Spring 5-6-2012

Examination of the Chromatin Structure of Xlr3b Using the Chromosome Conformation Capture Assay

Sarah Elise Conderino
University of Connecticut - Storrs, sarah.conderino@uconn.edu

Examination of the chromatin structure of *Xlr3b* using the chromosome conformation capture assay

Sarah Conderino

April 27, 2012

University of Connecticut

Honors Program

Thesis Advisor: Dr. Mike O'Neill

Honors Advisor: Dr. Chris Simon

Abstract

Imprinted genes contain epigenetic modifications that influence expression patterns based on parent-of-origin. Recent studies have shown that imprinted genes contribute to numerous human diseases and disorders. *Xlr3b*, an imprinted gene on the X chromosome, has been implicated in social and behavioral deficits characteristic of disorders such as Turner syndrome and autism. The imprinting mechanism of this gene is still unknown, and this study analyzed the native chromatin structure of *Xlr3b* through the chromosome conformation capture assay to determine if there are any long-range interactions that regulate the expression of this gene. Brain tissue from a mouse model of Turner syndrome (39, X^m) was used in this protocol, and the samples were analyzed through PCR amplification with primers designed to capture interacting fragments. No long-range interactions were found with the maternal copy of *Xlr3b*, indicating that the expression is not promoted by a distant enhancer. However, it remains a possibility that the imprinting mechanism of *Xlr3b* is regulated by insulating interactions within the paternal chromosome.

Background

Genomic imprinting is an epigenetic phenomenon where genes show differential expression patterns based on parent-of-origin markers. Imprinting control regions (ICRs) are *cis*-regulatory elements that control the expression of multiple imprinted genes that are clustered over large distances (1Mb) in the genome[1]. ICRs are modified, often through DNA methylation of CpG islands, to distinguish parental origin. This results in the activation of one parental allele while the other remains silenced. Interest and research in the area of genomic imprinting has increased due to its implications in numerous disorders and diseases. It has been shown that imprinted genes are important for normal development, and the over-expression or absence of these genes can result in growth abnormalities and other neurological or behavioral disorders[2].

Imprinting disorders are known to result in neurological and behavioral abnormalities, such as with Prader-Willi and Angelman syndromes. Prader-Willi syndrome is characterized

[1] Ruvinsky, "Basics of Gametic Imprinting."
[2] Bartolomei and Ferguson-Smith, "Mammalian Genomic Imprinting."

behaviorally by obsessive-compulsive manners, psychiatric disturbances, and temper tantrums, while Angelman syndrome is characterized by severe mental retardation, developmental delays, and an overly happy demeanor. These disorders were the first human diseases to recognize the involvement of imprinted genes, and both syndromes result from errors in the chromosomal region 15q11q13. The imprinted genes in this region are only expressed on one allele based on parent-of-origin modifications, and the chromosomal errors result in the loss of all gene expression in these disorders. This can occur from gene deletions, imprinting defects, or uniparental disomy of this chromosome. In the case of Prader-Willi syndrome, there is a loss of paternally expressed genes, and in Angelman syndrome, there is a loss of maternally expressed genes[3].

Recent studies have also shown that genomic imprinting may have a similar role in Turner syndrome and autism. A study conducted by Skuse, et al. investigated the social-cognitive dysfunction levels of girls diagnosed with Turner syndrome, a condition of X monosomy[4]. The researchers found that subjects who inherited a maternal X chromosome had significant deficits in social cognition compared to those who had inherited a paternal X chromosome. This difference led to the conclusion that there is an imprinted locus near the centromere of the X chromosome that is maternally silenced. The researchers noticed a similar trend among normal boys and girls (with boys showing poorer social cognition), and stipulated that males are more prone to having developmental disorders such as autism because they inherit a single maternal X chromosome, where this locus would be silenced.

A later study conducted by Raefski and O'Neill used a mouse model for Turner syndrome to discover X-linked imprinted genes[5]. Researchers found three imprinted genes: *Xlr3b*, *Xlr4b*, and *Xlr4c*. Using allele-specific expression assays, these genes were shown to be maternally expressed and paternally silenced. It was also determined that females with a single maternal X had higher levels of expression of these genes than normal females. This indicates that these imprinted genes are subject to X inactivation, so that a normal female would not express these genes in cells that had inactivated the maternal X. These findings conflict with the conclusion drawn by Skuse, et al., since there was no evidence of maternally repressed genes on

[3] Buiting, "Prader–Willi Syndrome and Angelman Syndrome."
[4] Skuse et al., "Evidence from Turner's Syndrome of an Imprinted X-linked Locus Affecting Cognitive Function."
[5] Raefski and O'Neill, "Identification of a Cluster of X-linked Imprinted Genes in Mice."

the X chromosome. The researchers proposed that the social impairments related to autism and Turner syndrome were due to the cumulative effect of the overexpression of maternal X-linked genes, which would bring the individual closer to the critical threshold of this impairment.

Further research has investigated the gene *Xlr3b* and its effect on cognition in mice. Davies, et al., examined reversal learning abilities with a Y maze using a mouse model for Turner syndrome[6]. It was found that mice with a single maternal X had greater difficulties with reversal learning (where a previously correct answer was now incorrect, or a previously incorrect answer was now correct). This parent-of-origin effect on cognitive function was attributed to an imprinted, X-linked gene. The researchers also confirmed that *Xlr3* paralogs had a greater expression in $39X^{m}O$ embryos than $39X^{p}O$ embryos, with *Xlr3b* specifically showing this difference in expression throughout development. This led to the conclusion that *Xlr3b* may be important in cognitive function and may be related to behavioral inflexibility.

Since *Xlr3b* has been implicated in impairments affecting cognitive function, understanding the imprinting mechanism of this gene has become increasingly important. Previous studies have assessed parent specific CpG methylation patterns of *Xlr3b* to determine if Differentially Methylated Regions (DMRs) within an ICR could be responsible for this imprint[7]. Analyses have shown that these CpG islands within and surrounding *Xlr3b* are completely methylated in both the paternal and maternal copies, showing no differential patterns. This indicates that unlike the majority of imprinted autosomal genes, *Xlr3b* expression is not regulated by a DMR within this region. However, the possibility remains that the imprinting mechanism is controlled by an ICR located greater than 150kb from this gene. This type of interaction has been seen before, and the *Igf2/ H19* locus will serve as a model for this imprinting mechanism.

The *Igf2/H19* locus has been extensively studied to understand the long-range relationship between the ICR and the repression or activation of genes. The ICR lies greater than 80kb upstream from *H19*, and both genes share enhancer sequences downstream from *H19*. For this locus, the ICR is methylated on the paternal chromosome, but is unmethylated on the maternal chromosome. The insulator protein CTCF will bind to the unmethylated ICR and will

[6] Davies et al., "Xlr3b Is a New Imprinted Candidate for X-linked Parent-of-origin Effects on Cognitive Function in Mice."
[7] Carone and Connecticut, *An X-linked Imprinted Cluster Defies the Classical Mechanisms of Epigenetic Regulation*.

block any interaction between the enhancer and *Igf2*. This results in the repression of *Igf2* and the activation of *H19* on the maternal chromosome. CTCF does not bind to the methylated ICR, so the enhancer interacts to allow the expression of *Igf2* instead of *H19* on the paternal chromosome[8]. This relationship has been visualized using chromosome conformation assays, and it was hoped that using a similar method could be used to understand the epigenetic regulation of *Xlr3b*.

The chromosome conformation capture assay allows for the detection of interactions between chromatin segments that are located hundreds of kilobases away from each other. The native, *in vivo*, interactions are induced through a formaldehyde crosslinking step, and the sample is subsequently restriction digested and then ligated together. This process results in the formation of DNA loops which contain the known region of interest and the interacting fragment[9]. It is likely that some loops will be composed of neighboring segments from random ligation rather than a physical interaction. It is therefore necessary to analyze the ligation products to see if the observed interaction occurs due to random chance or as a result of the formaldehyde crosslinking[10]. This procedure has been used with various methods of analysis to further understand the *in vivo* chromatin structure of the *Igf2/H19* locus.

