from the Ube3a:YFP protein and address direct measurement of total Ube3a protein levels, the experiment was replicated using primary cortical neurons derived from wild-type mice. Again, the topotecan treatment increased

Ube3a protein levels by approximately 2-fold (**Figure 15A-B**). Consistent with prior observations, the paternal allele was reduced but not completely silenced *in vivo* or *in vitro* (~1.8 to 3.2% of the maternal allele, P < 0.001 [data not shown]). Taken together, these data show a complex interplay between the *Ube3a-AS* and *Ube3a* gene expression, suggesting the imprint may in fact regulate dosage of Ube3a in neurons.

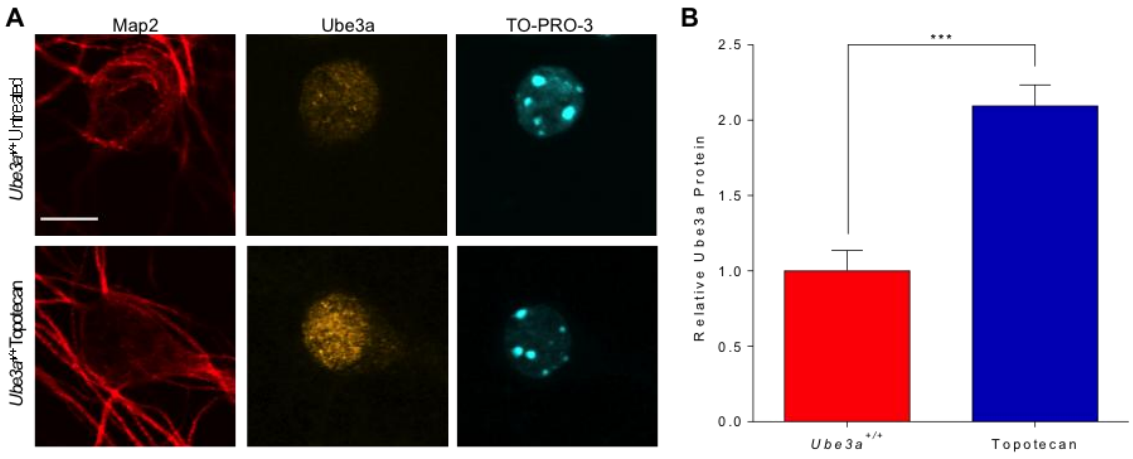


Figure 15. Pharmacological inhibition of the *Ube3a-AS* doubles Ube3a protein levels in neurons. A) Immunofluorescent images of primary cortical neurons derived from *Ube3a*$^{+/+}$ mice and *Ube3a*$^{+/+}$ neurons treated with topotecan. Map2, microtubule associated protein 2; TO-PRO-3 Iodide nuclear stain. Scale bar, 10μm. **B)** Relative Ube3a expression in primary neurons with Ube3a$^{+/+}$ set equal to 1. n = 37 *Ube3a*$^{+/+}$, 38 *Ube3a*$^{+/+}$ topotecan. *** P < 0.001.

4.6 Biallelic Expression of the Opossum *UBE3A* Gene Does Not Correlate with Increased UBE3A Protein Levels

Our observations of increased maternal-allelic expression during the acquisition of the imprint in neurons and elevated Ube3a protein levels in the brain suggested the imprint

did not evolve to negatively regulate Ube3a in neurons; however, the two-fold increase of Ube3a in neurons expressing both parental alleles implied the imprint may in fact have evolved to negatively regulate Ube3a. Therefore, we tested the hypothesis that the evolutionary acquisition of the *UBE3A* imprint correlated with decreased UBE3A

A

```
Human    -  MATACKRS-GEPQSDDIEASRMKRAAAKHLIERYYHQLTEGCGNEACTNEFCASCPTFLRMDNNAAAIKALELYKINAKLCDPHPSKKGASSAYLE-NSK - 098
Opossum  -  MATACKRSPGEPHSDDIEASRMKRAAAKHLIERYYHQLTEGCGNEVCTNEFCASCPTFLRMDNNAAAIKALELYKINAKLCDPHPSKKGTSSAYLENNSK - 100
Mouse    -  MATACKRSPGESQSEDIEASRMKRAAAKHLIERYYHQLTEGCGNEACTNEFCASCPTFLRMDNNAAAIKALELYKINAKLCDPHPSKKGASSAYLE-NSK - 099
            ******** **.:*:*******************************.*****************************************:****** ***

Human    -  GAPNNSCSEIKMNKK---GARIDFKDVTYLTEEKVYEILELCREREDYSPLIRVIGRVFSSAEALVQSFRKVKQHTKEELKSLQAKDEDKDEDEKEKAAC - 195
Opossum  -  GAHNNSCTDRKMNKKEVQGPRDDFKDVTYLTEEKVYEILELCREKEDFSPLIRVIGRVFSSAEALVQSFRKAKQHTKEELKSLQGKDEDKDEDEKEKAAC - 200
Mouse    -  GASNNS--EIKMNKK---EGK-DFKDVIYLTEEKVYEIYEFCRESEDYSPLIRVIGRIFSSAEALVLSFRKVKQHTKEELKSLQEKDEDKDEDEKEKAAC - 193
            ** *** _: *****    _:_***** ********** *:***_**:*********:******** ****.************_***************

Human    -  SAAAMEEDSEASSSRIGDS-SQGDNNLQKLGPDDVSVDIDAIRRVYTRLLSNEKIETAFLNALVYLSPNVECDLTYHNVYSRDPNYLNLFIIVMENRNLH - 294
Opossum  -  SAAAMEEDSEASSSRIGDSNSQGDNNLQKLGPDEVSVDIEAIRRVYNRLLSNEKIETAFLNALVYLSPNVECDLTYHNVYSRDPNYLNLFIIVMENGNLH - 300
Mouse    -  SAAAMEEDSEASSSRMGDS-SQGDNNVQKLGPDDVTVDIDAIRRVYSSLLANEKLETAFLNALVYLSPNVECDLTYHNVYTRDPNYLNLFIIVMENSNLH - 292
            ***************:*** ******:******:*:***:******.  **:***:************************:***************_***

Human    -  SPEYLEMALPLFCKAMSKLPLAAQGKLIRLWSKYNADQIRRMMETFQQLITYKVISNEFNSRNLVNDDDAIVAASKCLKMVYYANVVGGEVDTNHNEEDD - 394
Opossum  -  SPEYLEMALPLFCKAMSKLPLAAQAKLVRLWSKYSADQIRRMMETFQQLITYKVISNEFNSRNLVNDDDAIVAASKCLKMVYYANVVGGEVDTDHNEEED - 400
Mouse    -  SPEYLEMALPLFCKAMCKLPLEAQGKLIRLWSKYSADQIRRMMETFQQLITYKVISNEFNSRNLVNDDDAIVAASKCLKMVYYANVVGGDVDTNHNEEDD - 392
            ***************.**** **.**:******.*************************************************************:***:****:*

Human    -  EEPIPESSELTLQELLGEERRNKKGPRVDPLETELGVKTLDCRKPLIPFEEFINEPLNEVLEMDKDYTFFKVETENKFSFMTCPFILNAVTKNLGLYYDN - 494
Opossum  -  EEPIPESSELTLQELLGEERRNKKGPRVDPLETELGVKTIDCRKPLIPFEEFINEPLNDVLEMDKDYTFFKVETENKFSFMTCPFILNAVTKNLGLYYDN - 500
Mouse    -  EEPIPESSELTLQELLGDERRNKKGPRVDPLETELGVKTLDCRKPLISFEEFINEPLNDVLEMDKDYTFFKVETENKFSFMTCPFILNAVTKNLGLYYDN - 492
            ****************:*******************:*******.**********:*****************************************

Human    -  RIRMYSERRITVLYSLVQGQQLNPYLRLKVRRDHIIDDALVRLEMIAMENPADLKKQLYVEFEGEQGVDEGGVSKEFFQLVVEEIFNPDIGMFTYDESTK - 594
Opossum  -  RIRMYSERRITVLYSLVQGQQLNPYLRLKVRRDHIIDDALVRLEMIAMENPADLKKQLYVEFEGEQGVDEGGVSKEFFQLVVEEIFNPDIGMFTYDESTK - 600
Mouse    -  RIRMYSERRITVLYSLVQGQQLNPYLRLKVRRDHIIDDALVRLEMIAMENPADLKKQLYVEFEGEQGVDEGGVSKEFFQLVVEEIFNPDIGMFTYDEATK - 592
            ****************************************************************************************************:**

Human    -  LFWFNPSSFETEGQFTLIGIVLGLAIYNNCILDVHFPMVVYRKLMGKKGTFRDLGDSHPVLYQSLKDLLEYEGNVEDDMMITFQISQTDLFGNPMMYDLK - 694
Opossum  -  LFWFNPSSFETEGQFTLIGIVLGLAIYNNCILDVHFPMVVYRKLMGKKGTFRDLADSHPVLYQSLKDLLEYEGSVEDDMMITFQISQTDLFGNPMMYDLK - 700
Mouse    -  LFWFNPSSFETEGQFTLIGIVLGLAIYNNCILDVHFPMVVYRKLMGKKGTFRDLGDSHPVLYQSLKDLLEYEGSVEDDMMITFQISQTDLFGNPMMYDLK - 692
            ************************************************.*****************.************************

Human    -  ENGDKIPITNENRKEFVNLYSDYILNKSVEKQFKAFRRGFHMVTNESPLKYLFRPEEIELLICGSRNLDFQALEETTEYDGGYTRDSVLIREFWEIVHSF - 794
Opossum  -  ENGDKIPITNENRKEFVSLYADYILNKSVEKQFKAFRRGFHMVTNESPLKYLFRPEEIELLICGSRNLDFQALEETTEYDGGYTRDSLIIREFWEIVHSF - 800
Mouse    -  ENGDKIPITNENRKEFVNLYSDYILNKSVEKQFKAFRRGFHMVTNESPLKYLFRPEEIELLICGSRNLDFQALEETTEYDGGYTRESVVIREFWEIVHSF - 792
            *****************.**:*********************************************************************:*::***********

Human    -  TDEQKRLFLQFTTGTDRAPVGGLGKLKMIIAKNGPDTERLPTSHTCFNVLLLPEYSSKEKLKERLLKAITYAKGFGML - 872
Opossum  -  TDEQKRLFLQFTTGTDRAPVGGLGKLKMIIAKNGPDTERLPTSHTCFNVLLLPEYSSKEKLKERLLKAITYAKGFGML - 878
Mouse    -  TDEQKRLFLQFTTGTDRAPVGGLGKLKMIIAKNGPDTERLPTSHTCFNVLLLPEYSSKEKLKERLLKAITYAKGFGML - 870
            ***************************************************************************
```

B

	Conservative	Semi-conservative	Non-conserved	Insertion/Deletion	Total Mismatches	Percent Identity
Mouse versus Human	17	6	9	4	36	95%
Opossum versus Human	16	8	5	6	35	96%
Mouse versus Opossum	27	10	11	8	56	93%

Figure 16. The mouse, human, and opossum Ube3a/UBE3A protein sequences are highly conserved. A) Alignment of Ube3a orthologs among human, mouse, and opossum. "*", match. ":", conserved amino acid (AA) substitution. "." semi-conserved AA substitution. " ", non-conserved AA substitution or insertion/deletion. "_", call not made because of mismatch between all three species of varying conservation. B) Summary of amino acid mismatches and percent identity between species.

protein levels in the brain. As an evolutionary outlier, we used the gray, short-tailed opossum (*Monodelphis domestica*), a metatherian mammal lacking an orthologous region of the human 15q11-13 imprinted region. The opossum UBE3A protein sequence is highly similar (36-56/870; 93-95% [**Figure 16**]) to the mouse and human UBE3A/Ube3a protein sequences and thus considered a suitable evolutionary outlier for this study. Furthermore, *UBE3A* is biallelically expressed in marsupials [105]. We examined UBE3A steady state protein levels in the cortex, hippocampus, heart, and lung. Similar to the mouse (**Figure 7B**), UBE3A protein levels were elevated in the opossum brain relative to the peripheral organs ($P = 0.001$ [**Figure 17A]**). Direct comparisons between UBE3A/Ube3a protein levels in opossum and mouse brain, however, revealed a 1.7 fold increase of Ube3a in the mouse brain relative to the opossum brain ($P = 0.008$ [**Figure 17B]**). To account for possible discrepancies in the affinity of the UBE3A antibody between the two species, we compared the relative ratios of UBE3A/Ube3a protein among brain and peripheral organs. The ratio of Ube3a protein in the mouse brain was greater than the opossum brain (**Figure 17C**), suggesting that the imprinted expression of *Ube3a* in the mouse brain produces more Ube3a protein than biallelic expression in the opossum brain. Collectively, these data suggest that the *UBE3A* imprint did not evolve to negatively regulate UBE3A protein levels in the brain.

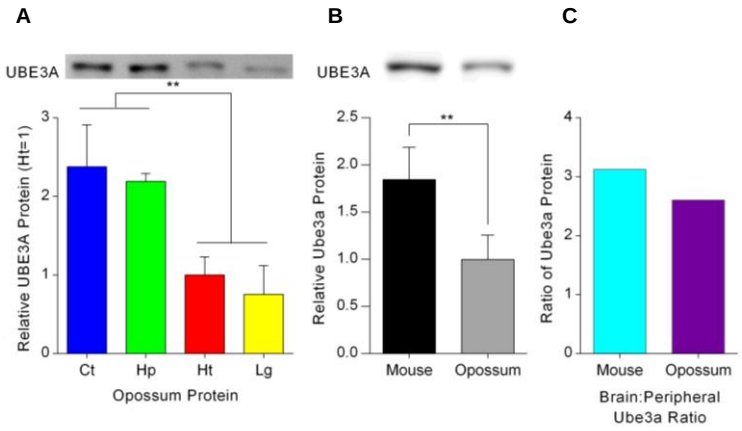


Figure 17. *UBE3A* allelic transcript levels in brain and UBE3A protein levels in brain and peripheral tissues of opossum compared to mouse. **A)** Normalized UBE3A protein levels isolated from wild-type opossum tissues. Levels are shown as a ratio of expression in tissues relative to heart, normalized by Ponceau total protein stain. n = 4. *** P < 0.001. Ct, cortex; Hp, hippocampus; Ht, heart; Lg, lung. **B)** Normalized Ube3a/UBE3A protein levels isolated from wild-type mouse and opossum cortex. Levels are set to average UBE3A = 1, normalized by Ponceau total protein stain. n = 4. .** P < 0.01. **C)** Ratio of average Ube3a/UBE3A expression in brain relative to peripheral tissues in mouse and opossum.

4.7 The *Ube3a-Snrpn* snoRNA Clusters Are Functionally Coupled to the *Ube3a-AS*

4.7.1 The Mouse Ube3a-Snrpn snoRNAs Evolved from Snord119

Given evolutionary constraint of the *Ube3a* imprint in neurons and the absence of any evidence indicating that it functions to negatively regulate Ube3a protein levels, we then hypothesized that the *Ube3a-AS* may in fact be under selection and functional. To determine a functional link between the *Ube3a-Snrpn* snoRNAs and the *Ube3a-AS*, we first examined the evolutionary history of the snoRNA genes/clusters: *Snord107* (1 copy), *Snord64* (1 copy), *Snord116* (27 copies, 10 unique), and *Snord115* (136 copies, 56 unique) (**Appendix D**). Alignment of the snoRNAs revealed that each contained

canonical C, D, C`, and D` boxes (**Figure 18A**). Phylogenetic analysis of the mouse *snoRNAs*, their human orthologs, and the *Snord119* gene (mouse, human, and opossum) indicated that the snoRNAs in the region originated from *Snord119* [140] (**Figure 18B).** This is consistent with the evolutionary history of the orthologous cluster in humans (**Figure 18B and Appendix E [12]**). The evolutionary history also appeared to correlate with their organization along the chromosome (5`-*Snrpn* to 3`-*Ube3a* [**Appendix E**]).