Kurukuti, et al., examined the impact of CTCF binding on the chromatin structure of the *Igf2/H19* locus using this methodology[11]. The researchers focused primarily on long range interactions, and examined a region of over 160kb. Analysis of the products showed that the ICR was in close proximity to the differentially methylated region 1 (DMR1) and the matrix attachment region 3 (MAR3) on the maternal chromosome. The researchers proposed that these interactions silence the maternal copy of *Igf2* by preventing access to the enhancers, unlike the paternal allele where the enhancers have contact with this region. It was also observed that mutations in the ICR that disrupted CTCF binding would cause the DMR1 to become methylated. The maternal chromosome was then observed behaving similarly to the paternal chromosome, with a loss of these looping interactions that silence *Igf2*. It was concluded that CTCF binding performs a critical role in the maternal silencing of *Igf2*, and that it also regulates

[8] Bartolomei and Ferguson-Smith, "Mammalian Genomic Imprinting."
[9] Dekker, "Capturing Chromosome Conformation."
[10] Hagège et al., "Quantitative Analysis of Chromosome Conformation Capture Assays (3C-qPCR)."
[11] Kurukuti et al., "CTCF Binding at the H19 Imprinting Control Region Mediates Maternally Inherited Higher-order Chromatin Conformation to Restrict Enhancer Access to Igf2."

epigenetic marks on the DMR1. Observing chromosome conformations allowed researchers to recognize the importance of CTCF in the epigenetic regulation of the chromatin structure of *Igf2/H19*.

In another study conducted by Qiu, et al., this technique was used to understand the silencing of the maternal *Igf2* allele[12]. The ligation products were analyzed through PCR amplification, quantification on polyacrylamide-urea gels, and direct sequencing. The researchers found that the maternal *Igf2* allele formed a complex knotted structure where there were multiple loops involving the ICR, DMR1, and enhancers. Previously, it had been proposed that a single chromatin loop was responsible for the maternal silencing at *Igf2*. The researchers suggested that the silencing of *Igf2* on the maternal allele is due to the binding of CTCF in addition to the multiple chromosome loops that affectively block the enhancers from interacting with this gene.

The chromosome conformation capture protocol was successful at identifying multiple short and long range interactions in the *Igf2/H19* locus. This assay has allowed researchers to observe interactions between enhancer sequences and expressed genes, as well as looping interactions that contribute to the silencing of genes. It is hoped that this protocol can be used to determine if similar interactions are occurring in *Xlr3b*. Comparing X^m and X^p samples may show an enhancing interaction that is present in the maternal sample but absent in the paternal sample, or a complex looping system in the X^p sample that blocks its expression. It is unknown where an interaction would be occurring or if any interaction will be seen. The objective of this study is to observe a difference between the maternal and paternal copies of this gene, giving insight into the imprinting mechanism of *Xlr3b*.

Methods

The protocol used was taken from *Quantitative Analysis of chromosome conformation capture assays*, which is detailed with all modifications in the following section[13]. *Xlr3b* is over

[12] Qiu et al., "A Complex Deoxyribonucleic Acid Looping Configuration Associated with the Silencing of the Maternal Igf2 Allele."
[13] Hagège et al., "Quantitative Analysis of Chromosome Conformation Capture Assays (3C-qPCR)."

12kb in length. EcoRI was chosen as the restriction enzyme for the digestion because it leaves cohesive ends to facilitate an effective ligation, and because this enzyme cuts *Xlr3b* six times, giving manageable regions ranging from approximately 1kb-3kb in length. Samples were taken from a mutant strain of *Mus musculus domesticus* (from C3H inbred strains) that produce X monosomic females. *Patchy fur* (Paf) males produce sperm that lack a sex chromosome due to an inversion on the X chromosome in the pseudo-autosomal boundary. Paf males crossed with wild-type C57BL/6J females produce 39, X^m mice at a frequency of around 30%[14]. 39, X^m mice were identified using a PCR genotyping analysis of strain-specific X-linked markers.

Cross-linking

Whole brains were taken from neonatal 39,X^m mice, were weighed, and placed in a petri dish on ice with 250µl of ice cold PBS per 50mg of tissue. Brains were chopped for two minutes with a blade, and the tissue was then transferred with a pipet to a 50mL falcon tube on ice. The tissue was homogenized by passing it through an 18G needle ten times and a 21G needle 20 times. The tube was taken off ice, and 500µl of room-temperature PBS was added. Next, 13.5µl of 37% formaldehyde was added for every 500µl PBS to cross-link the DNA. This was incubated for ten minutes, while gently swirling the solution every two minutes. To quench the cross-linking reaction, 1.25M glycine was added to a final concentration of 0.125M, and this was incubated for five minutes, while gently swirling every two minutes.

To wash and isolate the cells, the solution was aliquot in half into two 1.5mL micro-centrifuge tubes, and was centrifuged at 4˚C, 200g, for ten minutes. The supernatant was removed and the cell pellet was resuspended in 500µl of ice cold PBS. This was centrifuged again at 4˚C, 200g, for ten minutes. The PBS was removed, and the cell pellet was washed with an additional 500µl of ice cold PBS, followed by centrifuging at 4˚C, 200g, for ten minutes, and removing the PBS to isolate the cell pellet.

Cell Lysis

10mL of fresh lysis buffer was made at the concentrations of 10mM Tris-HCl (pH 8.0), 10mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, and 1× complete protease inhibitor. 150µl of ice cold

[14] Raefski and O'Neill, "Identification of a Cluster of X-linked Imprinted Genes in Mice," -.

lysis buffer was added to the tissue. After a gentle vortex, this was incubated for ten minutes on ice. The solution was centrifuged at 4˚C, 400g, for five minutes, and the supernatant was removed. The pelleted nuclei were frozen in liquid nitrogen and stored at -80˚C until future need.

Digestion

The pelleted nuclei were resuspended in 500µl of 1.2× Buffer 2. 2.5µl of 20% SDS was added for a final concentration of 0.2%, and this was incubated for one hour at 37˚C while shaking at 900 r.p.m. 3.75µl of 20% Triton X-100 was added for a final concentration of 1.5%. This was incubated for one hour at 37˚C while shaking at 900 r.p.m. 5µl of the sample was removed and stored at -20˚C as the undigested control to later be used to calculate digestion efficiency. EcoRI was then added to the sample at a concentration of 2.7 units per 1µg of DNA, and the sample was incubated overnight at 37˚C while shaking at 900 r.p.m. After digestion, 5µl of sample was removed as the digested control for calculating digestion efficiency.

Ligation

40µl of 20% SDS was added to the sample, and it was incubated for 25 minutes at 65˚C while shaking at 900 r.p.m. The sample was transferred to a 15mL falcon tube with 5.421mL of water and 704µl of 10× ligation buffer (660mM Tris-HCl, 50mM DTT, 50mM $MgCl_2$, 10mM ATP). Next, 375µl of 20% Triton X-100 was added, and the sample was incubated for one hour at 37˚C while shaking at 250 r.p.m. 100 units of 400U/µl T4 DNA Ligase was added, and the sample was incubated for four hours at 16˚C followed by 30 minutes at room temperature. The crosslinks were reversed by adding 300µg of Proteinase K and incubating at 50˚C overnight.

DNA Purification

The sample was incubated for 30 minutes with 30µl of 10mg/mL RNase A. The DNA was isolated from the sample using a phenol-chloroform extraction with 3.5mL phenol and 3.5mL 24:1 chloroform:isoamyl alcohol. After mixing vigorously, the sample was centrifuged for 15 minutes at 2,200g at room temperature. The supernatant was transferred into a new 15mL falcon tube, and the extraction and centrifuging was repeated using 7mL of chloroform:isoamyl alcohol. The supernatant was transferred into a 50mL falcon tube. 7mL of distilled water, 1.5mL of 2M sodium acetate (pH 5.6) and 35mL of ethanol were added, and this was placed at -80˚C

for one hour. Next, this solution was centrifuged for 45 minutes at 2,200g at 4˚C. The supernatant was removed, and 10mL of 70% ethanol was added to the DNA pellet. This was centrifuged for 15 minutes at 2,200g at 4˚C. Afterwards the supernatant was removed and the pellet was air dried for five minutes. The DNA pellet was then dissolved in 150µl of 10mM Tris-HCl (pH 8.0), and was stored at -20˚C until ready for analysis.

Results

Digestion Efficiency

It was necessary to test the digestion efficiency to ensure that a substantial amount of ligation products could be formed. The undigested and digested controls were treated with 500µl 1× PK buffer and 1µl 20mg/mL PK at 50˚C overnight to reverse the cross-links, followed by adding 1µl of 1mg/mL RNase A at 37˚C for 2 hours to degrade any RNA in the samples. The DNA was then isolated with a phenol-chloroform extraction and ethanol precipitation. 250µl of phenol and 250µl of chloroform:isoamyl alcohol was added to the controls, and these were centrifuged for five minutes at 16,100g at room temperature. The supernatants were transferred into new tubes with 50µl 2M sodium acetate and 1.5mL ethanol. These were placed at -80˚C for one hour, followed by centrifuging for 20 minutes at 16,100g at 4˚C. The supernatants were removed, and the cell pellets were washed with 500µl of 70% ethanol. These were centrifuged for 4 minutes at 16,100g at room temperature. The ethanol was removed, and the pellets were briefly air-dried before resuspending in 60µl of water.

The two controls were compared using a real-time PCR quantification with primers designed to amplify three regions containing EcoRI sites in the genes *F8a* and *Igf2*, and a control primer set in *F8a* that did not contain an EcoRI site. The digestion efficiency was calculated using the equation:

$$\% \text{ restriction} = 100\text{-}100/2\text{^}((Ct_R\text{-}Ct_C)DIG - (Ct_R\text{-}Ct_C)UND)$$

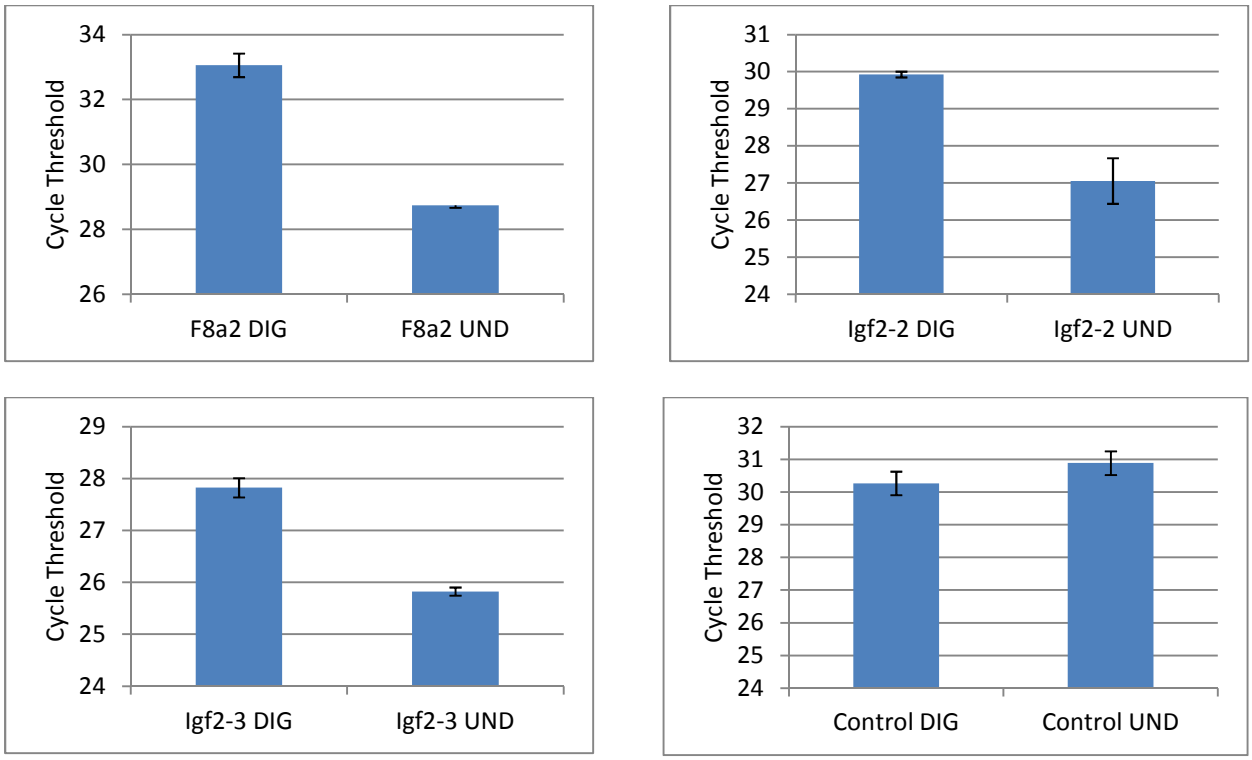


Figure 1: Ct values from real-time PCR used to compute digestion efficiency. The digested (DIG) and undigested (UND) controls for sample X^m830(1) were compared with a real-time PCR with primer sets in *F8a* and *Igf2* that spanned EcoRI restriction sites and a control primer set in *F8a* that did not span an EcoRI site. The digested control had an amplification threshold approximately 2-4 cycles later than the undigested control at primers spanning the restriction site and one cycle earlier at the control primer set.

Samples ideally should have a digestion efficiency of at least 80%, and any samples with digestion efficiencies below 60-70% were discarded. The sample X^m830 (1) passed this test with digestion efficiencies of 96.7% and 83.7-91.1% at *F8a* and *Igf2* respectively (Figure 1).

DNase/Exonuclease

It was also necessary to test the ligation efficiency before proceeding to the analysis of the samples. This was done by using the 454 Rapid Paired End Library Prep kit to degrade any linear DNA so that the sample would only contain circularized fragments. First, an aliquot of the sample was diluted to a concentration of 1ng/µl at a final volume of 100µl. Next, 1.1µl of 100mM ATP and 5µl of 10 U/µl plasmid safe, ATP-dependent DNase were added to the sample to degrade double stranded linear DNA. Single stranded DNA was removed by using 3µl of 20 U/µl Exonuclease I. The sample was then incubated for a half an hour at 37˚C. 1µl of 5ng/µl

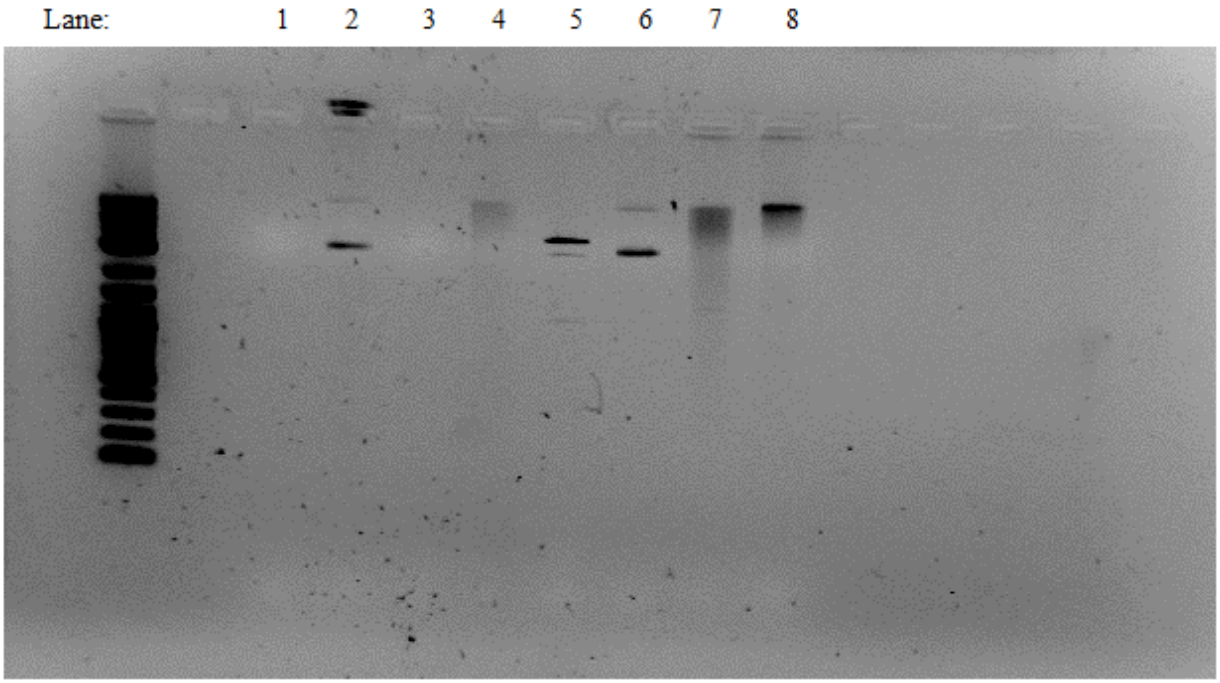


Figure 2: Comparing DNase/Exonuclease treatment in digested and undigested control and sample. The GFP vector was used as a control to compare to the X^m830(1) sample. Samples from lanes 1-4 were treated with the 454 Rapid Paired End Library Prep, and samples from lanes 5-6 did not undergo this treatment. Lanes 1 and 5 have GFP vector that was digested with EcoRI, and lanes 2 and 8 contain undigested GFP vector. Lanes 3 and 7 have the sample that was digested with EcoRI, and lanes 4 and 8 contain undigested sample.

glycogen was used as a carrier to maximize the DNA yield since the sample was so dilute. This DNA was then washed and eluted using the standard QIAquick purification kit.