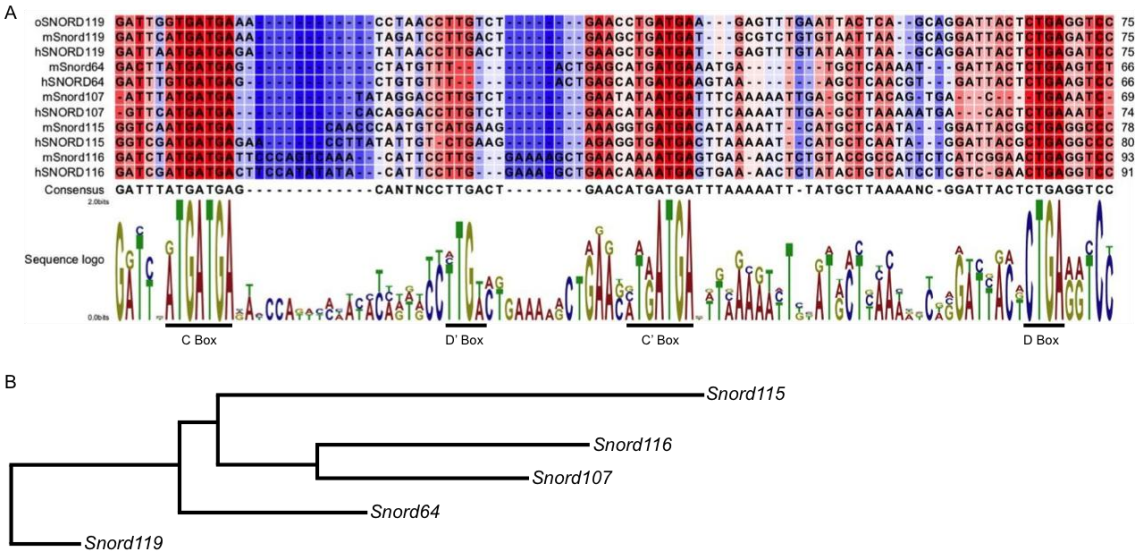


Figure 18. The human 15q11-q13 and mouse 7C *SNORD/Snord* gene clusters arose from *Snord119*. **A**) Representative sequence alignments of mouse (m) and human (h) *SNORDs - 64, 107, 115,* and *116* - to the ancestral sequence, *SNORD119*, from mouse, human, and opossum (o). Consensus sequence was extracted based on base conversation (red shading is high and blue shading is low) between all sequences. Sequence logo displays all bases represented at each position, stacked vertically with their heights representing relative base frequency: A = adenine, red; T = thymine, green; C = cytosine, blue; and, G = guanine, gold. Conserved characteristic C/D box snoRNA sequences are underlined C, D', C', D. **B**) Schematic representation of tree clustering analysis of all human and mouse SNORD sequences (Appendix E). Horizontal distance is representative of relative sequence divergence from branch points.

The presence of identical/high percent identity *Snord115* and *Snord116* gene copies and the lack eutherian conservation at the host-gene exons (**Appendix F**) and intervening sequences suggested the snoRNAs have been under selection and, perhaps, homogenized.

4.7.2 The Snurf-Snrpn, snoRNA host-genes/snoRNAs and Ube3a-AS RNAs Are Likely Transcribed as a Single Polycistronic Transcript

We next analyzed ChIP-seq and RNA-seq data (described above) to examine whether the *Snurf-Snrpn*, *Snord* gene clusters, and *Ube3a-AS* transcripts were assembled into a single or multiple individual transcript(s). Consistent with prior studies, we did not detect features (e.g., trimethylated histone 3 lysine 4, acetylated histone 3 lysine 9, and RNA polymerase II enriched regions) 3` of *Snurf-Snrpn* that were consistent independent promoter being used to transcribe the *snoRNAs* and *Ube3a-AS*, further supporting the notion that the genes in the *Snrpn-Ube3a* region are transcribed as an single unit (**APPENDIX G** [12, 27]).

4.7.3 The Snurf-Snrpn, snoRNA host-gene Exons, and Ube3a-AS Exhibit Brain-Specific Alternative Splicing Events and Polyadenylation Patterns

Analysis of the RNA-seq data revealed extensive alternative splicing (canonical 5`-GT-AG-3`) of the *Snurf-Snrpn* coding and upstream untranslated region (UTR) exons in the brain; however, there were no spliced transcripts detected between *Snurf-Snrpn* and the

downstream *Snord64* host-gene exons (**APPENDIX H-I**). We did detect spliced transcripts among the *Snord116* host-gene exons and *Ipw* and between the 3` exon of *Ipw* and a single 5 ` *Snord115* host-gene exon (**APPENDIX J-L**). At the 3` end of the *Snord115* cluster, we also detected spliced transcripts between the *Snord115* host-gene exons and the 5` end of the *Ube3a-AS* (**Figure 19** and **APPENDIX M**). In the *Ube3a-AS*, we detected numerous splicing events, including a single continuous transcript of 7 exons that extended from the 3` end of the *Snord115* cluster to a region approximately 30 kb 5` of the *Ube3a* transcriptional start-site (TSS). Most of the *Ube3a-AS* exons were alternatively spliced and, except for one 5` - 3` splice event in *Ube3a* exon 6, primarily located in intronic regions at the 3` end of *Ube3a* (**Figure 19** and **APPENDIX N**).

Analysis of strand-specific polyA-seq data identified numerous brain specific polyadenylation (polyA) signals in the *Ube3a-Snrpn* region (**Figure 19** and **APPENDIX H-O**). Most of the brain-specific polyAs were present on the minus strand and thus represented polyA tails of RNAs transcribed in the same orientation as the *Snurf-Snrpn*, *snoRNA*, and *Ube3a-AS* RNA transcripts. There were numerous polyAs located within the *Snord116* cluster (minus strand) and one polyA located at the 5` end of the *Snord115* cluster (minus strand). The remaining brain-specific polyAs were located in *Ube3a* introns 4 (plus and minus strand) and 11 (minus strand), the *Ube3a* TSS (minus strand), and a region 5` of the *Ube3a* TSS. The polyAs located upstream of *Ube3a* were immediately 3` of the distal *Ube3a-AS* exon (**APPENDIX M-O**). Collectively, these data suggest polyadenylation decouples the *Snurf-Snrpn* and *Snord* host-gene transcripts (i.e., *Snord64*, *Snord116* and *Snord115*) and splicing excises functional snoRNA species

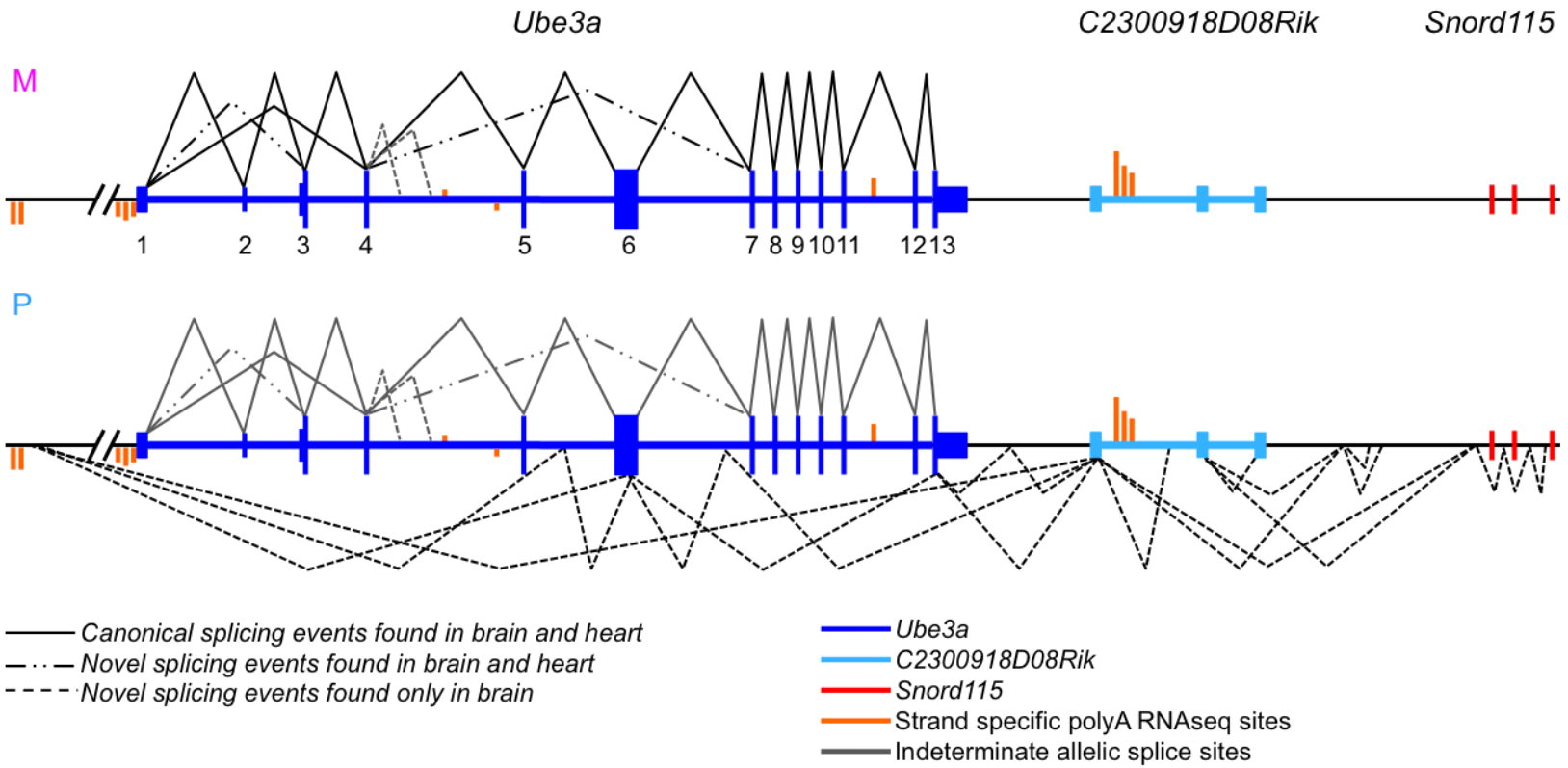


Figure 19. Alternative splicing patterns of Ube3a and the Ube3aAS. The schematic diagrams depicts alternative splicing patterns detected in brain and heart by RNAseq and RT-PCR on the maternal (M) and paternal (P) alleles. Numbers represent exons of Ube3a. Line composition differentiates tissue specificity of splicing and line color differentiates allelic specificity. Lines above gene depictions represent 5' → 3' splicing from left to right on the plus strand of DNA and lines below the genes reflect 5' → 3' splicing from right to left on the minus strand of DNA. Orange bars represent strand specific polyadenylation sites (APPENDIX X), their heights are proportional signal strength by RNAseq, and bars above the line are specific to the plus strand and bars below the line are specific to the minus strand. Schematic is not to scale but spans approximately 170 kilobases (kb) and the hash marks represent a gap of approximately 21kb

in the brain. The *Snord115* host-gene exon(s), however, are coupled to the *Ube3a-AS*, which is processed into multiple alternatively spliced, polyadenylated transcripts, suggesting it has additional functions aside from silencing *Ube3a* paternal-allelic expression in neurons.

4.8 The *Ube3a* Gene Is Extensively Alternatively Spliced in the Brain

Analysis of the RNA-seq data also confirmed the expression in both the brain and heart of the known *Ube3a* transcript variants representing transcript variants 1 and 3, which are not distinguishable at the 5` exons, and transcript variant 2. In the brain, we also detected a novel transcript that spliced from exon 4 to a region in intron 4, which was located 5` of the brain-specific polyA signals described above (plus strand) (**Figure 19** and **APPENDIX N-O**). Although we were unable to determine the parent of origin of this novel transcript, prior studies have shown that the paternal *Ube3a* allele transcribes a 5` truncated transcript that terminates in intron 4 [20, 31].

We next used RT-PCR to confirm the *Ube3a* transcript variants and identify additional transcripts that were undetectable in the RNA-seq data. Using combinations of primers specific to each *Ube3a* exon (**APPENDIX P**), we confirmed the presence of the known *Ube3a* splice variants (sense and antisense) and the novel 5` truncated transcript. The RT-PCR also showed that the *Ube3a* transcript variant 1/3 was the most abundant transcript in the brain. We also detected novel alternatively spliced transcripts that reflected splicing events between exons 1 and 3 and between exons 4 and 7 in both brain

and heart. The 1-3 splice event preserves the Kozak consensus sequence and start site of *Ube3a* transcript variant 2. The 4-7 splice event lacked the translational start codon of transcript variants 1/3; however, the 5` end of the transcript contained the 5` UTRs associated with transcript variant 1/3 and a novel Kozak consensus sequence preceding an in frame methionine. This variant was more abundant in heart than in brain, and, if translated, would produce a truncated protein containing the HECT domain (**APPENDIX P**). Both novel splicing events were confirmed by Sanger sequencing.

Since the novel 4-7 splice variant was predicted to skip the exon deleted (exon 5) in the *Ube3a* knockout mouse model, we next used a TaqMan RT-PCR assay to analyze brain- and heart-RNA (mRNA and toRNA) isolated from wild-type and *Ube3a* maternal- and paternal-deficient mice ($Ube3a^{+/+}$, $Ube3a^{-/+}$, $Ube3a^{+/-)}$. In the heart samples of both $Ube3a^{+/-}$ and $Ube3a^{-/+}$ mice, there was approximately a 50% reduction in toRNA and mRNA, which was anticipated. Analysis of the cortical samples, however, revealed relatively equal amounts of toRNA and mRNA among all three genotypes, suggesting the presence of either a novel promoter or an unknown splicing pattern not involving exon 5. Additional RT-PCR analyses in each genotype confirmed the presence of the 4-7 splice variant. A novel splice variant between exons 4 and 6 was also detected, but the transcript variant was only detected in samples isolated from the $Ube3a^{-/+}$ and $Ube3a^{+/-}$ mice, suggesting it was an atypical splicing event (**APPENDIX P**).

Visual examination of the RNA-seq data revealed that there was a disproportionate number of reads aligning to the *Ube3a* introns in the brain relative to the heart. Genome-wide analysis of exon/intron read-depth ratios revealed the brain samples had increased intronic read-depth ratios (Wilcoxon signed rank-test), suggesting the reads reflected an abundance of immature RNA transcripts in the brain relative to the heart. Further analysis of informative variants in the intronic reads revealed that most were derived from the paternal *Ube3a* allele but some reflected maternal-allelic expression. We also observed a "saw-tooth" pattern of read-depth ratios across some of the introns, particularly intron 4, suggesting increased levels of alternative splicing events involving the flanking exons in the brain, which was consistent with the number of alternatively spliced Ube3a sense and antisense transcripts identified in the brain (**Figure 20** and **APPENDIX P)** [141].

Collectively, these data demonstrate the existence of novel alternatively spliced *Ube3a* transcript variants (heart and brain), which are also expressed in the *Ube3a* knockout mouse model. Furthermore, these data show that different rates of *Ube3a* RNA processing between the heart and brain, which likely reflects different patterns of alternative splicing events between the two tissues.

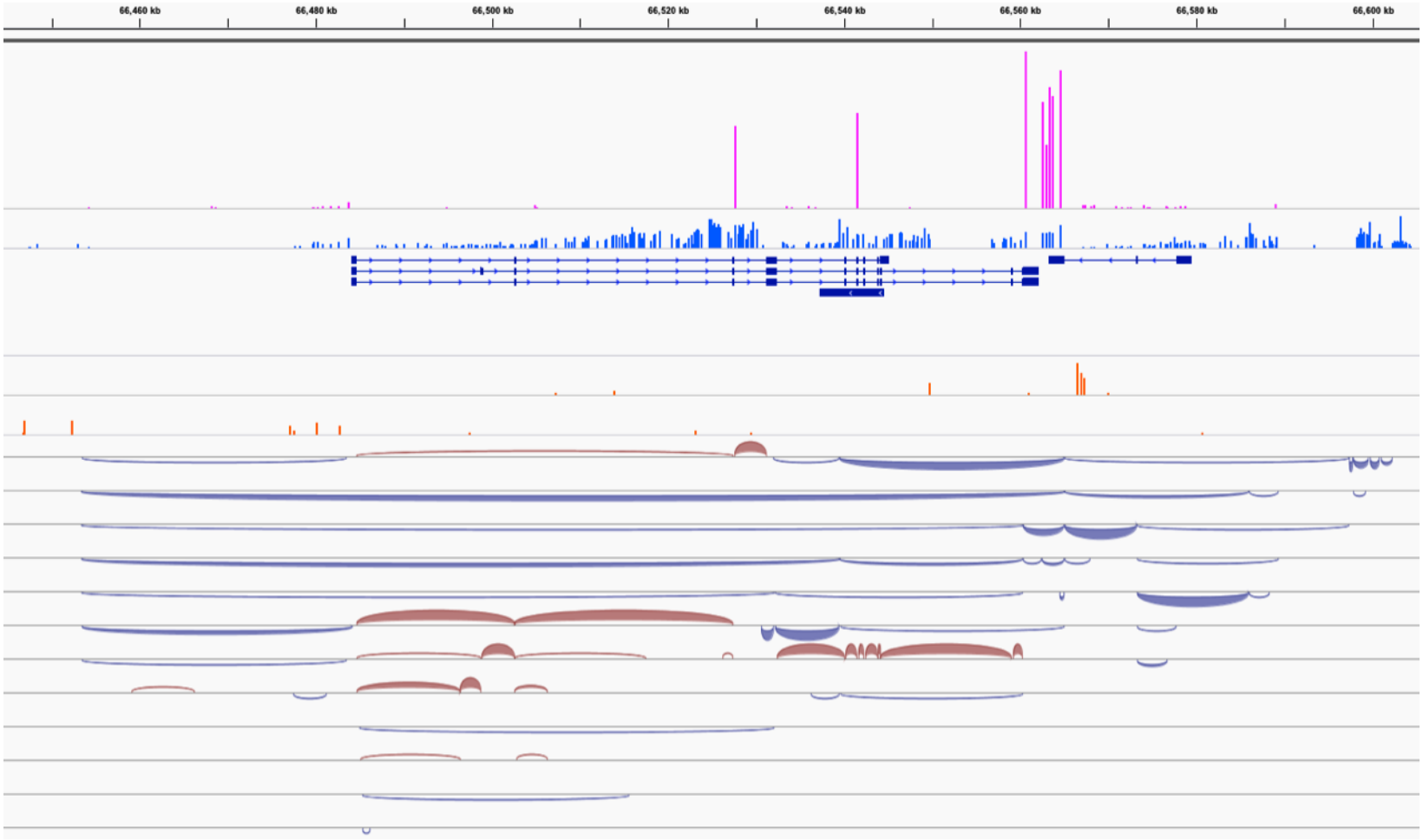


Figure 20. RNAseq coverage of informative allelic single nucleotide polymorphisms across *Ube3a* in the brain. Image depicts the relative coverage of maternal (pink bars) and paternal (blue bars) of allelic single nucleotide polymorphisms (SNPs) across the introns of *Ube3a*. The positive slope of predominately paternal SNPs suggests transcription of the minus strand consistent with Ube3a-AS expression. Enrichment of these same SNPs flanking exons suggests alternative splicing at these sites. Orange bars represent relative intensity of strand specific (plus strand, top; minus strand bottom) polyadenylation sites. Red arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the plus strand. Blue arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the minus strand.