To determine if ligation occurred, the aliquot of sample that had been DNased and exonucleased was compared to an equivalent aliquot of sample that had not been through this process. GFP vector was used as a control and was treated with the same protocol. These samples were run out on a 1% agarose gel along with aliquots of sample and control that were digested with 1 unit of EcoRI for one hour at 37˚C (Figure 2). This gel showed that the 454 Rapid Paired End Library Prep was effective at degrading linearized DNA and did not affect circularized fragments. The undigested GFP vector remained intact while the digested GFP vector was completely degraded and no longer visible. The digested sample was also completely degraded compared to the undigested sample.

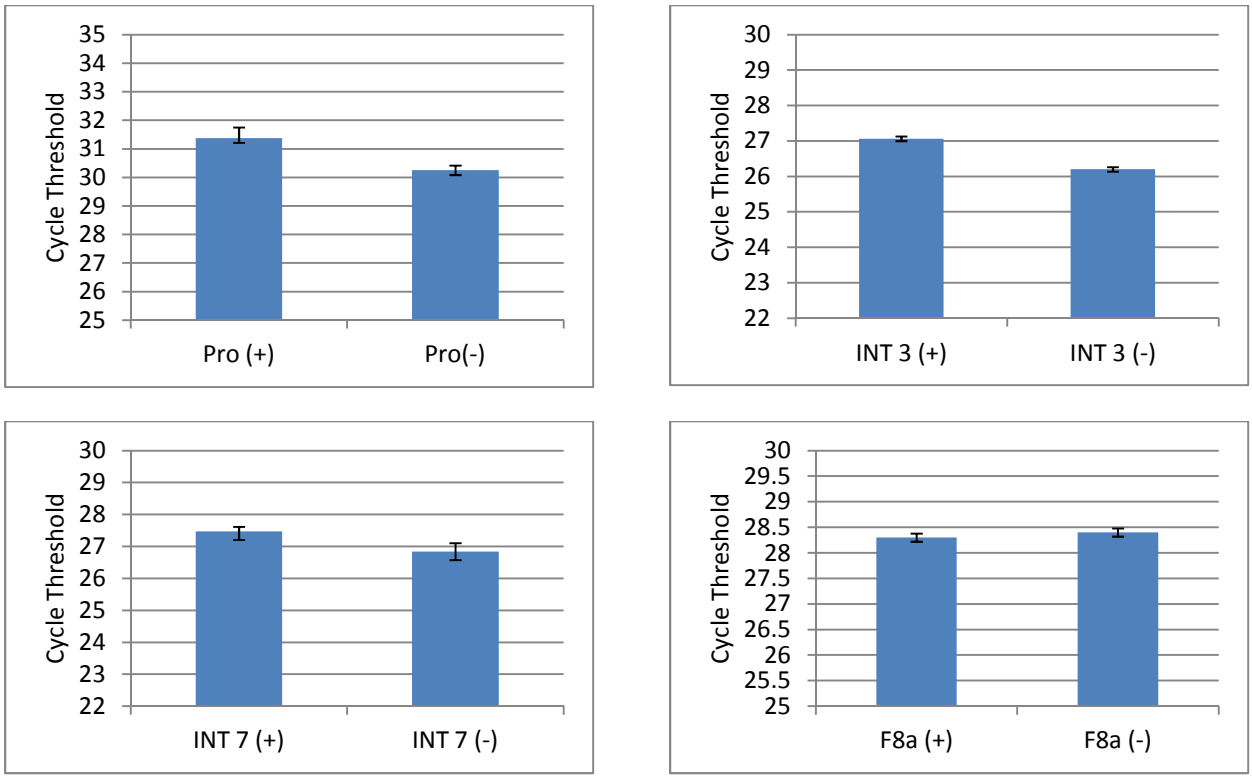


Figure 3: Ct values from real-time PCR comparing DNased/Exonucleased and untreated samples. The X^m830(1) sample was treated with the 454 Rapid Paired End Library Prep and was compared to an aliquot of untreated sample. Primers were designed amplify approximately 150 base pairs spanning an EcoRI site in *Xlr3b* (promoter, intron 3, and intron 7) and *F8a*. The sample treated with DNase and exonuclease (+) had a Ct value less than one cycle later than the untreated sample (-) at these sites.

Furthermore, it appeared that the sample was mainly composed of circularized DNA. The gel shows that the sample that underwent this treatment had a smear that corresponded to the smear of untreated sample. This smear was fainter than the untreated sample for numerous reasons. Some of the sample did not ligate to form circular fragments and was therefore degraded, but the aliquot also became more dilute after the DNase/Exonuclease treatment and some DNA was most likely lost during the purification. For these reasons, although it may seem from the picture that a small portion of the sample remained after the DNase/Exonuclease treatment, this is an underestimation of true concentration of circularized fragments.

The DNased and exonucleased sample and the untreated sample were then compared with a real-time PCR to provide an additional quantification assessment. This test showed that there was not a significant difference in amplification thresholds between the two samples. At

four different locations spanning EcoRI sites, the DNase/Exonuclease sample had a Ct value less than one cycle later than the untreated sample (Figure 3). This indicates that a fraction of the sample was degraded through this treatment, but the majority of the sample was intact and was therefore circularized. Through these assessments, it was determined that the ligation was effective and the sample could be further analyzed for unknown interactions.

Control Interaction

It was also necessary to test the sample to see if a known interaction could be visualized and sequenced. Although the procedure was taken from a paper based on the *Igf2/H19* locus, these genes have very low expression in brain and would therefore be inadequate as a control in this study. Instead, the *COX* subunit genes and *Tf* genes served as the control since these genes are highly expressed in brain. The ten nuclear *COX* subunit genes code for the enzyme cytochrome c oxidase along with the *Tf* genes, which allow for the transcription of three mitochondrial-encoded *COX* subunit genes. A study conducted by Dhar, et al., used the chromosome conformation capture assay to find that the mitochondrial transcription factors and nuclear *COX* subunits physically interacted with each other during transcription[15].

The analysis from this paper was replicated to see if these same interactions were present in the X^m830 (1) sample. Primers were designed 50-150 base pairs away from EcoRI sites on each gene. A standard PCR was conducted using forward primers from *COX4i1* and *COX6a1* paired with reverse primers from *Tfam* and *Tfb2m*. This allows for the selective amplification of looping interactions between the *COX* subunit gene and *Tf* gene that have been ligated together at the EcoRI site. The cycling conditions followed the strict procedure in the paper to give only the specific interactions of interest: denaturation for 2 minutes at 94˚C, followed by 36 repeats of 30 seconds at 94˚C, 15 seconds at 59˚C, and 15 seconds at 72˚C, and a final elongation step of 1 minute at 72˚C.

The PCR products were then run out on a 1% agarose gel (Figure 4). This gel shows that the interaction between *COX4i1* and *Tfb2m* and the interaction between *COX6a1* and *Tfb2m* were detected. The desired product for each interaction should have been approximately 200

[15] Dhar, Ongwijitwat, and Wong-Riley, "Chromosome Conformation Capture of All 13 Genomic Loci in the Transcriptional Regulation of the Multisubunit Bigenomic Cytochrome C Oxidase in Neurons."