5. CONCLUSIONS

Loss-of-function or dysregulated expression of *UBE3A* in the brain has devastating consequences, which strongly suggest that *UBE3A* is a dosage sensitive gene. It is thus perplexing that the *UBE3A* gene is imprinted in neurons where it is critically important; unless, the imprint evolved as an important mechanism to regulate the *UBE3A* expression levels in neurons. Indeed a gene dosage/gene expression model has been proposed to explain the evolution of genomic imprinting at some loci [142, 143]. Most studies to date involving UBE3A have focused on what function UBE3A has in neurons; few studies, if any, have addressed why *UBE3A* is imprinted in neurons. In this dissertation, it was hypothesized that genomic imprinting of *UBE3A* evolved as a mechanisms to negatively regulate UBE3A levels in neurons. The data presented here, however, rejects this hypothesis. Based on our findings, we propose a novel hypothesis in which the imprint evolved to facilitate neuron specific co-expression of the *UBE3A* and *SNORD115-SNORD116* genes and, perhaps, regulate the expression of neuron specific UBE3A isoforms. Although further studies are necessary to confirm this hypothesis, the findings presented here significantly advance our understanding of why *UBE3A* is specifically imprinted in neurons.

We show a non-linear relationship between the number of active *Ube3a* alleles and the amount of total Ube3a protein produced in the cell (**Figure 21**). Independent of the number of expressed alleles, the *Ube3a* gene is expressed at higher levels in neurons of

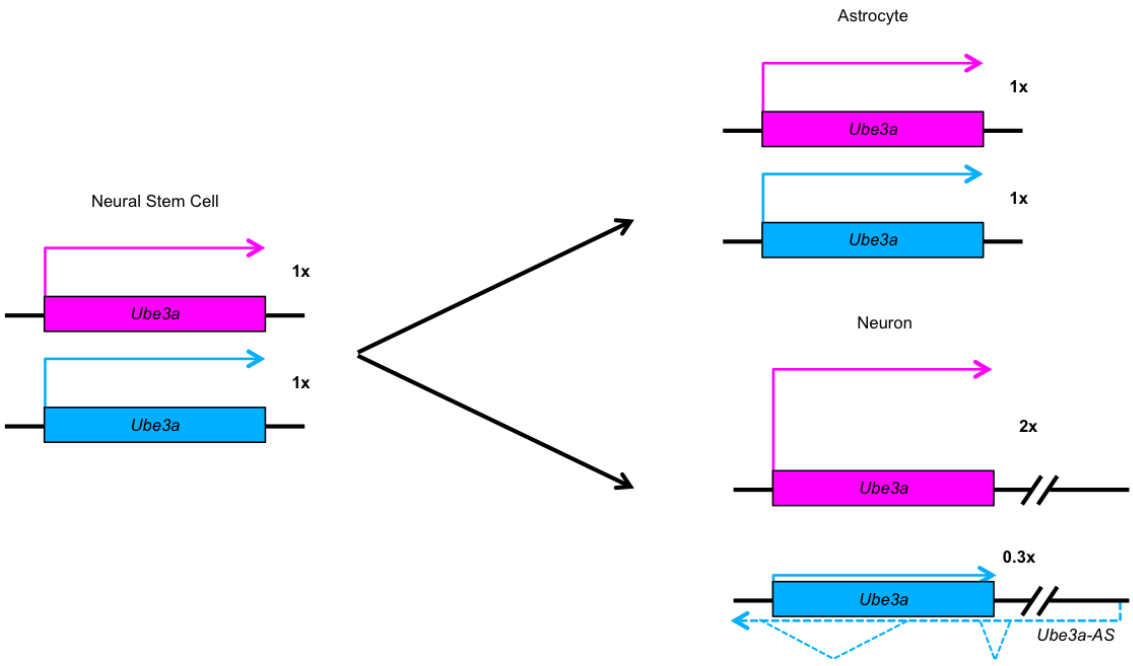


Figure 21. The number of active functional alleles of Ube3a does not correlate with expression levels in neuron compared to non-imprinted cell types. Ube3a is biallelically expressed in neural stem cells (NSCs) (colored right angle arrows, 1x). Upon differentiation (black arrows) NSCs give rise to astrocytes and neurons. Astrocytes biallelically express *Ube3a* and do not appear to change total protein levels from NSCs. Conversely neurons express only the maternal allele, yet total Ube3a levels increase during the course of differentiation. This leads to increased maternal expression, but relatively constant expression of Ube3a in developing neurons (2x maternal 2.3x total). Furthermore this unexpected expression pattern of maternal *Ube3a* coincides with expression of the paternal Ube3a-AS which is expressed across the length of *Ube3a* and exhibits alternative splicing (dashed blue arrow). This suggests that the *Ube3a* imprint did not evolve strictly to regulate Ube3a dosage. Angled arrow heights reflect relative levels of expression. Pink = maternal and Blue = paternal.

appears to be evolutionarily conserved. The preservation of this phenomenon supports current beliefs that Ube3a plays a critical role in neural function and synapse formation [129-136]. Interestingly, although increased expression of Ube3a in neurons is conserved between mouse and opossum, its total dosage does not appear to be similarly restricted. This is further complicated by the knowledge that, at least in humans, the expression of *UBE3A* is tightly regulated. This suggests that although total Ube3a/UBE3A protein levels have some critical level in the cell, the levels may be

dictated to some degree by the expression of *Ube3a/UBE3A* relative to other genes in the cell, possibly one of its theorized targets or cofactors discussed previously.

The observations of increased maternal-allelic expression during neurogenesis and the inability of Ube3a to negatively regulate biallelic expression in neurons suggest that an independent mechanism evolved to compensate for the loss of paternal-allelic expression in neurons. This postulation is supported if we simultaneously consider the evolution of the *Ube3a-Snrpn* region as it exists in placental mammals. The formation of the locus was due to a chance rearrangement that was most likely not immediately detrimental to the fitness of the organism; however, the ability of some individuals/organisms to regain a level of *Ube3a* expression consistent with levels that existed previous to the imprint provided some sort of competitive benefit. Once this was accomplished through the establishment of the imprint, the need for autoregulation of the gene product was no longer needed as over expression of the gene product was highly unlikely secondary to the selective pressures preserving the imprint and expression of paternal genes to overall organism fitness (see below) and under expression of the gene would likely result in offspring that would not be fit to reproduce.

Our observations that *Snurf-Snrpn*, the *Snord* snoRNA gene clusters, and the *Ube3a-AS* are expressed as a single polycistronic transcript and that the *Ube3a-AS* is transcriptionally coupled to the *Snord115* host-gene exons and processed suggests that the *Ube3a-AS* is either a functional component of the *Snord115* locus, functions as an

independent RNA, or a protein-coding transcript. Previous studies have shown that the *Snord115* host gene exons form an RNA cloud that is restricted to the *Ube3a-Snrpn* locus[116, 117]. The transcriptional link that we demonstrate between the *Ube3a-AS* and the *Snord115* host-gene transcript suggests that the *Ube3a-AS* functions as the anchor for the localization of the RNA cloud. This would be supported by the extensive coverage of the *Ube3a* introns seen exclusively from the paternal allele. Alternatively, the *Ube3a-AS* could simply act as a 3' UTR for *Snord115.* Future studies involving the analysis of the Snord115 host gene RNA cloud would need to be done using a *Ube3a-AS* knockout model or in cells treated with topotecan in order to define the fundamental role of the *Ube3a-AS*. The idea that the *Ube3a-AS* has an independent function is supported in the study reported by *Meng et al*. (2014) [18]. Insertion of a termination signal in the 3` UTR of Ube3a (*Ube3a-ATS* knockout mouse [KO]) partially reactivates *Ube3a* paternal-allelic expression in the brain and improves, but does not fully restore, the neurological deficits observed in the AS mouse model ($Ube3a^{-/+}$). The authors also showed increased expression of the imprinted *Necdin* gene, which is located in the 7c imprinted region but approximately 3000 kb downstream of *Ube3a*. Although the relationship between the *Ube3a-AS* and *Necdin* is unclear, it is tempting to speculate that the *Ube3a-AS*, or an imprint dependent *Ube3a* isoform negatively regulates *Necdin* expression in the brain. Finally, based on current data, we cannot exclude the possibility that some component(s) of the *Ube3a-AS* transcript may contain protein coding sequences. The extent of alternative isoforms predicted in our study made full evaluation of this possibility prohibitive. Furthermore, recent discoveries of small

protein coding genes (e.g., less than 40 amino acids) prove the existence of small open reading frames, which were previously believed to be non-functional [144-146]. Nevertheless, our findings support the notion that the imprint did not simply evolve as a mechanism to down-regulate the level of UBE3A in neurons.

Based on our findings, we propose that genomic imprinting of *SNURF-SNRPN-SNORDs-UBE3A-AS* and *UBE3A* evolved to allow co-expression of these genes in neurons. This idea is consistent with the Complementation Hypothesis proposed by Kaneko-Ishin *et al.* [147]. In the Complementation Model, genomic imprinting evolved as a mechanism to maintain the expression of at least one allele of two genes that compete for transcriptional elements/resources. The genes may be functionally linked or not; however, they do rely on the same genetic elements to allow or regulate expression of their respective genes. In this model, biallelic expression of one gene (e.g., Gene-A) outcompetes another gene (Gene-B) for elements that both require for transcription; as a result, Gene-B is inactivated on both alleles. Genomic imprinting of the regulatory element in one parental germ-line (i.e., the creation of a differentially methylated region or imprinting control element) would inhibit the expression of Gene-A on the allele carrying the germ-line modification. As a result, Gene-B would be expressed on this chromosome. Indeed, expression of *Snurf-Snrpn*, *Snord snoRNAs*, and *Ube3a-AS* is negatively regulated by the methylation of the PWC-ICE on the maternal chromosome [148]. Loss of the paternal methylation at the PWS-ICE results in biallelic expression of the polycistronic transcript and repression of both parental *Ube3a* alleles. There is no

known phenotype associated with biallelic expression of the polycistronic transcript as paternal uniparental disomy results in Angelman syndrome, not a unique disease; however, decreased or increased *Ube3a* expression is detrimental. Thus, loss of expression of the polycistronic transcript on the maternal chromosome *via* genomic imprinting protects the maternal *Ube3a* allele from being silenced. Based on our observations, we propose that the Complementation Model likely drove the evolution of imprinting of 15q11-q13 and the orthologous regions in other placental mammals. If expression of the *Ube3a-AS* also gives rise to imprint dependent transcripts of *Ube3a* or itself (e.g., through alternative splicing events), there may be other advantages resulting from the creation of an imprinted region, as seen at the GNAS locus. Further studies are necessary to demonstrate that the *Ube3a-AS* expands the repertoire of functionally meaningful Ube3a isoforms.

Furthermore, our understanding of the diseases associated with 15q11-q13 support that the imprint evolved to allow co-expression *UBE3A* and the *SNORD115-SNORD116* genes. Many attempts to produce Prader-Willi mouse models failed because loss of paternal expression in this region led to failure to thrive, poor suckling, and soft palate defects, leading to neonatal mortality [149]. Conversely, maternal inheritance of these same defects showed no such phenotype. This and other observations suggest the existence of evolutionary constraint acting on the paternally expressed transcripts in the region. Furthermore, phenotypes observed in Angelman syndrome show only minimal correlation with genetic cause. In fact genetic changes leading to functional duplication

of paternally expressed genes (imprinting defects or paternal uniparental disomy) have a less severe phenotype [40, 121, 122]. Therefore, these paternally expressed transcripts are not overtly dosage sensitive, but their expression is essential for life.

Previous studies indicate equal allelic expression of the 5' end of *Ube3a* in the brain [20, 31]. We detected a novel exon at the end of the region predicted to be biallelically expressed. This exon is upstream of a polyadenylation signal and detected by cDNA PCR in only *Ube3a*$^{-/-}$ mice. This may have occurred secondary to primer dilution in other animals because of the relative prevalence of major isoforms. These findings do not explain the observation of equal toRNA and mRNA levels in the brains of *Ube3a* in *Ube3a*$^{+/-}$ and *Ube3a*$^{-/+}$ mice; however, another report suggested the presence of Mef2 bound promoters within *Ube3a* that were distinct from the canonical *Ube3a* promoter [150]. One of these promoters is found downstream of the AS mouse deletion and coincides with predicted novel human protein isoforms arising from 3' exons of *Ube3a*. Furthermore, the discovery of other unique transcripts demonstrated splicing around the AS mouse deletion, which may also be contributing to the RNA findings.

The post-transcriptional processing and the extent of transcription across *Ube3a* of the *Ube3a-AS* shown here is novel. It further highlights the complexity of the *Ube3a-AS* transcript and demonstrates that it does not simply function to down-regulate expression of *Ube3a*. This data suggests that previous quantifications and investigation of the *Ube3a-AS* may have been misleading as they would be synonymous to measuring a

coding transcripts' steady state levels by measuring the presence of its introns. We do however, note significant coverage of the introns of *Ube3a* almost exclusively from the paternal allele, which we presume to be the antisense transcript, suggesting some increased stability or alternate functionality of the unspliced portions of the antisense overlapping *Ube3a* compared to the unspliced portions of the maternal sense *Ube3a* transcript. The idea of increased stability is supported by the existence of the Snord115 host gene cloud that localizes to the locus [116, 117]. Furthermore, alternative functionality cannot be excluded in the face of the ever expanding discovery of novel RNA forms. Although searches for sequence and structure homology failed to reveal functional RNA motifs in the *Ube3a-AS* introns, the possibility still exists for yet unidentified or unannotated RNA structures imbedded in these transcripts.

The importance of appropriate *Ube3a* expression in the brain is unequivocal; however, the importance of the various Ube3a isoforms is incompletely understood. We identified two additional *Ube3a* transcript variants. Although their functions are unknown, they both contain open reading frames that would generate smaller Ube3a proteins with a functional HECT domain. The characterization of these novel isoforms requires further analysis. It is possible that these variants may be directed to specific areas of the cell based on their 5` sequence or have differing target specificities based on amino acid changes outside of the HECT domain.

The approach of alternative transcript identification used here is limited. This approach is restricted to identifying alternative splice transcripts, but is unable to detect truncated transcripts or transcripts using previously unidentified upstream or downstream coding or non-coding exons. These pitfalls do not mitigate the findings presented here but do require appreciation of other transcriptional possibilities. Furthermore, the transcriptional activity of the *Ube3a*$^{+/-}$ mice and the relative paucity of coding SNPs identified in hybrid mice prohibited investigation of allele specific transcript variants.

Angelman syndrome, chromosome 15q11-q13 duplication syndrome, and Prader-Will syndrome have overlapping phenotypes of intellectual disability, social communication deficits (e.g., autism spectrum disorder), and motor developmental delay/impairment. Given that these syndromes arise from dysregulation of transcriptionally coupled genes located in the 15q11-q13 imprinted region, understanding their cellular roles in the brain and the manner in which they are functionally linked is not only important to understand the pathogenesis and management of these syndromes, but to other neurodevelopmental conditions with overlapping features as well. The findings presented here shed light on the function of the *UBE3A* imprint, refine our knowledge of Ube3a expression, provide novel data on the complex structure and post-transcriptional processing of the *SNURF-SNRPN-SNORDs-UBE3A-AS* polycistronic transcript, suggest the existence of previously unappreciated *Ube3a* isoforms, and provide a foundation for future studies.

REFERENCES

[1] Russo, V.E.A., R.A. Martiensen, A.D. Riggs, *Epigenetic Mechanisms of Gene Regulation*. 1996. New York, USA: Cold Spring Harbor Laboratory Press.