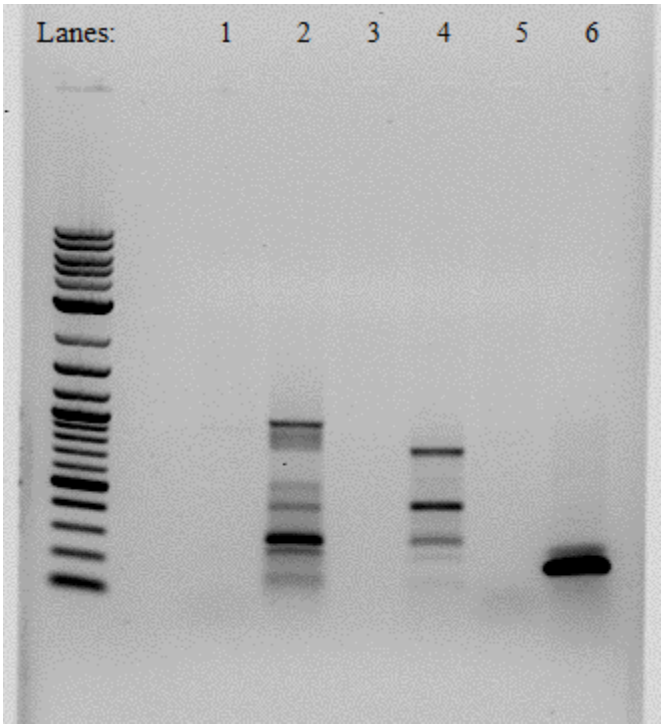


Figure 4: Gel picture of control interactions between *COX* subunit genes and *Tf* genes. Forward primers from the *COX* subunit genes were paired with reverse primers from the *Tf* genes. Lane 1: *COX4i1 / Tfam*, Lane 2: *COX4i1 / Tfb2m*, Lane 3: *COX6a1 / Tfam*, Lane 4: *COX6a1 / Tfb2m*, Lane 6: Positive control.

base pairs in length. The larger bands seen on the gel were possibly due to unsuccessful digestion at the restriction site closest to one of the primers. This would give a larger product with additional length until the next successful digestion site. These PCR products were then cloned using the StrataClone PCR Cloning Kit and a standard TA cloning protocol. Sanger sequencing was then used to identify the nucleotide sequences of these products, and analysis of these sequences confirmed that the interaction of interest was being detected.

Analysis of X^m Interactions

The X^m830 (1) sample was analyzed using PCR amplifications followed by cloning and sequencing. This assessment looked at the five regions of *Xlr3b* that were created from the EcoRI digestion: INT1-EX4, INT4-INT5, INT5-INT7, INT7-EX9, and EX9-UTR3. Primers were designed 50-100 base pairs away from the restriction sites facing outwards (Figure 6A). This would allow for the amplification of primarily the unknown interacting sequence rather than the "bait" or known region of *Xlr3b*. These primers were used in an Expand High Fidelity PCR

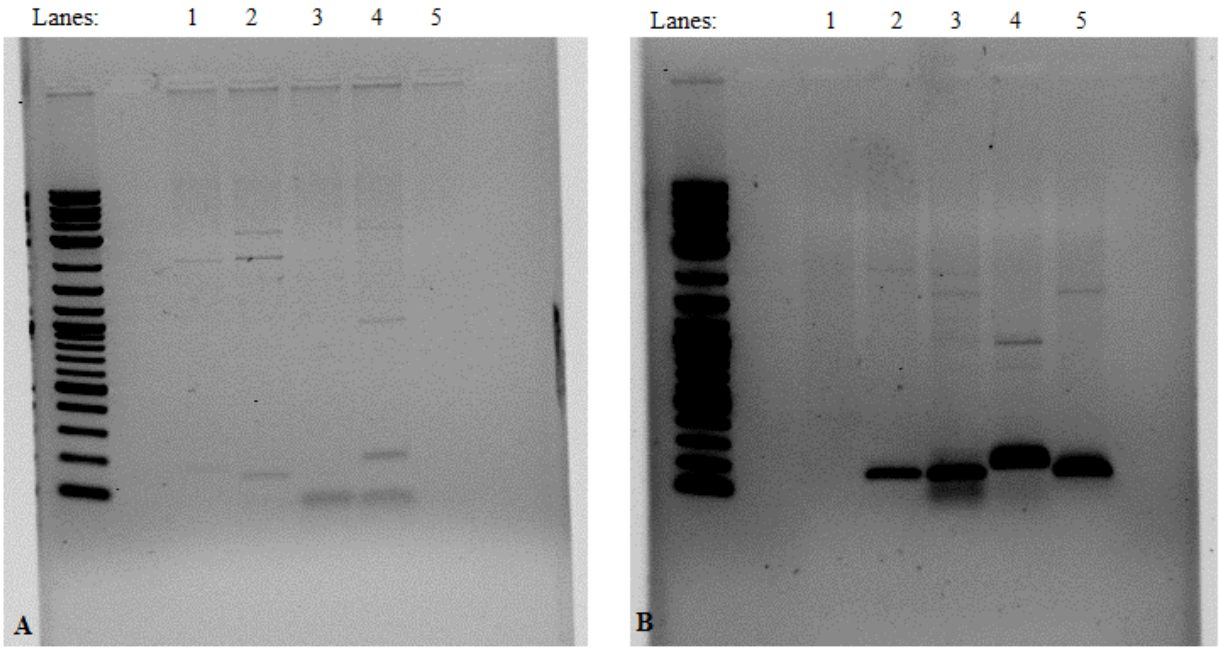


Figure 5: High-fidelity PCR (A) and Standard PCR (B) with X^m830(1) sample. Lane 1: INT1-INT4, Lane 2: INT4-INT5, Lane 3: INT5-INT7, Lane 4:INT7-EX9, Lane 5: EX9-UTR3. Bands at 100-200 base pairs correspond to sequences with no insert. Bands from 1kb-3kb represent products that may have captured an insert in between the primers.

(using the Roche kit and protocol) and in a standard PCR with an extended elongation time of four minutes and annealing temperature of 60˚C. Products from these PCRs were run out on 1% agarose gels (Figure 5). Both gels showed distinct bands, with the smaller sized bands more easily visible, especially with the standard PCR. Fainter bands of approximately 1kb to 3kb in length were visible in some of the products.

These PCR products were then cloned and sequenced using the StrataClone PCR Cloning Kit. The majority of the sequences obtained from this procedure showed that the "bait" region ligated to itself (Figure 6B). The sequence began at the forward primer, and after the digestion site the sequence continued as the beginning of the same *Xlr3b* region up to the reverse primer. There was no insert or captured fragment in these clones. These sequences were approximately 150 base pairs in length and accounted for the smaller, dark bands on the gel. There were a small percentage of the clones where the sequences continued past the EcoRI site into the next region of *Xlr3b* (Figure 6C). At the subsequent restriction site, the sequence went back to the beginning

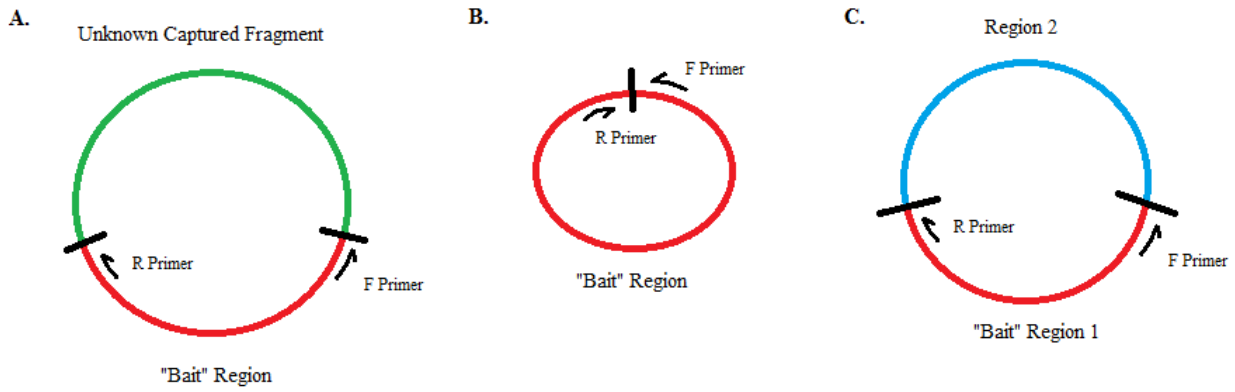


Figure 6: Diagram of PCR amplification of unknown fragment. (A) Diagram of theoretical interaction: Primers were designed in the known regions of *Xlr3b* near the outer EcoRI restriction sites. The forward and reverse primers both faced towards the captured fragment to amplify only this unknown region. (B) Diagram of observed amplification: The majority of the sequences showed that the "bait" ligated to itself with no interacting fragment in between the two restriction sites. (C) Diagram of observed amplification: Few sequences showed that the circular fragments were composed of two neighboring *Xlr3b* regions.

of the original "bait" region. This represents occurrences when the circularized fragment was composed of two neighboring *Xlr3b* regions and could account for the larger bands of 1-3kb on the gel. Using this method, an unknown interacting fragment was not captured and sequenced.

Since no interactions were found with Sanger sequencing, the sample was treated in order to go into high-throughput sequencing. This would give a greater depth of reads with hundreds of millions of short sequences. The sample was prepared with the SOLiD Fragment Library Barcoding kit and protocol. A total of approximately 2µg of DNA sample was used to enter this procedure, and after shearing and size selection, there was approximately 500ng of DNA left in the SOLiD library. This falls within the expected recovery range of 20-30%. The library was diluted to a concentration of approximately 3ng/µl, and this was then analyzed with a Bioanalyzer to determine the size and relative concentrations of the fragments in the library (Figure 7 and Figure 8).