[2] Hirasawa, R. and R. Feil, *Genomic imprinting and human disease.* Essays Biochem, 2010. **48**(1): p. 187-200.

[3] Bartolomei, M.S. and S.M. Tilghman, *Genomic imprinting in mammals.* Annu Rev Genet, 1997. **31**: p. 493-525.

[4] Suzuki, S., et al., *The evolution of mammalian genomic imprinting was accompanied by the acquisition of novel CpG islands.* Genome Biol Evol, 2011. **3**: p. 1276-83.

[5] Hsieh, T.F., et al., *Regulation of imprinted gene expression in Arabidopsis endosperm.* Proc Natl Acad Sci U S A, 2011. **108**(5): p. 1755-62.

[6] Coolon, J.D., et al., *Genomic imprinting absent in Drosophila melanogaster adult females.* Cell Rep, 2012. **2**(1): p. 69-75.

[7] Spencer, H.G. and A.G. Clark, *Non-conflict theories for the evolution of genomic imprinting.* Heredity (Edinb), 2014. **113**(2): p. 113-8.

[8] Holliday, R., *Genomic imprinting and allelic exclusion.* Dev Suppl, 1990: p. 125-9.

[9] Nicholls, R.D., et al., *Evaluation of potential models for imprinted and nonimprinted components of human chromosome 15q11-q13 syndromes by fine-structure homology mapping in the mouse.* Proc Natl Acad Sci U S A, 1993. **90**(5): p. 2050-2054.

[10] Rapkins, R.W., et al., *Recent assembly of an imprinted domain from non-imprinted components.* PLoS Genet, 2006. **2**(10): p. e182.

[11] Bielinska, B., et al., *De novo deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch.* Nat Genet, 2000. **25**(1): p. 74-8.

[12] Runte, M., et al., *The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A.* Hum Mol Genet, 2001. **10**(23): p. 2687-700.

[13] Nakao, M., et al., *Imprinting analysis of three genes in the Prader-Willi/Angelman region: SNRPN, E6-associated protein, and PAR-2 (D15S225E).* Hum Mol Genet, 1994. **3**(2): p. 309-15.

[14] Vu, T.H. and A.R. Hoffman, *Imprinting of the Angelman syndrome gene, UBE3A, is restricted to brain.* Nat Genet, 1997. **17**(1): p. 12-3.

[15] Rougeulle, C., H. Glatt, and M. Lalande, *The Angelman syndrome candidate gene, UBE3A/E6-AP, is imprinted in brain.* Nat Genet, 1997. **17**(1): p. 14-5.

[16] Albrecht, U., et al., *Imprinted expression of the murine Angelman syndrome gene, Ube3a, in hippocampal and Purkinje neurons.* Nat Genet, 1997. **17**(1): p. 75-8.

[17] Yamasaki, K., *Neurons but not glial cells show reciprocal imprinting of sense and antisense transcripts of Ube3a.* Human Molecular Genetics, 2003. **12**(8): p. 837-847.

[18] Meng, L., et al., *Truncation of Ube3a-ATS unsilences paternal Ube3a and ameliorates behavioral defects in the Angelman syndrome mouse model.* PLoS Genet, 2013. **9**(12): p. e1004039.

[19] Kishino, T., *Imprinting in neurons.* Cytogenet Genome Res, 2006. **113**(1-4): p. 209-14.

[20] Numata, K., et al., *Highly parallel SNP genotyping reveals high-resolution landscape of mono-allelic Ube3a expression associated with locus-wide antisense transcription.* Nucleic Acids Res, 2011. **39**(7): p. 2649-57.

[21] Ohta, T., et al., *Imprinting-mutation mechanisms in Prader-Willi syndrome.* Am J Hum Genet, 1999. **64**(2): p. 397-413.

[22] Buiting, K., et al., *A 5-kb imprinting center deletion in a family with Angelman syndrome reduces the shortest region of deletion overlap to 880 bp.* Hum Genet, 1999. **105**(6): p. 665-6.

[23] Kantor, B., et al., *Establishing the epigenetic status of the Prader-Willi/Angelman imprinting center in the gametes and embryo.* Hum Mol Genet, 2004. **13**(22): p. 2767-79.

[24] Kaufman, Y., et al., *Protein-binding elements establish in the oocyte the primary imprint of the Prader-Willi/Angelman syndromes domain.* Proc Natl Acad Sci U S A, 2009. **106**(25): p. 10242-7.

[25] Landers, M., et al., *Regulation of the large (approximately 1000 kb) imprinted murine Ube3a antisense transcript by alternative exons upstream of Snurf/Snrpn.* Nucleic Acids Res, 2004. **32**(11): p. 3480-92.

[26] Cavaille, J., et al., *Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization.* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14311-6.

[27] Yin, Q.F., et al., *Long noncoding RNAs with snoRNA ends.* Mol Cell, 2012. **48**(2): p. 219-30.

[28] Rougeulle, C., et al., *An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript.* Nat Genet, 1998. **19**(1): p. 15-6.

[29] Chamberlain, S.J. and C.I. Brannan, *The Prader-Willi syndrome imprinting center activates the paternally expressed murine Ube3a antisense transcript but represses paternal Ube3a.* Genomics, 2001. **73**(3): p. 316-22.

[30] Kohama, C., et al., *ES cell differentiation system recapitulates the establishment of imprinted gene expression in a cell-type-specific manner.* Hum Mol Genet, 2012. **21**(6): p. 1391-401.

[31] Meng, L., R.E. Person, and A.L. Beaudet, *Ube3a-ATS is an atypical RNA polymerase II transcript that represses the paternal expression of Ube3a.* Hum Mol Genet, 2012. **21**(13): p. 3001-12.

[32] Huang, H.S., et al., *Topoisomerase inhibitors unsilence the dormant allele of Ube3a in neurons.* Nature, 2012. **481**(7380): p. 185-9.

[33] Knoll, J.H., et al., *Angelman and Prader-Willi syndromes share a common chromosome 15 deletion but differ in parental origin of the deletion.* American Journal of Medical Genetics, 1989. **32**(2): p. 285-290.

[34] Ledbetter, D.H., et al., *Chromosome 15 abnormalities and the Prader-Willi syndrome: a follow-up report of 40 cases.* Am J Hum Genet, 1982. **34**(2): p. 278-85.

[35] Ledbetter, D.H., et al., *Deletions of chromosome 15 as a cause of the Prader-Willi syndrome.* N Engl J Med, 1981. **304**(6): p. 325-9.

[36] Battaglia, A., et al., *The inv dup(15) syndrome: a clinically recognizable syndrome with altered behavior, mental retardation, and epilepsy.* Neurology, 1997. **48**(4): p. 1081-6.

[37] Kishino, T., M. Lalande, and J. Wagstaff, *UBE3A/E6-AP mutations cause Angelman syndrome.* Nat Genet, 1997. **15**(1): p. 70-3.

[38] Matsuura, T., et al., *De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome.* Nat Genet, 1997. **15**(1): p. 74-7.

[39] Sutcliffe, J.S., et al., *The E6-Ap ubiquitin-protein ligase (UBE3A) gene is localized within a narrowed Angelman syndrome critical region.* Genome Res, 1997. **7**(4): p. 368-77.

[40] Williams, C.A., et al., *Angelman syndrome 2005: updated consensus for diagnostic criteria.* Am J Med Genet A, 2006. **140**(5): p. 413-8.

[41] Cook, E.H., Jr., et al., *Autism or atypical autism in maternally but not paternally derived proximal 15q duplication.* Am J Hum Genet, 1997. **60**(4): p. 928-34.

[42] de Smith, A.J., et al., *A deletion of the HBII-85 class of small nucleolar RNAs (snoRNAs) is associated with hyperphagia, obesity and hypogonadism.* Hum Mol Genet, 2009. **18**(17): p. 3257-65.

[43] Sahoo, T., et al., *Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster.* Nat Genet, 2008. **40**(6): p. 719-21.

[44] Scheffner, M., et al., *The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53.* Cell, 1993. **75**(3): p. 495-505.

[45] Scheffner, M. and O. Staub, *HECT E3s and human disease.* BMC Biochem, 2007. **8 Suppl 1**: p. S6.

[46] Huibregtse, J.M., M. Scheffner, and P.M. Howley, *Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins.* Mol Cell Biol, 1993. **13**(8): p. 4918-27.

[47] Yamamoto, Y., J.M. Huibregtse, and P.M. Howley, *The human E6-AP gene (UBE3A) encodes three potential protein isoforms generated by differential splicing.* Genomics, 1997. **41**(2): p. 263-6.

[48] Kishino, T. and J. Wagstaff, *Genomic organization of the UBE3A/E6-AP gene and related pseudogenes.* Genomics, 1998. **47**(1): p. 101-7.

[49] Martinez-Noel, G., et al., *Identification and proteomic analysis of distinct UBE3A/E6AP protein complexes.* Mol Cell Biol, 2012. **32**(15): p. 3095-106.

[50] Kumar, S., A.L. Talis, and P.M. Howley, *Identification of HHR23A as a substrate for E6-associated protein-mediated ubiquitination.* J Biol Chem, 1999. **274**(26): p. 18785-92.

[51] Rotin, D. and S. Kumar, *Physiological functions of the HECT family of ubiquitin ligases.* Nat Rev Mol Cell Biol, 2009. **10**(6): p. 398-409.

[52] Singhmar, P. and A. Kumar, *Angelman syndrome protein UBE3A interacts with primary microcephaly protein ASPM, localizes to centrosomes and regulates chromosome segregation.* PLoS One, 2011. **6**(5): p. e20397.

[53] Jana, N.R., *Understanding the pathogenesis of Angelman syndrome through animal models.* Neural Plast, 2012. **2012**(710943): p. 1-10.

[54] Yang, Y., et al., *Ubiquitin-dependent proteolysis of trihydrophobin 1 (TH1) by the human papilloma virus E6-associated protein (E6-AP).* J Cell Biochem, 2007. **101**(1): p. 167-80.

[55] Khan, O.Y., et al., *Multifunction steroid receptor coactivator, E6-associated protein, is involved in development of the prostate gland.* Mol Endocrinol, 2006. **20**(3): p. 544-59.

[56] Ramamoorthy, S. and Z. Nawaz, *E6-associated protein (E6-AP) is a dual function coactivator of steroid hormone receptors.* Nucl Recept Signal, 2008. **6**(e006): p. 1-9.

[57] Crouse, H.V., *The Controlling Element in Sex Chromosome Behavior in Sciara.* Genetics, 1960. **45**(10): p. 1429-43.

[58] Surani, M.A.H., S.C. Barton, and M.L. Norris, *Development of Reconstituted mouse eggs suggests imprinting of the genome during gametogenesis.* Nature, 1984. **308**(5959): p. 548-550.

[59] McGrath, J. and D. Solter, *Completion of mouse embryogenesis requires both the maternal and paternal genomes.* Cell, 1984. **37**(1): p. 179-183.

[60] DeChiara, T.M., E.J. Robertson, and A. Efstratiadis, *Parental imprinting of the mouse insulin-like growth factor II gene.* Cell, 1991. **64**(4): p. 849-59.

[61] Weaver, J.R. and M.S. Bartolomei, *Chromatin regulators of genomic imprinting.* Biochim Biophys Acta, 2014. **1839**(3): p. 169-77.

[62] Morison, I.M., J.P. Ramsay, and H.G. Spencer, *A census of mammalian imprinting.* Trends Genet, 2005. **21**(8): p. 457-65.

[63] Prickett, A.R. and R.J. Oakey, *A survey of tissue-specific genomic imprinting in mammals.* Mol Genet Genomics, 2012. **287**(8): p. 621-30.

[64] Reik, W. and E.R. Maher, *Imprinting in Clusters: Lessons from Beckwith-Weidaman syndrome.* Trends in Genetics, 1997. **13**(8): p. 330-334.

[65] Barlow, D.P. and M.S. Bartolomei, *Genomic imprinting in mammals.* Cold Spring Harb Perspect Biol, 2014. **6**(2): p. a018382.

[66] Koerner, M.V., et al., *The function of non-coding RNAs in genomic imprinting.* Development, 2009. **136**(11): p. 1771-83.

[67] Watt, F. and P.L. Molloy, *Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter.* Genes Dev, 1988. **2**(9): p. 1136-43.

[68] Dindot, S.V., et al., *Epigenetic profiling at mouse imprinted gene clusters reveals novel epigenetic and genetic features at differentially methylated regions.* Genome Res, 2009. **19**(8): p. 1374-83.

[69] Bartolomei, M.S. and A.C. Ferguson-Smith, *Mammalian genomic imprinting.* Cold Spring Harb Perspect Biol, 2011. **3**(7): p. a002592.

[70] Timothy Bestor, A.L., Robert Mattaliano, Vernon Ingram, *Cloning and Sequencing of a cDNA Encoding DNA methylatrasferase of Mouse Cells.* Journal of Molecular Biology, 1998. **203**: p. 971-983.

[71] Okano M, X.S., Li E, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.* Nature Genetics, 1998. **19**(3): p. 219-220.

[72] Justyna Turek-Plewa, P.P.J., *The Role of Mammalian DNA Methyltransferases in the Regulation of Gene Expression.* Cellular and Molecular Biology Letters, 2005. **10**: p. 631-647.

[73] Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.* Cell, 1999. **99**(3): p. 247-57.

[74] Santos, F., et al., *Dynamic reprogramming of DNA methylation in the early mouse embryo.* Dev Biol, 2002. **241**(1): p. 172-82.

[75] Hata, K., et al., *Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice.* Development, 2002. **129**(8): p. 1983-93.

[76] Kaneda, M., et al., *Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting.* Nature, 2004. **429**(6994): p. 900-3.

[77] Leonhardt, H., et al., *A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei.* Cell, 1992. **71**(5): p. 865-73.

[78] Lewis, J.D., et al., *Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA.* Cell, 1992. **69**(2): p. 905-914.

[79] Meehan, R.R., et al., *Identification of a Mammalian Protein that Binds Specifically to DNA Containing Methylated CpGs.* Cell, 1989. **58**(3): p. 499-507.

[80] Bell, A.C. and G. Felsenfeld, *Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene.* Nature, 2000. **405**(6785): p. 482-5.

[81] Thomas, J.O. and R.D. Kornberg, *An octamer of histones in chromatin and free in solution.* Proc Natl Acad Sci U S A, 1975. **72**(7): p. 2626-30.

[82] Philip, M., M. Jamaluddin, and H.S. Chandra, *Nucleosome core protein: asymmetric dissociation of the octamer.* Biochim Biophys Acta, 1980. **607**(3): p. 480-9.

[83] Hebbes, T.R., A.W. Thorne, and C. Crane-Robinson, *A direct link between core histone acetylation and transcriptionally active chromatin.* The EMBO Journal, 1988. **7**(5): p. 1395-1402.

[84] Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.

[85] McEwen, K.R. and A.C. Ferguson-Smith, *Distinguishing epigenetic marks of developmental and imprinting regulation.* Epigenetics Chromatin, 2010. **3**(1): p. 2.

[86] Qu, H. and X. Fang, *A brief review on the Human Encyclopedia of DNA Elements (ENCODE) project.* Genomics Proteomics Bioinformatics, 2013. **11**(3): p. 135-41.

[87] Schmidt, J.V., J.M. Levorse, and S.M. Tilghman, *Enhancer competition between H19 and Igf2 does not mediate their imprinting.* Proc Natl Acad Sci U S A, 1999. **96**(17): p. 9733-8.

[88] Thorvaldsen, J.L., K.L. Duran, and M.S. Bartolomei, *Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2.* Genes & Development, 1998. **12**(23): p. 3693-3702.

[89] Hark, A.T., et al., *CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus.* Nature, 2000. **405**(6785): p. 486-9.

[90] Ferguson-Smith, A.C., et al., *Parental-origin-specific epigenetic modification of the mouse H19 gene.* Nature, 1993. **362**(6422): p. 751-5.

[91] Zwart, R., et al., *Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes.* Genes Dev, 2001. **15**(18): p. 2361-6.

[92] Sleutels, F., R. Zwart, and D.P. Barlow, *The non-coding Air RNA is required for silencing autosomal imprinted genes.* Nature, 2002. **415**(6873): p. 810-3.

[93] Latos, P.A., et al., *Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing.* Science, 2012. **338**(6113): p. 1469-72.

[94] Santoro, F., et al., *Imprinted Igf2r silencing depends on continuous Airn lncRNA expression and is not restricted to a developmental window.* Development, 2013. **140**(6): p. 1184-95.

[95] Nagano, T., et al., *The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin.* Science, 2008. **322**(5908): p. 1717-20.

[96] Zhang, H., et al., *Long noncoding RNA-mediated intrachromosomal interactions promote imprinting at the Kcnq1 locus.* J Cell Biol, 2014. **204**(1): p. 61-75.

[97] Golding, M.C., et al., *Depletion of Kcnq1ot1 non-coding RNA does not affect imprinting maintenance in stem cells.* Development, 2011. **138**(17): p. 3667-78.

[98] Mancini-Dinardo, D., et al., *Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes.* Genes Dev, 2006. **20**(10): p. 1268-82.

[99] Williamson, C.M., et al., *Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster.* Nat Genet, 2006. **38**(3): p. 350-5.

[100] Williamson, C.M., et al., *Uncoupling antisense-mediated silencing and DNA methylation in the imprinted Gnas cluster.* PLoS Genet, 2011. **7**(3): p. e1001347.

[101] Coombes, C., et al., *Epigenetic properties and identification of an imprint mark in the Nesp-Gnasxl domain of the mouse Gnas imprinted locus.* Mol Cell Biol, 2003. **23**(16): p. 5475-88.

[102] Liu, J., et al., *Identification of the control region for tissue-specific imprinting of the stimulatory G protein alpha-subunit.* Proc Natl Acad Sci U S A, 2005. **102**(15): p. 5513-8.

[103] Williamson, C.M., et al., *A cis-acting control region is required exclusively for the tissue-specific imprinting of Gnas.* Nat Genet, 2004. **36**(8): p. 894-9.

[104] Ball, S.T., et al., *Gene Dosage Effects at the Imprinted Cluster.* PLoS One, 2013. **8**(6): p. e65639.

[105] R, D., et al., *Convergent and divergent evolution of genomic imprinting in the marsupial Monodelphis domestica.* BMC Genomics, 2012. **13**(394).

[106] MacIsaac, J.L., et al., *Tissue-specific alternative polyadenylation at the imprinted gene Mest regulates allelic usage at Copg2.* Nucleic Acids Res, 2012. **40**(4): p. 1523-35.

[107] Zhang, Z., et al., *Comparative analyses of genomic imprinting and CpG island-methylation in mouse Murr1 and human MURR1 loci revealed a putative imprinting control region in mice.* Gene, 2006. **366**(1): p. 77-86.

[108] Wood, A.J. and R.J. Oakey, *Genomic Imprinting in Mammals: Emerging Themes and Established Theories.* PLOS Genetics, 2006. **2**(11): p. e147.

[109] Jiang, Y.H., et al., *Altered ultrasonic vocalization and impaired learning and memory in Angelman syndrome mouse model with a large maternal deletion from Ube3a to Gabrb3.* PLoS One, 2010. **5**(8): p. e12278.

[110] Barlow, D.P., *Methylation and imprinting: from host defense to gene regulation?* Science, 1993. **260**(5106): p. 309-10.

[111] Haig, D., *Genomic imprinting and kinship: how good is the evidence?* Annu Rev Genet, 2004. **38**: p. 553-85.

[112] Glessner, J.T., et al., *Autism genome-wide copy number variation reveals ubiquitin and neuronal genes.* Nature, 2009. **459**(7246): p. 569-73.

[113] TA, G., et al., *Concerted regulation and molecular evolution of the duplicated SNRPB'/B and SNRPN loci.* Nucleic Acids Research, 199. **27**(23): p. 4577-4584.

[114] Chai, J.H., et al., *Retrotransposed genes such as Frat3 in the mouse Chromosome 7C Prader-Willi syndrome region acquire the imprinted status of their insertion site.* Mamm Genome, 2001. **12**(11): p. 813-21.

[115] Bittel, D.C., et al., *Whole genome microarray analysis of gene expression in an imprinting center deletion mouse model of Prader-Willi syndrome.* Am J Med Genet A, 2007. **143**(5): p. 422-9.

[116] Vitali, P., et al., *Long nuclear-retained non-coding RNAs and allele-specific higher-order chromatin organization at imprinted snoRNA gene arrays.* J Cell Sci, 2010. **123**(Pt 1): p. 70-83.

[117] Powell, W.T., et al., *A Prader-Willi locus lncRNA cloud modulates diurnal genes and energy expenditure.* Hum Mol Genet, 2013. **22**(21): p. 4318-28.

[118] Cassidy, S.B., et al., *Prader-Willi syndrome.* Genet Med, 2012. **14**(1): p. 10-26.

[119] Steffenburg, S., et al., *Autism in Angelman syndrome: a population-based study.* Pediatr Neurol, 1996. **14**(2): p. 131-6.

[120] Angelman, H., *'Puppet children'-a report of three cases.* Developmental Medicine and Child Neurology, 1965. **7**(6): p. 681-688.

[121] Williams, C.A., *The behavioral phenotype of the Angelman syndrome.* Am J Med Genet C Semin Med Genet, 2010. **154C**(4): p. 432-7.

[122] Williams, C.A., D.J. Driscoll, and A.I. Dagli, *Clinical and genetic aspects of Angelman syndrome.* Genet Med, 2010. **12**(7): p. 385-95.

[123] Williams, C.A., A. Lossie, and D. Driscoll, *Angelman syndrome: mimicking conditions and phenotypes.* Am J Med Genet, 2001. **101**(1): p. 59-64.

[124] Daily, J.L., et al., *Adeno-associated virus-mediated rescue of the cognitive defects in a mouse model for Angelman syndrome.* PLoS One, 2011. **6**(12): p. e27221.

[125] Metzger, M.B., V.A. Hristova, and A.M. Weissman, *HECT and RING finger families of E3 ubiquitin ligases at a glance.* J Cell Sci, 2012. **125**(Pt 3): p. 531-7.

[126] Mulherkar, S.A., J. Sharma, and N.R. Jana, *The ubiquitin ligase E6-AP promotes degradation of alpha-synuclein.* J Neurochem, 2009. **110**(6): p. 1955-64.

[127] Mishra, A., et al., *E6-AP promotes misfolded polyglutamine proteins for proteasomal degradation and suppresses polyglutamine protein aggregation and toxicity.* J Biol Chem, 2008. **283**(12): p. 7648-56.

[128] Nuber, U., S.E. Schwarz, and M. Scheffner, *The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as its own substrate.* Eur J Biochem, 1998. **254**(3): p. 643-9.

[129] Yashiro, K., et al., *Ube3a is required for experience-dependent maturation of the neocortex.* Nat Neurosci, 2009. **12**(6): p. 777-83.

[130] Smith, S.E., et al., *Increased gene dosage of Ube3a results in autism traits and decreased glutamate synaptic transmission in mice.* Sci Transl Med, 2011. **3**(103): p. 103ra97.

[131] Dalva, M.B., *Ephecting excitatory synapse development.* Cell, 2010. **143**(3): p. 341-2.

[132] Margolis, S.S., et al., *EphB-mediated degradation of the RhoA GEF Ephexin5 relieves a developmental brake on excitatory synapse formation.* Cell, 2010. **143**(3): p. 442-55.

[133] Greer, P.L., et al., *The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc.* Cell, 2010. **140**(5): p. 704-16.

[134] Kuhnle, S., et al., *Role of the ubiquitin ligase E6AP-UBE3A in controlling levels of the synaptic protein Arc.* Proc Natl Acad Sci U S A, 2013. **110**(22): p. 8888-8893.

[135] Dindot, S.V., et al., *The Angelman syndrome ubiquitin ligase localizes to the synapse and nucleus, and maternal deficiency results in abnormal dendritic spine morphology.* Hum Mol Genet, 2008. **17**(1): p. 111-8.

[136] Jiang, Y.H., et al., *Mutation of the Angelman ubiquitin ligase in mice causes increased cytoplasmic p53 and deficits of contextual learning and long-term potentiation.* Neuron, 1998. **21**(4): p. 799-811.

[137] Gassmann, M., et al., *Quantifying Western blots: pitfalls of densitometry.* Electrophoresis, 2009. **30**(11): p. 1845-55.

[138] Shetty, A.K., *Progenitor cells from the CA3 region of the embryonic day 19 rat hippocampus generate region-specific neuronal phenotypes in vitro.* Hippocampus, 2004. **14**(5): p. 595-614.

[139] Hilgenberg, L.G. and M.A. Smith, *Preparation of dissociated mouse cortical neuron cultures.* J Vis Exp, 2007. **10**: p. 562-6.

[140] Nahkuri, S., et al., *Molecular evolution of the HBII-52 snoRNA cluster.* J Mol Biol, 2008. **381**(4): p. 810-5.

[141] Ameur, A., et al., *Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain.* Nat Struct Mol Biol, 2011. **18**(12): p. 1435-40.

[142] Solter, D., *Differential imprinting and expression of maternal and paternal genomes.* Annu Rev Genet, 1988. **22**: p. 127-46.

[143] Moore, T. and D. Haig, *Genomic imprinting in mammalian development: a parental tug-of-war.* Trends Genet, 1991. **7**(2): p. 45-9.

[144] Frith, M.C., et al., *The abundance of short proteins in the mammalian proteome.* PLoS Genet, 2006. **2**(4): p. e52.

[145] Yang, X., et al., *Discovery and annotation of small proteins using genomics, proteomics, and computational approaches.* Genome Res, 2011. **21**(4): p. 634-41.

[146] Galindo, M.I., et al., *Peptides encoded by short ORFs control development and define a new eukaryotic gene family.* PLoS Biol, 2007. **5**(5): p. e106.

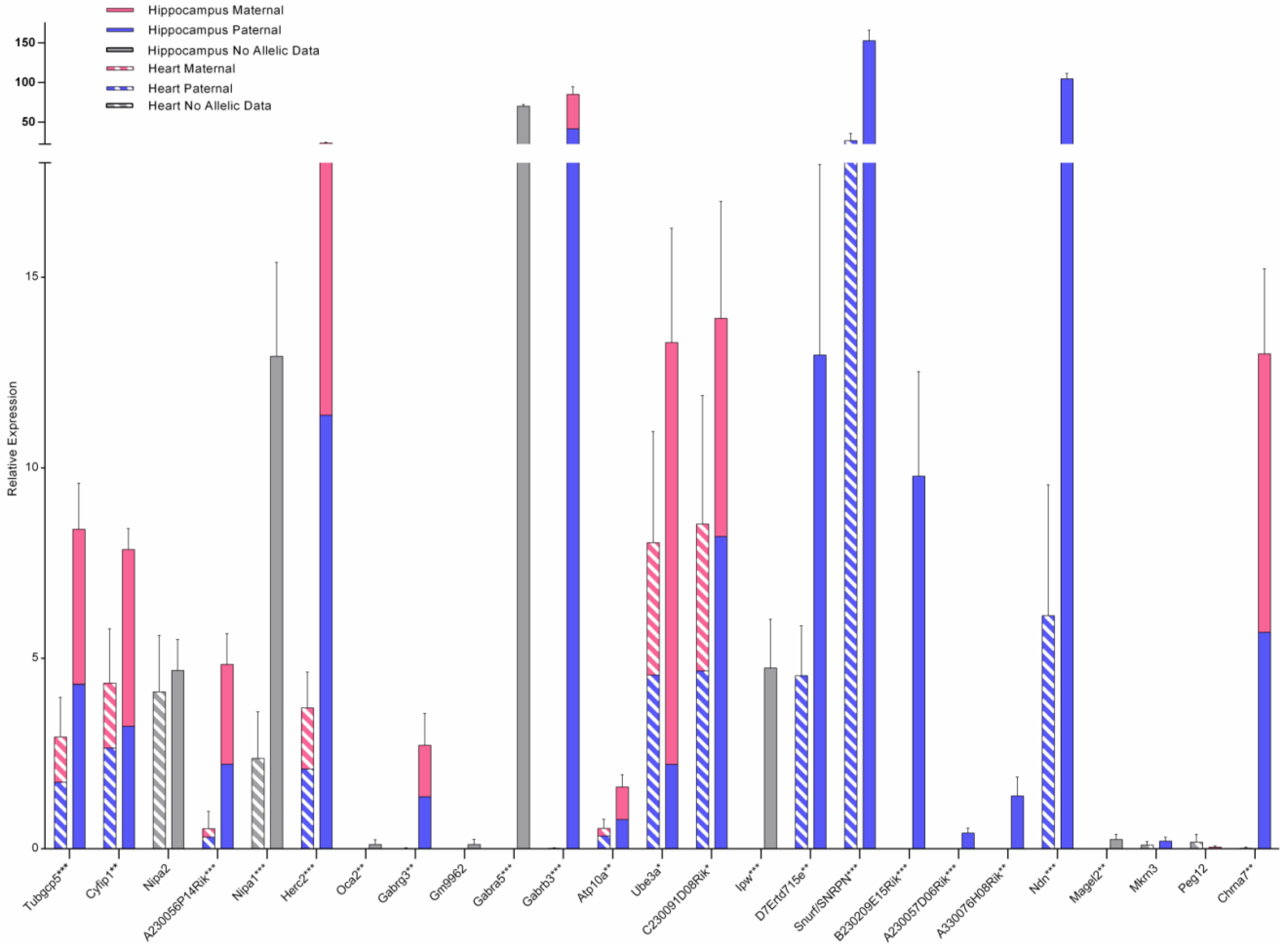
[147] Kaneko-Ishino, T., T. Kohda, and F. Ishino, *The regulation and biological significance of genomic imprinting in mammals.* J Biochem, 2003. **133**(6): p. 699-711.

[148] Wu, M.Y., et al., *An unexpected function of the Prader-Willi syndrome imprinting center in maternal imprinting in mice.* PLoS One, 2012. **7**(4): p. e34348.

[149] Ding, F., et al., *Lack of Pwcr1/MBII-85 snoRNA is critical for neonatal lethality in Prader-Willi syndrome mouse models.* Mamm Genome, 2005. **16**(6): p. 424-31.

[150] Flavell, S.W., et al., *Activity-dependent regulation of MEF2 transcription factors suppresses excitatory synapse number.* Science, 2006. **311**(5763): p. 1008-12.

APPENDIX A



APPENDIX A. The mouse 7C imprinted region is upregulated in the heart. Graph of normalized allelic expression for all RefSeq annotated Genes in the mouse 7C region in brain and heart from RNA-seq analysis. Genes are presented in genetic order by start position. Only unambiguous SNVs in exons and 3' untranslated regions were used for ratios. Gray bars indicate the absence of informative coding SNVs for that gene. n = 6. Significance values are indicated on gene names representing comparisons between tissues, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

APPENDIX B

Please see attached Microsoft Excel file entitled “APPENDIX”, sheet “APPENDIX B”. This table summarizes the analysis of RNA-seq data for all genes in the 7c region. This table contains 37 columns detailing gene coordinates and sizes, per animal and averaged expression values, significance tests, fold change calculations by two alternative methods, informative coding single nucleotide polymorphism (SNP) counts and allelic contributions, and calculated allelic expression values.

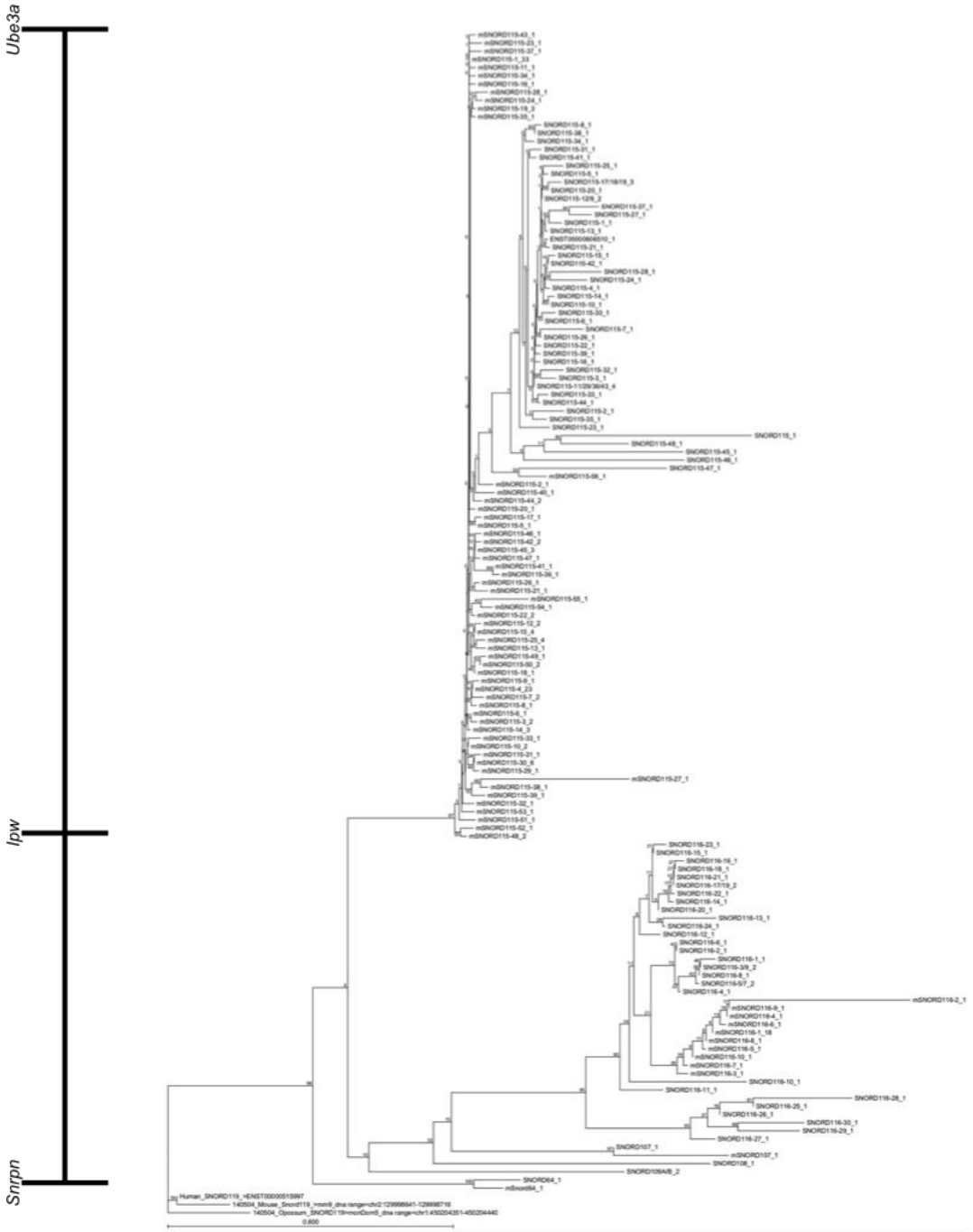
APPENDIX C

Please see attached Microsoft Excel file entitled “APPENDIX”, sheet “APPENDIX C”. This table summarizes all single nucleotide polymorphisms called amongst all 6 mice in the 7c region for both brain and heart and details the total coverage and allelic frequency of each polymorphism.