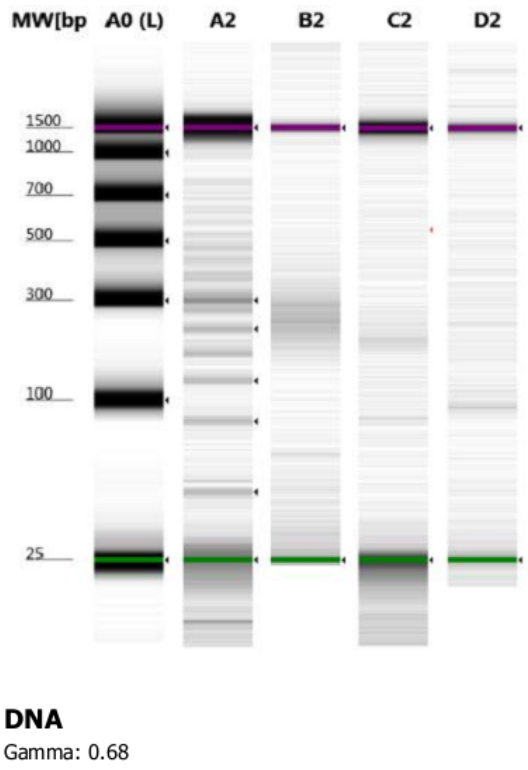
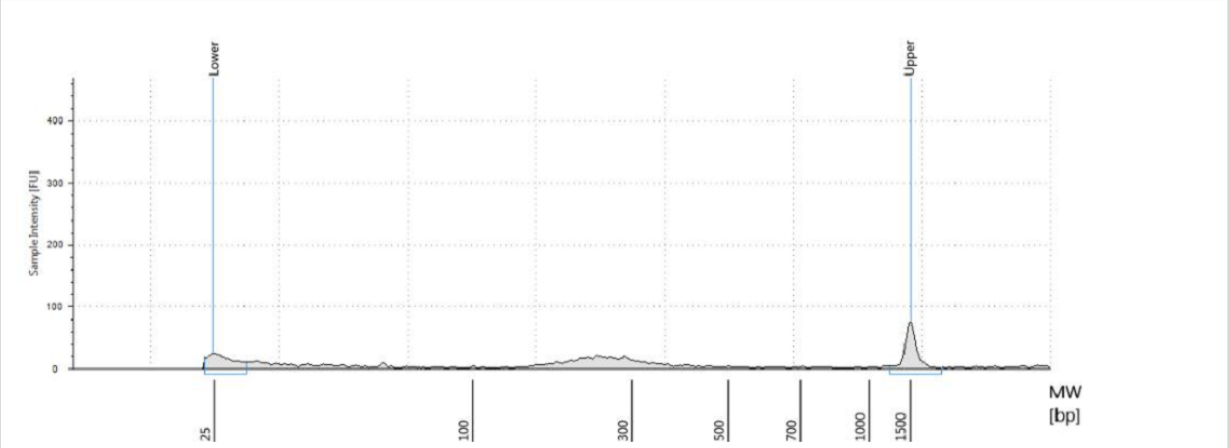


Figure 7: Gel picture from Bioanalyzer. X^{m}830(1) SOLiD library is in lane B2.



□avelengt□	M□ [bp]	Conc□[pg/µl]	Molarit□[pmol/l]	□ of Integrated Area	□ea□Comment	Observations
Sampe	25	405	24500			Lower Marer
Sampe	1500	580	586			pper Marer

□rom [bp]	□o [bp]	Average Si□e [bp]	Conc□[pg/µl]	Molarit□[pmol/l]	□ of □otal	□egion Comment	□avelengt□
138	418	267	1020	5800	50.9		Sampe

Figure 8: Graph of intensities by molecular weight for SOLiD library. The lower limit is at 25bp and upper limit is at 1500bp. The average size fragment in the library is 267bp.

Discussion

The PCR amplification and Sanger sequencing of the X^m830 (1) sample showed that there were no interactions with *Xlr3b*. From these results it would appear that unlike at the *Igf2/H19* locus, there is no long-range enhancer interaction on the maternal X chromosome that facilitates the differential expression of *Xlr3b*. Ideally, the paternal and maternal chromosomes should be compared to see if there are any differences in chromatin structure that could account for parent-of-origin expression patterns. This study only analyzed the maternal X chromosome due to time limitations. It is still possible that the paternal X chromosome will show an insulating interaction that is responsible for the silencing of the gene. It may also prove beneficial to look at a different tissue source (such as liver) where the *Xlr3b* imprint is stronger than in brain. This may allow for the capture of an interaction that was not seen in this study.

The results showed that some of the circularized fragments were composed of two neighboring *Xlr3b* regions. Ligation between two separate neighboring regions is highly likely since the probability of random ligation increases with genomic proximity[16]. However, this also could have occurred if EcoRI did not cut at the restriction site separating the two regions of interest. This single adjoined fragment could then ligate to itself, similar to what was seen with the majority of the sequences. The digestion efficiency was calculated by looking at only three locations, so it is possible that this was not representative of the entire genome and that the restriction enzyme did not have equal access to all genes. It may be necessary to optimize the digestion by testing various levels of SDS and triton-X. A more efficient digestion would increase the number of ligation products that could be formed, which could be beneficial if an interaction was a rare event.

It is possible that there was an interaction between the maternal copy of *Xlr3b* and an unknown region that was not captured in this analysis. The high fidelity PCR is only optimized to amplify fragments up to 5kb in length. If the genome was completely random, a six-base cutter like EcoRI would on average cut the sequence at about every 4kb, but realistically the length of an interacting region is unknown and could be larger than 5kb. Also, as stated above, the digestion efficiency may have been over-estimated, which would make the captured fragment

[16] Hagège et al., "Quantitative Analysis of Chromosome Conformation Capture Assays (3C-qPCR)."

even longer and more difficult to amplify. Cloning and sequencing also favors inserts of smaller length over longer ones, creating a bias towards seeing self-circularization of the region over ligation to a large interacting fragment. This problem could be addressed by using two restriction enzymes to increase the frequency of digestion. This would create fragments of smaller length that would be easier to PCR amplify, clone, and sequence.

It is also possible that the four and a half hour incubation was not long enough to allow for two ligation events to occur. Although it appeared that the majority of the sample was circularized from the gel picture and real-time PCR, the sequencing showed that these products were largely the result of a single ligation event of the fragment to itself. If there was a long-range interaction with *Xlr3b*, the two fragments may have ligated together at only one restriction site, resulting in a linear product. If this interaction was a rare event, it could account for the slight differences seen in the real-time PCR (Figure 3). With the primers designed within *Xlr3b*, the thresholds were one cycle later for the DNased/Exonucleased samples. The primer designed in *F8a* came up at the same threshold for both the treated and untreated sample. This could indicate that there were linear products in *Xlr3b*, possibly from a long-range interaction, that were absent in *F8a*, a non-imprinted gene.

If the product from an interacting sequence and *Xlr3b* was linear, it would not have been captured in this analysis. Primers were designed outwards from the "bait" so that only circularized sequences could be amplified. It would be impossible to PCR amplify or Sanger sequence this theoretical product because the interacting sequence is unknown, so primers could not be designed to capture the entire length of the linear fragment. In future analyses, it may prove beneficial to increase the ligation time to ensure that two ligation events can occur. This may yield an interaction that was not captured in this study.

Although the results of this study show no long-range interaction between *Xlr3b* and another sequence, further analysis implementing the proposals discussed above should be conducted to either give strength to this claim or to capture an interaction that was not seen in this study. Optimizing the digestion, using a combination of restriction enzymes, and increasing the ligation period could allow all unknown products to be analyzed with standard amplification and sequencing procedures. Furthermore, using a different tissue source and performing a comparison to the paternal chromosome could provide further insight into the imprinting

mechanism of this gene. The results from the Bioanalyzer show that the SOLiD library is within the correct size and concentration ranges to go into emulsion PCR. Analyzing the $X^m830(1)$ sample with this method of high-throughput sequencing should yield a great depth of reads that could capture a sequence that was not seen with Sanger sequencing. Continuing with the SOLiD sequencing will give a clearer picture of the chromatin structure of *Xlr3b* and may reveal interactions within the maternal copy of this gene that could affect its regulation.