APPENDIX D

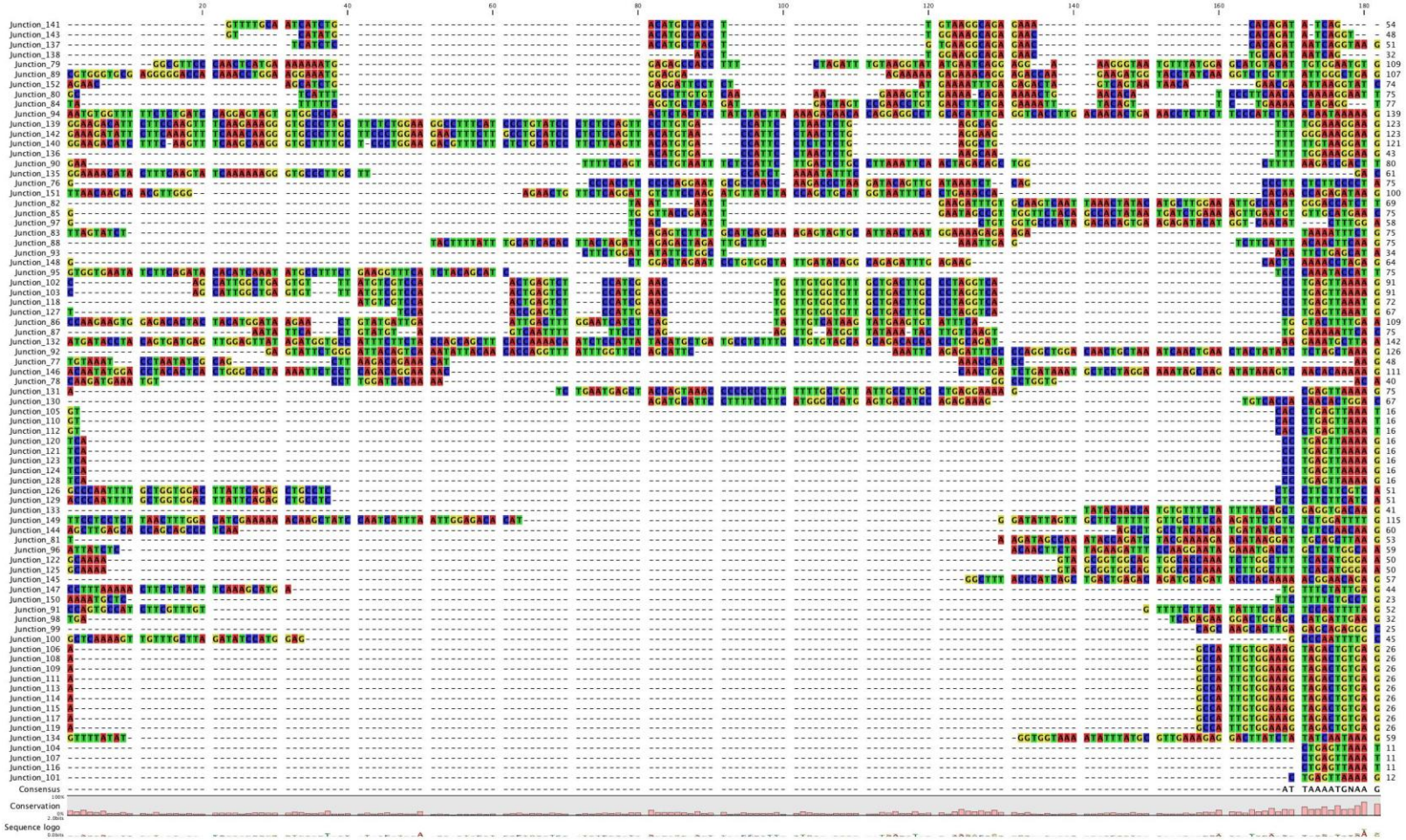
Please see attached Microsoft Excel file entitled “APPENDIX”, sheet “APPENDIX D”. This table summarizes all the snoRNAs found in the *Ube3a-Snrpn* region. Ensembl annotations and genomic coordinates were used to extract sequences. Assigned naming was based on duplication frequency and location on the chromosome in the order of transcription.

APPENDIX E



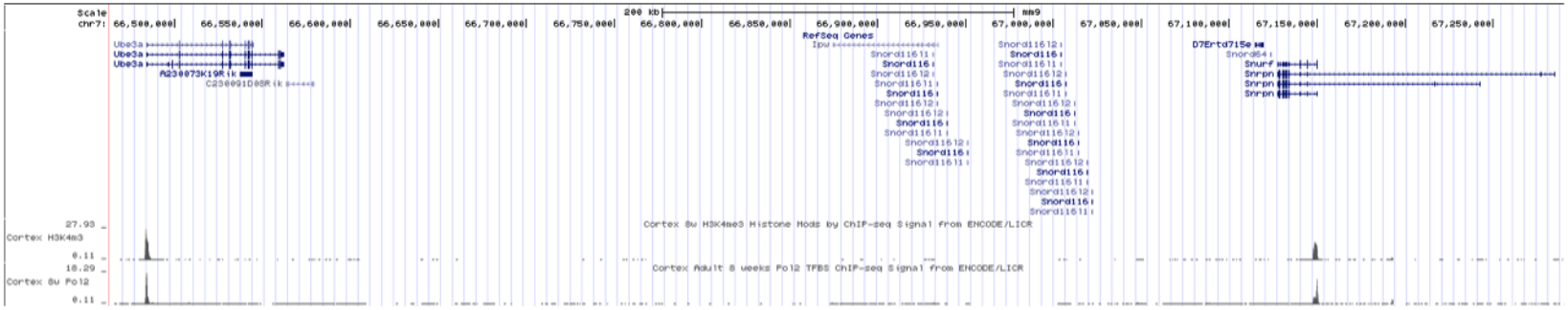
APPENDIX E. Human and mouse snoRNAs derive from the same ancestral snoRNA, *SNORD119*. **T**ree clustering analysis of all human (h) and mouse (m) Snrpn-Ube3a region SNORD sequence, rooted from ancestral human, mouse, and opossum (o) SNORD119. Horizontal distance is representative of relative sequence divergence from branch points

APPENDIX F



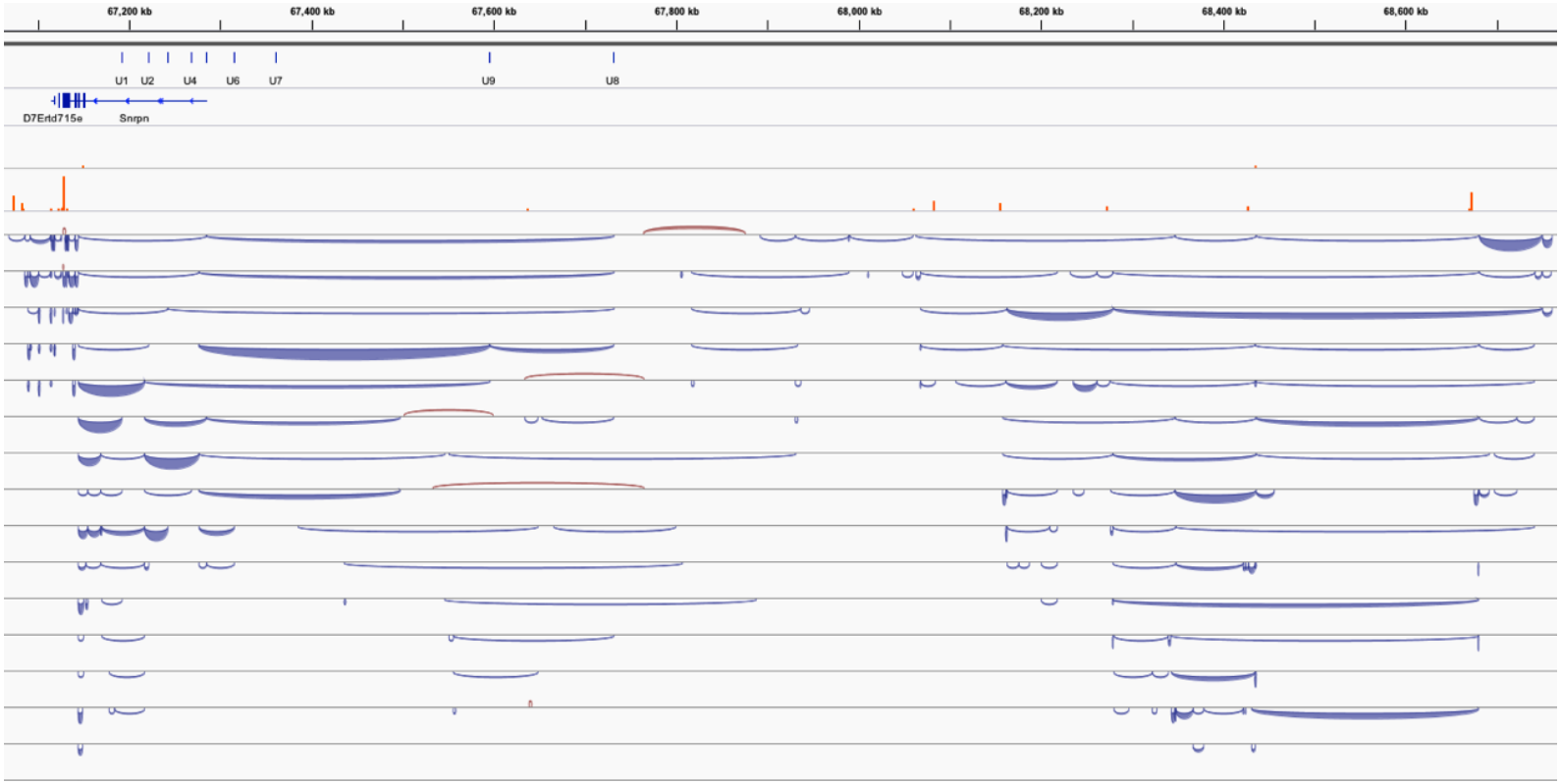
APPENDIX F. Snord host-gene exons do not reflect the same evolutionary conservation as their resident snoRNAs. Alignment of all junctional exons extracted from RNA-seq splicing events within the *Ube3a-Snrpn* intergenic region. No alignment or elements of conservation were detected amongst these exons.

APPENDIX G



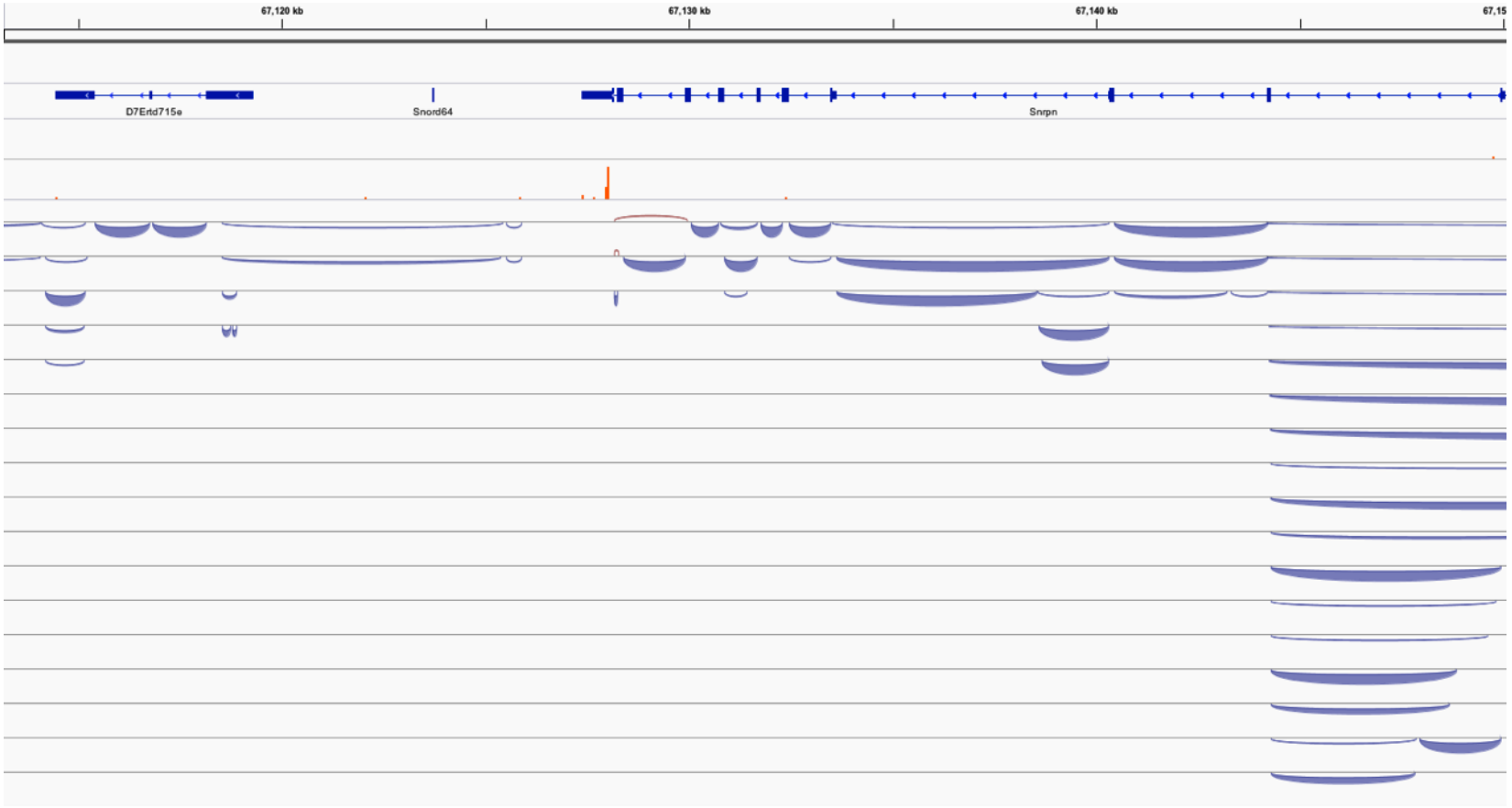
APPENDIX G. Analysis of Ube3a-Snrpn intergenic region Chip-Seq data does not indicate the presence of marks consistent with promoters. Screen capture of UCSC genome browser with RefSeq annotations. Top row of peaks represent trimethylated histone 3 lysine 4 RNA-seq read alignments in adult mouse whole brain. Bottom row of peaks represent RNA polymerase II (Pol2) RNA-seq read alignments in adult mouse whole brain. Note the enrichment of both elements over the promoters of both *Ube3a* and *Snrpn* and the absence of any such enrichment in the intergenic region.

APPENDIX H



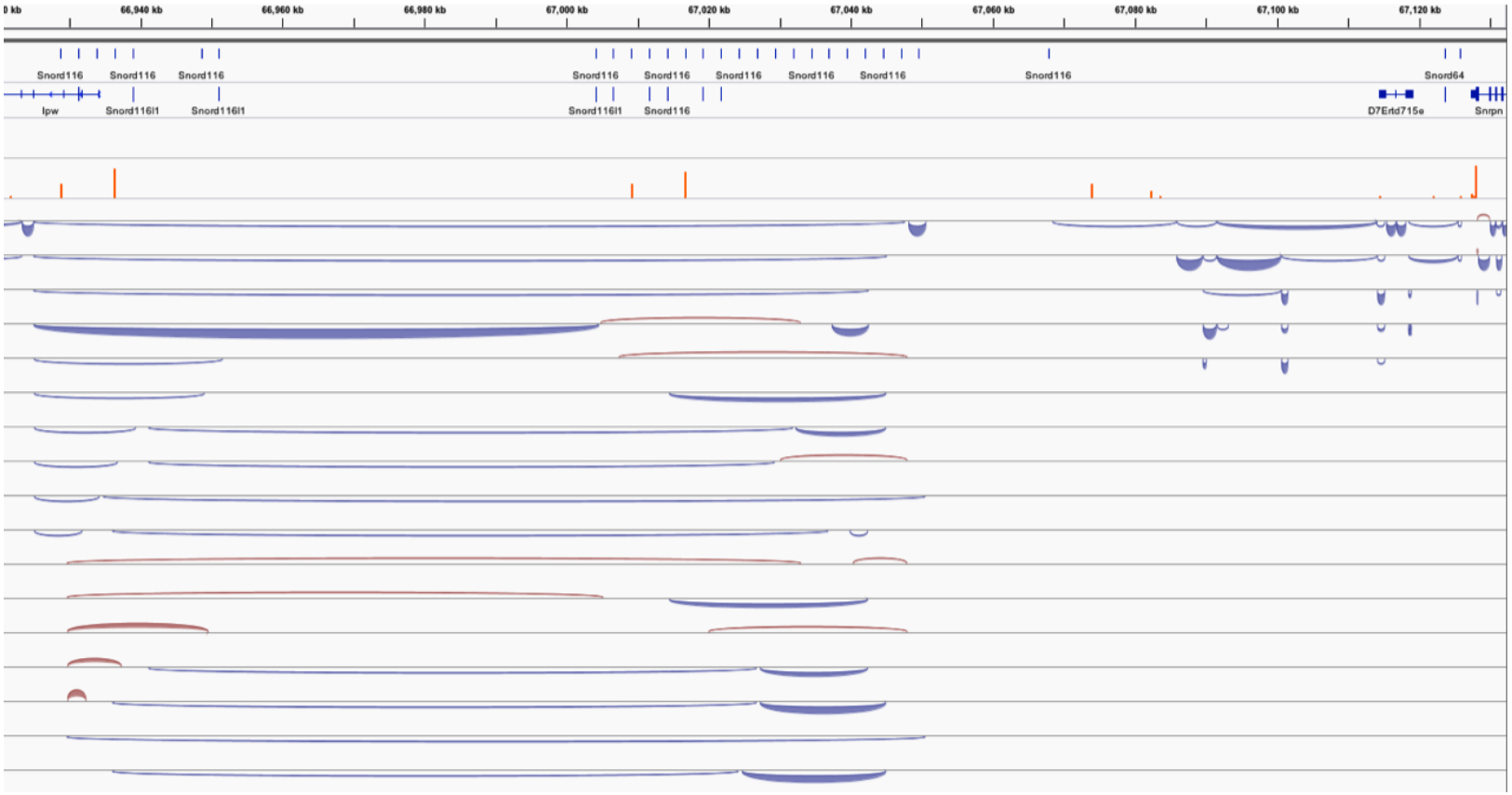
APPENDIX H. **The upstream exons of the *Snurf-Snrpn* exhibit extensive alternative splicing in the brain.** Illustration of alternative splicing events and strand specific polyadenylation sites in the *Snurf-Snrpn* upstream region. Top track line represents previously described alternative upstream exons of *Snurf-Snrpn* ("U" exons) [25]. RefSeq gene annotations are represented and labeled in blue. Orange bars represent relative intensity of strand specific (plus strand, top; minus strand bottom) polyadenylation sites. Red arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the plus strand. Blue arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the minus strand.

APPENDIX I



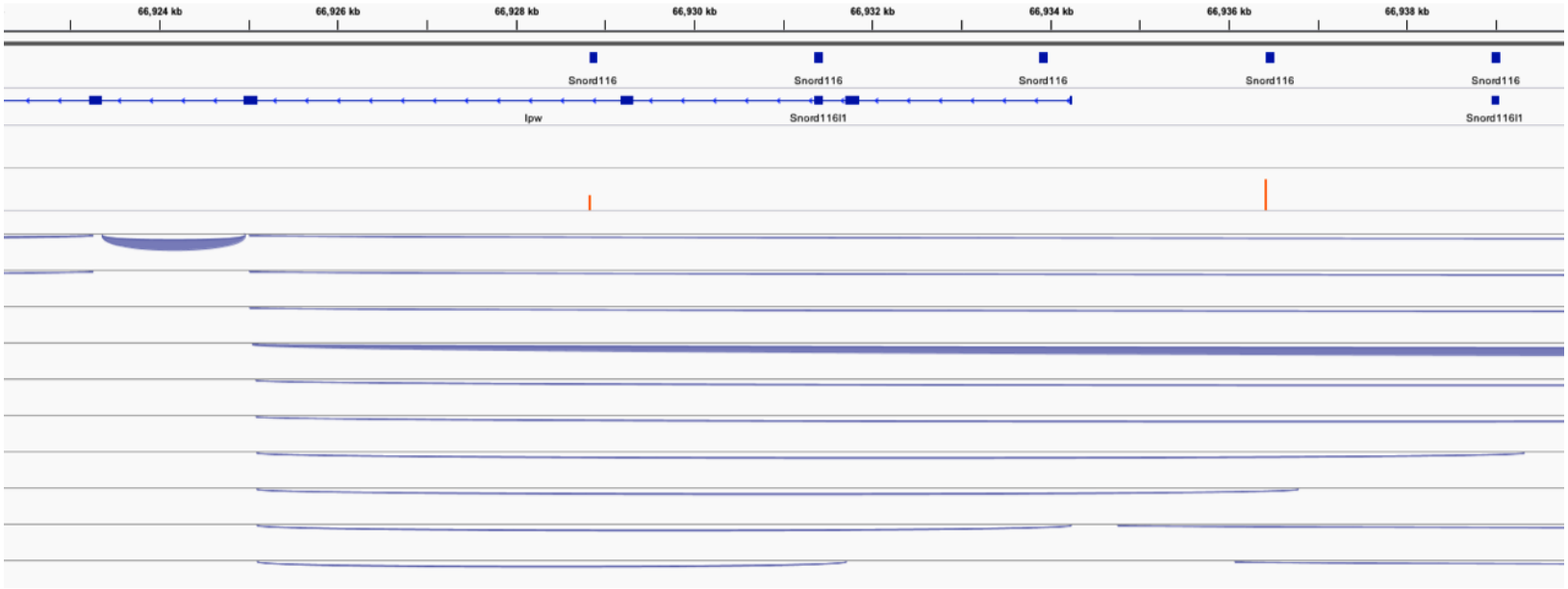
APPENDIX I. Alternative splicing flanking *Snord64* is not contiguous with *Snurf-Snrpn*. Illustration of alternative splicing events flanking *Snord64*. Alternative splicing events surrounding *Snord64* are contiguous with downstream splicing around *Snord116* but not with upstream splicing of *Snurf-Snrpn*. RefSeq gene annotations are represented and labeled in blue. Orange bars represent relative intensity of strand specific (plus strand, top; minus strand bottom) polyadenylation sites. Red arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the plus strand. Blue arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the minus strand.

APPENDIX J



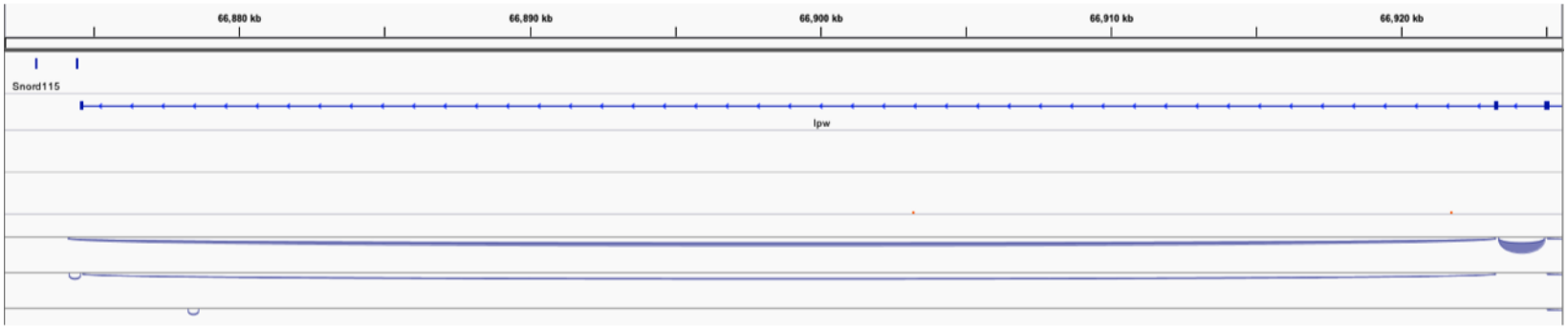
APPENDIX J. ***Snord116*** **host gene exons demonstrate alternative splicing and are linked with** ***Snord64*** **and** ***Ipw*****.** Illustration of alternative splicing events within the *Snord116* cluster. Alternative splicing within *Snord116* link it to *Snord64* upstream and *Ipw* downstream. RefSeq gene annotations are represented and labeled in blue. Orange bars represent relative intensity of strand specific (plus strand, top; minus strand bottom) polyadenylation sites. Red arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the plus strand. Blue arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the minus strand.

APPENDIX K



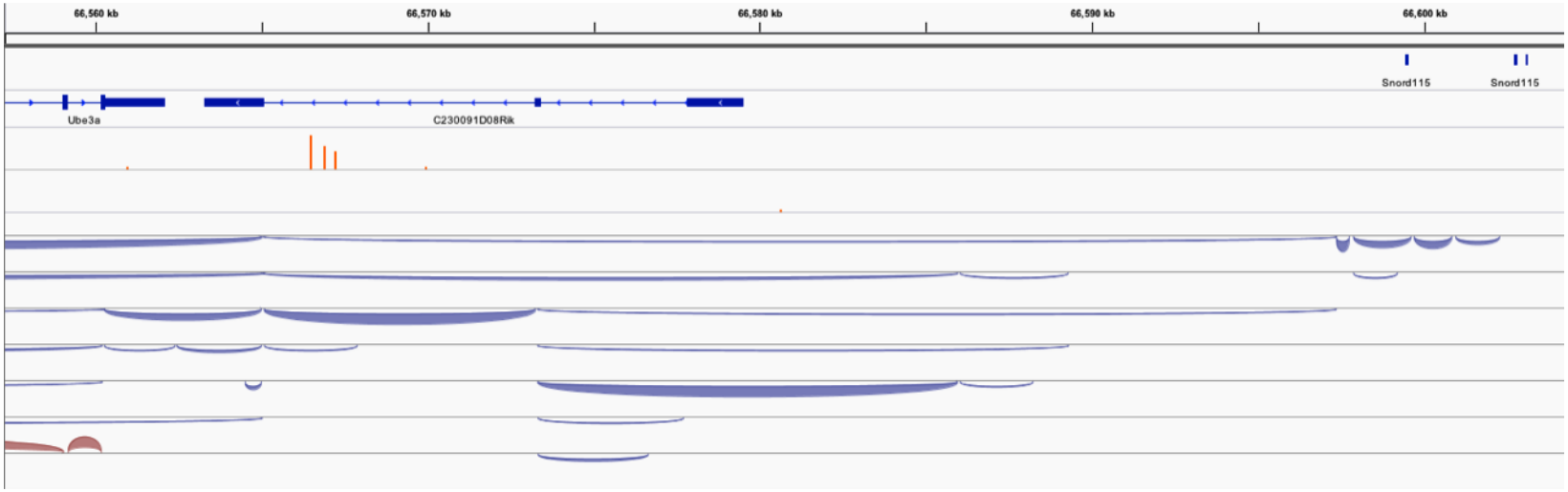
APPENDIX K. The *Snord116* clusters splice into *Ipw* and around strand specific polyadenylation signals. Illustration of alternative splicing events of *Snord116* into *Ipw*. Alternative splicing of Snord116 into Ipw leads to splicing around strand specific polyadenylation signals. RefSeq gene annotations are represented and labeled in blue. Orange bars represent relative intensity of strand specific (plus strand, top; minus strand bottom) polyadenylation sites. Red arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the plus strand. Blue arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the minus strand.

APPENDIX L



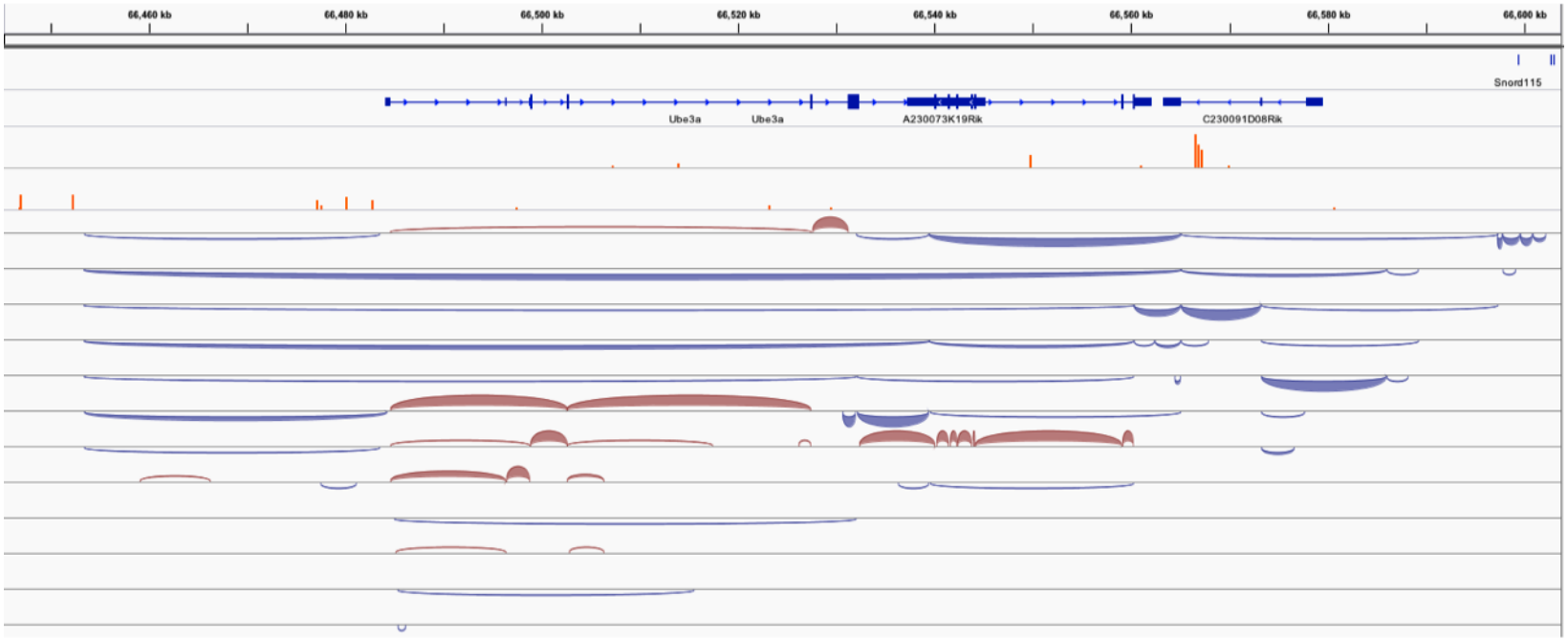
APPENDIX L. The *Snord116* clusters and *Ipw* transcripts are contiguous with the *Snord115* cluster. Illustration of alternative splicing events within *Ipw*. Alternative splicing of *Snord116* and *Ipw* link these transcripts with the 5' end of the *Snord115* cluster. RefSeq gene annotations are represented and labeled in blue. Orange bars represent relative intensity of strand specific (plus strand, top; minus strand bottom) polyadenylation sites. Red arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the plus strand. Blue arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the minus strand.

APPENDIX M



APPENDIX M. The *Snord115* cluster demonstrates alternative splicing into the 3' end of *Ube3a*. Illustration of alternative splicing events between *Snord115, Ube3a-AS, and Ube3a*. Alternative splicing of the 3' end of *Snord115* is contiguous with *Ube3a-AS* and overlaps the 3' end of *Ube3a*. RefSeq gene annotations are represented and labeled in blue. Orange bars represent relative intensity of strand specific (plus strand, top; minus strand bottom) polyadenylation sites. Red arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the plus strand. Blue arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the minus strand.