Works Cited

Bartolomei, Marisa S, and Anne C Ferguson-Smith. "Mammalian Genomic Imprinting." *Cold Spring Harbor Perspectives in Biology* 3, no. 7 (July 2011). http://www.ncbi.nlm.nih.gov/pubmed/21576252.

Buiting, Karin. "Prader–Willi Syndrome and Angelman Syndrome." *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* 154C, no. 3 (August 15, 2010): 365–376.

Carone, Benjamin R., and University of Connecticut. *An X-linked Imprinted Cluster Defies the Classical Mechanisms of Epigenetic Regulation*. ProQuest, 2008.

Davies, William, Anthony Isles, Rachel Smith, Delicia Karunadasa, Doreen Burrmann, Trevor Humby, Obah Ojarikre, et al. "Xlr3b Is a New Imprinted Candidate for X-linked Parent-of-origin Effects on Cognitive Function in Mice." *Nature Genetics* 37, no. 6 (June 2005): 625–629.

Dekker, JobRippe. "Capturing Chromosome Conformation." *Science* 295, no. 5558 (February 15, 2002): 1306.

Dhar, Shilpa S, Sakkapol Ongwijitwat, and Margaret T T Wong-Riley. "Chromosome Conformation Capture of All 13 Genomic Loci in the Transcriptional Regulation of the Multisubunit Bigenomic Cytochrome C Oxidase in Neurons." *The Journal of Biological Chemistry* 284, no. 28 (July 10, 2009): 18644–18650.

Hagège, Hélène, Petra Klous, Caroline Braem, Erik Splinter, Job Dekker, Guy Cathala, Wouter de Laat, and Thierry Forné. "Quantitative Analysis of Chromosome Conformation Capture Assays (3C-qPCR)." *Nature Protocols* 2, no. 7 (2007): 1722–1733.

Kurukuti, Sreenivasulu, Vijay Kumar Tiwari, Gholamreza Tavoosidana, Elena Pugacheva, Adele Murrell, Zhihu Zhao, Victor Lobanenkov, Wolf Reik, and Rolf Ohlsson. "CTCF Binding at the H19 Imprinting Control Region Mediates Maternally Inherited Higher-order Chromatin Conformation to Restrict Enhancer Access to Igf2." *Proceedings of the National Academy of Sciences of the United States of America* 103, no. 28 (July 11, 2006): 10684–10689.

Qiu, Xinwen, Thanh H. Vu, Qiucheng Lu, Jian Qun Ling, Tao Li, Aiju Hou, Shu Kui Wang, Hui Ling Chen, Ji Fan Hu, and Andrew R. Hoffman. "A Complex Deoxyribonucleic Acid

Looping Configuration Associated with the Silencing of the Maternal Igf2 Allele." *Molecular Endocrinology* 22, no. 6 (June 2008): 1476–1488.

Raefski, Adam S, and Michael J O'Neill. "Identification of a Cluster of X-linked Imprinted Genes in Mice." *Nature Genetics* 37, no. 6 (June 2005): 620–624.

Ruvinsky, A. "Basics of Gametic Imprinting." *Journal of Animal Science* 77 Suppl 2 (1999): 228–237.

Skuse, D. H., R. S. James, D. V. M. Bishop, B. Coppin, P. Dalton, G. Aamodt-Leeper, M. Bacarese-Hamilton, C. Creswell, R. McGurk, and P. A. Jacobs. "Evidence from Turner's Syndrome of an Imprinted X-linked Locus Affecting Cognitive Function." *Nature* 387, no. 6634 (June 12, 1997): 705–708.

Supplemental Materials

Map of *Xlr3b* locus with EcoRI restriction sites and primers:

EcoRI sites highlighted in dark blue
Forward primers highlighted in light blue
Reverse primers highlighted in yellow
Exons highlighted in green
UGNC highlighted in grey
UTR highlighted in red

XLR3B SPECIFIC
3B+OTHER PARALOG

>XLR3B
GTAATGCTGGGACCACAGTCTTTGGGAGAGGATGGTGTGGATCACAAGGAGACTTGCCTAAGAGGGCCTGTGATGTA
ATGTGCCTGAGTCTGGCTTTAGCCTCTGACTCGACTAGTAACCCCTGGAGACTCATTGCAGAAGTTGCAGAGATAGA
GAGCCTGGGATGTCTGAGCCGGACACTGCGGTCAAGGTACGGGCTCATAGTATCCAACATCTGCCCGTTTTGTACTG
TTTAGGGCCAGACTGAACACCCGAGGCCTCTGAGATACAGCCCAGGGCATTCTGCAATGGGAGCACCAATCAGAGAG
AAAAGGACACAGGACCCTTCAGGCATGCACCCTAGTGTACACTTGGGACTGGTAGGTTTTCCTATCCCTGGAGCATT
GGCTGTGCTCTGTTTGCGCAAGCTTACAGGAGCATCTCATCTTTGCACATGAAAATTCTTTTCCCTCACAGGACCCA
ATGGACCTCTCGCTGATGCACATTTAGGCCTCTTTCTGTTGCAGCGAATCAGAGATACTTACATAAGGAGATGGGCA
AAATGGAGATGGATTGTAGCTGTGGCCCAATGAAATGAAGCCACGGGCCATCCATCACCAGGAAAGCAAATGTGAGT
GGATGAGGCATGCTGGGCCCCTTGATTTATGCGCAGAGTGCTGGTACCCAGTGCGCGTGTGCAAGCTGGTACACCAG
CCAATCAGAGAACATTATATCAGCCAAGGCCCGACCAAGTGGGATGAACCTCTGAGTGGAAGGAGGCTGCCTTGTGG
AGAGGAGGGGACTAGGTGGGTGACGAGTTTTGCCGGGGGGTGGGGTGGGGGGGGGTGGGGACGTTCCACAGTGGGAC
TCCGCGGGTGGTTAGGAGGCACCCTTGTTGTCCTCGTCGTCATTGTAGTGGGACAGTGGGGTCACAGAATATGAGGG
TCCCGGGAAGCCTGTGAGGTGACTTCCGGAGCTTTCTTTTACTGTAAGAGGCGCCAGAGCTGAGAATGAGAAGCCAG
GAATGAGAAAAACTCCCAGTTTGGGGTTGTTTTTCGTTGTTCTTTGGTTTTGTTTGGTTTTTGTGGTGGTGGTGGTG
GTTTTTTTTTTTTTTTTTTTTTTGTTGTTGCTGGGTCTTGTTGTTGTTTGTTTGTTTTGGTTGTTTTTTTTT
GTTTGTTTGTTTGGTTCCTGACTGCACACCTGGATCCTCCTGTGACACTTCCATAAGCAGTATCTGAGCAGGATAGT
TAGTCTGAAGCGGAAGGGGGAAACCCTGGGCACTTTGCACAAACACATGGGCAACGACAGGGCCCAAGCTGGGAATG
GGTACCTGATTGGCAGGCCTCACACCTTTGGGAGTACCTGTAGGCAGCCTCAGGCATGCCTGCACAGTAGTGTTTCT
GATAATCAGAGACCTACTCCAGAATGTGGGGCCTCTATCTGAGGGGCGGGGGAAATAGCCTCTCAGAGGAATGTATT
GAAGTGTGGTCACCTGGGTGAGTGACAGGGAATTAGCAGAGTCAGTCTTGAGCCTCTGATGGTGAGAAAGCTCTGTC
AGCCTGGTGTATCTGGCCTTTTGGGGAGCAAAGTGGGGTCTTGAGGCAGAAGGGACCAGGTTTAGCTGTGATGAGAT
GTTTGGAAGACTTTGCACCTGGTAGTAAAATAGACCAGTGAGACCAGTTGATGAAAATGGTTCCAGGAGCGTTGCAG
GTTTGAAGAGAGTTGGCTGGGACAGTGACTTGGAGAAAAGGAAGTAGTGACTCTCCCCTGTCCTCCAGGGGTCCTCA
GTGGCCACATCTGGGTCCTGCTTCTTTGCAGCCATGGGCCTTGCCTCCCTCAACACAAATGCCTGTGTTTTTGTATG
TTTTTCTCAAAGTGTCAAGATTTGGTCAAGATCTAGAATATTGGTGGGCCATAGGAGGTATTCCTAAAATCTGTGTT
GAATGTTGAAGGGGATGGGCCAGGACTGGTCCCCTTCACCAAAGGACAGCATCTCCATCCAAATCCCACAAGTACTA
TACTAGCTGAAACCACTTAACCCTGGGTATACACACCTTCCCAGCCAGTCAGAGAGATGTGCACTGGGAGGTGGTTC
CTACATAGGTATGACCAACCACCCTCTGAAAAGGAGGAGGCTCCAGGGGTAGAAGAGAACCCTGGGTGGATGATTGG
CTATTGGGGAACTCTGTTTAGGGGCTCTCCAGTTGATACAGGGTTCAGAAAAGAGCCTTAGCTCGTGTGGTAGGGCA
GTGGGATCACAAAGAGGCAGGTCCAAACCAAACCGGTGGGACCATCCTAGGACCTTCACTTTCCAAAGGAGGCCTGG
CAAGGCCACAGAGGGCAAGGCAAGGGAGCAGTCATGGGGCTGCATGGATTGGGAGGGCTAAATGGGACAGCAAGTGG
GCTGTGAGACATGGCCTCCCTGACTGATGAGTAGGACCTTAGTGAGTTGGAATTCCCAACCAGACATCATTAGGTCC
TTGTATTATCAGGGTGAGGGTATAGGGGTAGGAGAGCAGCAGAGGGTACTGATCTAGAGAGGTAGTGTGTCTGGGAG
CTCCAAGTCCCGATACAGAAAAGGGGCATGTCAGGACATTGGTGGCATGCTCTTGTCC**TCATTTTGCTCCAGCACAA**
GTGCCAGGATACTCAGCCACACCCATGGTGTCCTGTTCCCTGTTCCTATTTCCTTCTTCCAATTCTTCCCGGTATCT
CTACTTCCAGGACTCCTATCTAACACAAGAAGATTCTAGCTATGTCAAGCAGAAAAAGGAAGGCCACTGACACTGCT
GGCAGGCACTCAAGGATGGATCCAAATCTCTCCTCAGATGACAGCCAGAACCCAGGTGCAGTCGCAGCAGGTAAAAA