APPENDIX N

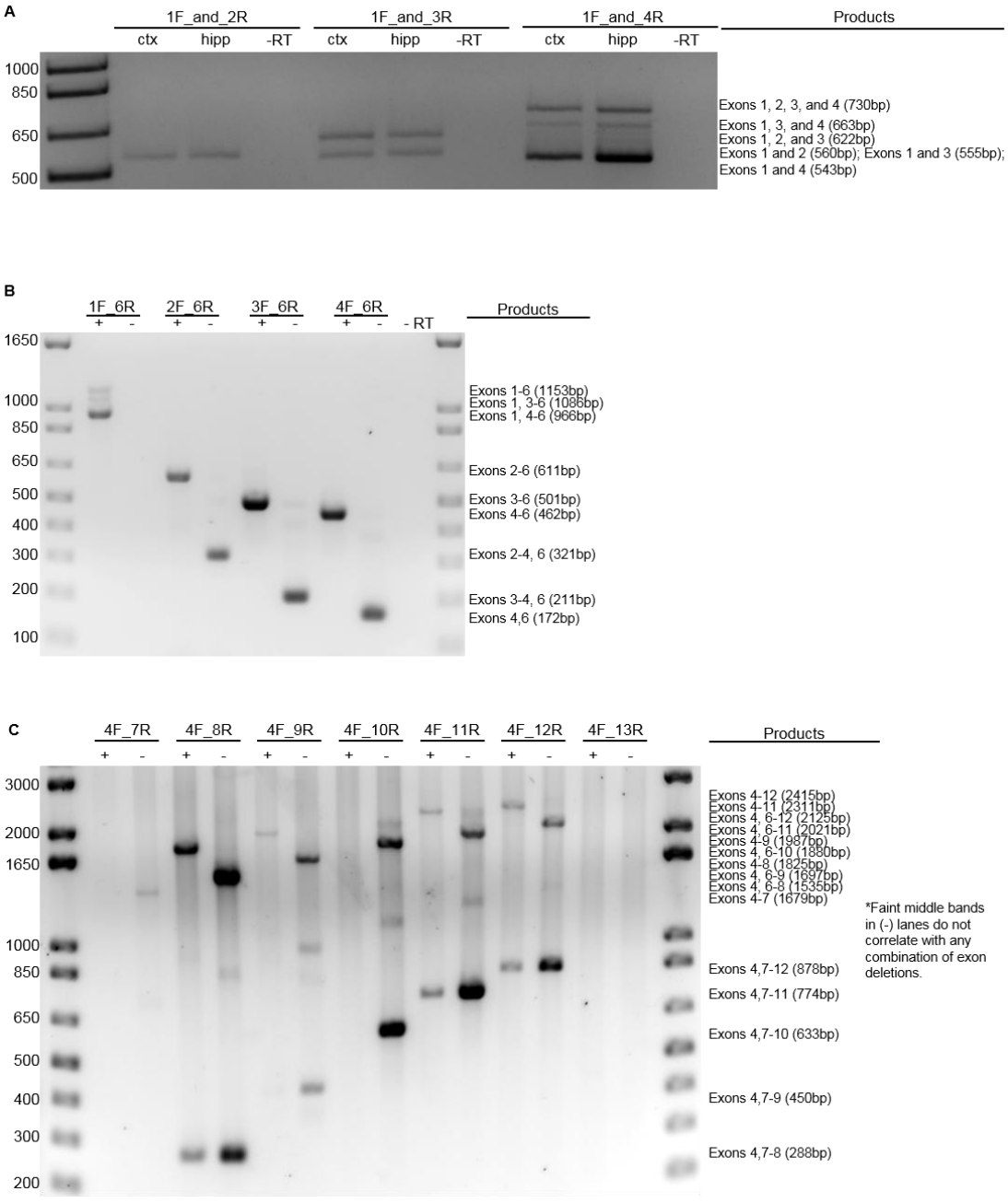


APPENDIX N. Both *Ube3a* and *Ube3a-AS* show extensive alternative splicing in the brain. Illustration of alternative splicing events across *Ube3a*. Alternative splicing of *Ube3a* reveals novel alternative transcripts at its 5' end. These novel transcripts appear to terminate in regions consistent with strand specific polyadenylation signals (orange bars). The *Ube3a-AS* shows even greater alternative splicing into and within *Ube3a*. These splice site are into and out of both introns and exons of *Ube3a*. Furthermore, these transcripts also terminate at strand specific polyadenylation signals within and upstream of *Ube3a*. Orange bars represent relative intensity of strand specific (plus strand, top; minus strand bottom) polyadenylation sites. Red arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the plus strand. Blue arcs represent canonical "GT-AG" splicing events in the 5' → 3' direction of the minus strand.

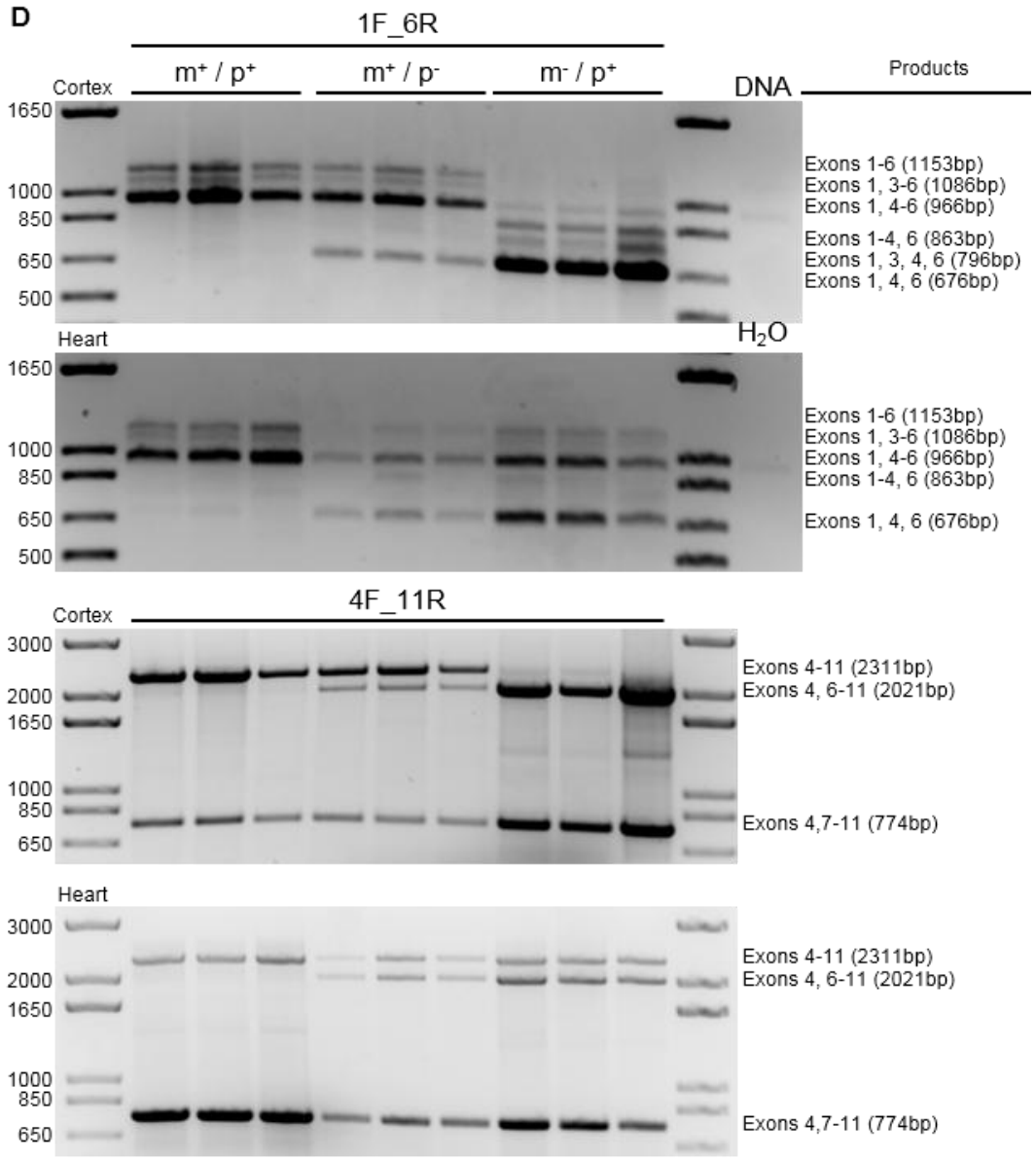
APPENDIX O

Please see attached Microsoft Excel file entitled “APPENDIX”, sheet “APPENDIX N”. This table summarizes the location and relative strength of strand specific polyadenylation signals identified by RNA-seq, provided by the UCSC genome browser.

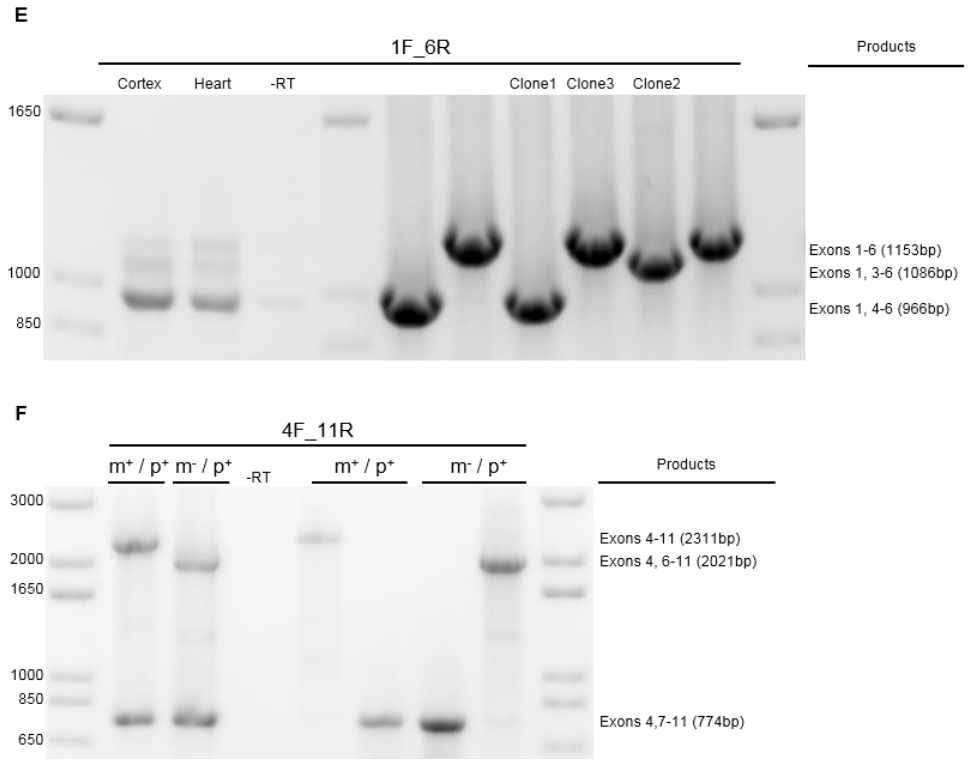
APPENDIX P



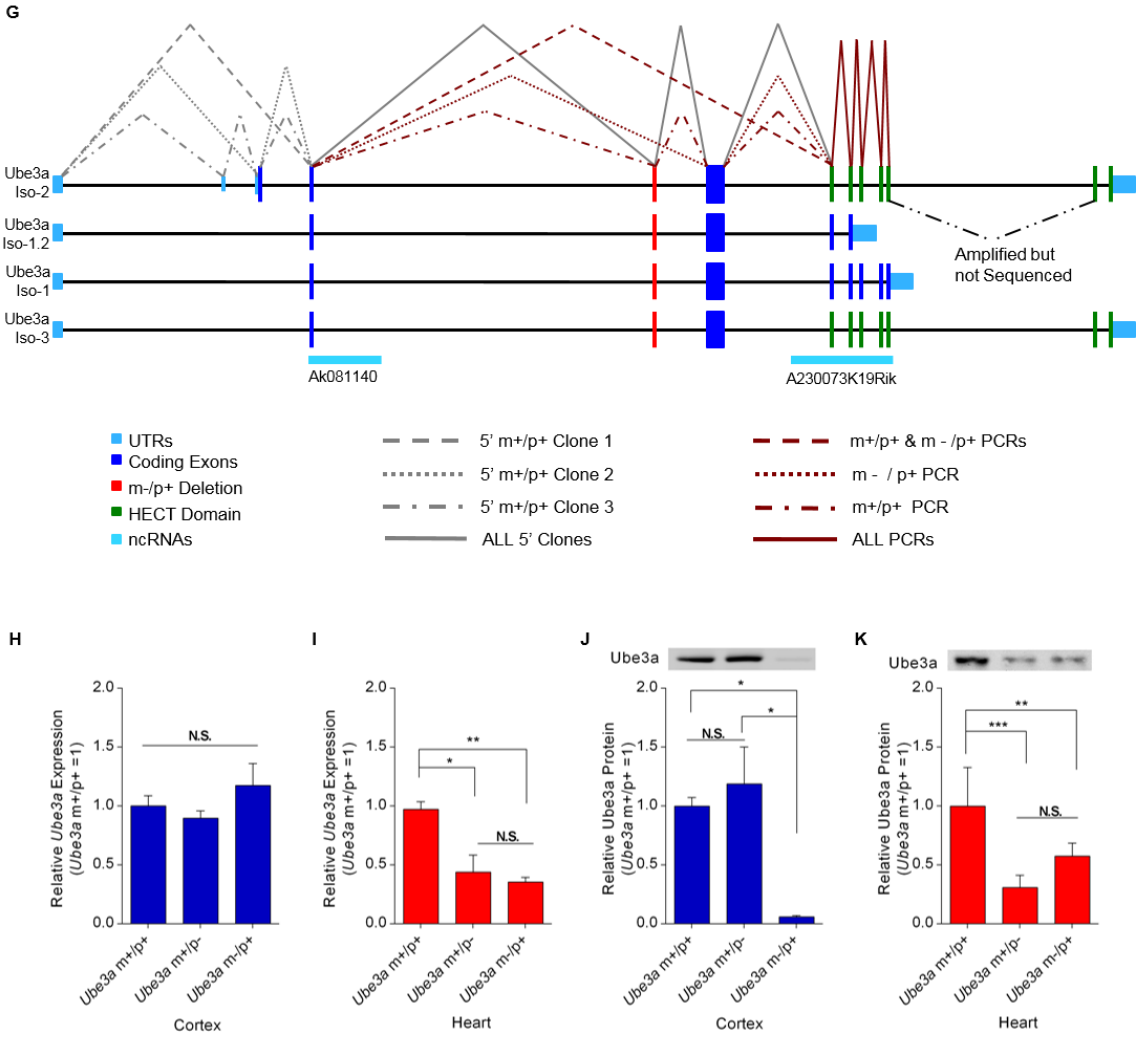
APPENDIX P. Confirmation and discovery of novel *Ube3a* transcripts in Ube3a$^{+/+}$, Ube3a$^{-/+}$, and Ube3a$^{+/-}$ mice. A-D) RT-PCR amplification products of PolyDT cDNA using exon specific primers indicated above images (as described in Appendix Q). Numbers to left of indicate size of ladder standard bands in basepairs. Products column indicates measured product size in base pairs (bp) and predicted exon composition based on size, primers used, and subsequent sanger sequencing. –RT; negative control of RT-PCR reaction run without reverse polymerase. A) Confirmation of alternative 5' splicing in both cortex (ctx) and hippocampus (hipp) of wildtype (m^+ / p^+) mice. B-C) Identification of novel transcripts spanning the Ube3a$^{+/-}$ deletion in the cortex of wildtype (+) and maternal deficient, Ube3a$^{-/+}$, mice (-).



APPENDIX Q. Continued. D) Identification of novel transcripts spanning the *Ube3a* deletion in both the cortex and heart of wildtype Ube3a$^{+/+}$ (m^+/p^+), paternal deficient Ube3a$^{+/-}$ (m^+/p^-), and maternal deficient Ube3a$^{-/+}$ (m^-/p^+) mice.



APPENDIX R. Continued. E) Restriction enzyme digest of clones (bold bands at right of image) used in sanger sequencing (clones 1-3) for RT-PCR products amplified between exons 1 and 6 (1F_6R) (faint bands to left of image). **F)** PCR amplification products from gel band excisions used in sanger sequencing (right side of image) from RT-PCR products amplified between exon 4 and 11 (4F_11R) (left side of image).



APPENDIX S. Continued. G) Schematic representation of known *Ube3a* transcripts with representation of experimentally validated splicing events identified in preceding experiments. Vertical boxes represent individual exons, with coloration specific to contribution to the transcript as indicated in the key. Lines indicate splicing events and connect exons that were shown to splice together, while line color and style refer to source of evidence as indicated in the key. **H-I)** TaqMan analysis of total RNA in the brain (**H**) and heart (**I**) of Ube3a$^{+/+}$, Ube3a$^{+/-}$, and Ube3a$^{-/+}$ mice. **J-K**) Western blots and analysis of brain (**J**) and heart (**K**) of the same mice in **H-I**. N = 6. * P < 0.05, ** P < 0.01, *** P < 0.001. N.S. = not significant.

APPENDIX Q

Please see attached Microsoft Excel file entitled “APPENDIX”, sheet “APPENDIX Q”. This sheet contains genomic coordinates and sequences for all primers used in the reverse transcription assays.

APPENDIX R

Please see attached Microsoft Excel file entitled “APPENDIX”, sheet “APPENDIX R”. This table summarizes all antibodies used in immunofluorescence and western blotting with dilutions used and vendor information.