AACAGAAGACAACAACAAAACCAAAACAACATCATGAAAAAACCCACAAACAAACAAACCAGAGGGGCTGGCAGAGG
GCTGTGTCCACCCCCAGCAGACATCCTAAAATCCCTTGGCAGGTGGCACACTGAGTGGATGTGGTGGTGTGTGGACA
TTTCCTTCTTGGACTAGTGGGGGGAGGTTCAGATGCAATTAACTGGTACTCAAGCTTGTTTGTGGCTTTAGCTTTCC
CCACCTGAAATTCCCACCACACAAGAATCCTGCAAGGCTCATAAGACAGTTCTCTGCAGGCAATATGCAGTGTGCCT
CTCTGGGAACACATGGGTCACATGTGGAAGGACAGTGCATATGCCTCCAACATCTGGTCCAAGAATCACTCAAGCCT
AATGAATTTTGTGTGTGTCTGTACTTCTGTGTGATATTAATATTCTCTCTCTCTCTCTCTGTCTCTCCCCATGCTTT
TTGTGCCCACTGTCACTCACTCTCTGCCTGCACGTGCTTGTGTGTGCTTGCATGCATGCATGCGTAAGAATGGGAGT
GTGTGTGCATGACTTCTTTTAAAAGTGAAAGTTACTGTAGGTGGCTATTGGCTCTAGAAACCCCACCGTGAGTCAGT
TTTCAATGGGAGATGATGGGAGAAGCTCACTTTGTCTCAAAACATTTTTCATCTTTGTCTTTAAATACAGCTAACAG
AGAAGTCCTTGATGCTGGTAGGGAGGACATTATTTCCTCAGGAACAGAGAGGCAACAGGCCAGGTGATCACCTTTTT
TGTTTGTTTGTTTCTTTCTCTCTCTATTTCTCTCTATTTCTCTTTCTCTCTTTTTCTTTTTCTCTCTATTTCTCTTT
CTTTTCTTCCTTCCTTCTCTCTCCCTCCCTCCATCCCTCCCTCCCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTC
TCTCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCCTTTTGCCCACTCATAAACAATGTTTAGAAAAG
TCAGGCTGGAAAGGTGGCTCAGTGAAAGGGTGCTCTATGTATAAACATGAGAATCTGAGGTCATATTCAGCCCCAAC
AGAAAAGGCCACACGTGGCCACTTAACACCTTTGAAGCCCAGAACTGAGGGCTGAGGTGTGGCAGAGAGCTTCTTTC
TGCATGCCAACCTACCAAAAATAAACTGAGAGATAAACAAACAAACAAACAAACAAATGCAGTGGTGGGTAAGGTGT
GTGAATCACAGGAATGTCCTGGATGCATTGAGTTTAAAGGCCACAACACATTTGTCCTGCAGTCCATCAGACGTTCA
GTGTGTTACAAAATGAAAATCTCATGCTACCAGCAAACTACAAGTTCTGTCACAGAACTTGTCTGACAGGAATAAGG
TGGAAATTGGTAGAGGAGGACCCCCTCCACACATACACACATATACATTACTACCTCTGATGGCTGCACACTTATAC
ACATGCATTTCTGTTGCATACAGAGAAATCGAGAAGAATCAGTGTCTGCTCTTGAACTGTGAACTGTATTGAGGCTG
TGTGGGGCACACTCTACTTCCCTGCTCATGCAACAAAGTTGGCCTTTGGTATACATTGAATTAAAGAGTTTATGTGA
TTTCT**TTGCAGGAAGGAAAAACAGG**ATTTGGTCCAAGAATTCGAAGGTATGTTCTAGGAGACAATTCGAATTTGAGA
AAATTGTTTGAGCCAAAGAAGTG**CAAGGCAAGGGAGTTGTCAC**TTGGCAGAGAAGAGAAGTGGGCTTACTGCCTAGT
TCTCAAACACACAGTGTGGAAATAGTGTTTATTACTCACCAAGTCTCTTACTTCACCTTAATTGTCCATGTGAAGAT
CCTGGCACCTGCAGTGGGACGATCCGTGCCTACTCTCTGTGAGATATTTACAAATTGGATGAACGTATCGGTCTCCC
AGTTTTTGATTCACACAAGTGATGCATCTGTTTTTGATACCACATTTTACTTCTCTCATTTGTATTCTAGTCCTGTT
CTTCTATATTTGGTTGGGGGCGGCAATGCTTGAAGAGGGCCAATCCTGGTTAAATAGCTGCCCACAATTTACAATGC
AAAGAGATAGCACAGGAAGTGGAAAGCTGAAGTCAGTGCTGTTAGCCTGTGGCTCTAGTATGATCAGTTGAGTCCTG
CATTAAATTGCTGTCAGAGGTTAGTCACATGAACTGTTAAACAAACTTATCAACCCGAGAAACAGCTGGGAGGTCTG
CAGGCTTTTCTGATGTTGAGTTGCATCTAGAAACTATCCTGAACCTCAGGATAGGGCTGTAGTCATGTGGAGGCCAT
TGCCTCATATTGTTCAGGGCTCTGGGTTCCATCCCCAATACCACAGAGTTTTCCATGAGTCAAGTTTTCCCTTTCAA
TTATTCAGGATGTAGCCGGGCGTGGTGGCCCAAGCCTTTAATCATAGCACTCAGGAGGGAGAGACAGATGGATTTCT
GAGTTCAACACCAGCACGGTCTACAGAGTGAGTTCCAGGACAGCCAGGGCTACACAGAGAAACCCTGTCTCAACCCC
CCCCCCCAAAATTTTAGGATGTCTACCTTATCAAATTTGTCAAATAGCCTCAAGTGCTACAATGGTTACTGTATTTT
TCATCTCACTGTCCAGTGATCTCTGAGCCGTTGGAGATTGTTGTAAATGGACATTTGTACTTTGCATCAAGCAGCCC
AACAGTTCCTCCTAGCATGTTCTAGGACTGGAACACTCTTAGTACATTGATTACATTCATACTTTATAAATGACATC
TAGTAAAGAAATTACACAGTGACTTGATACCCTTTTTGTAAGTCTTGTTAAGGCCCCTCTGTTATGTTTTGTATGTA
ATATCTAGACACTGAAGTCCAACTGTCCAATTTCATGGTCAGGTCAGAGCCACTGAAATTCTTAACGGAAAGAAAGG
AGGCAAATAACCATTGCATTGTCAGATCATCTCCACCTATTGATAGCAAGTGTCTTTCCAGTACCCTTGAGGGACCT
TCACAGTAATGCCTACAAGATTATGTCCCCTTACAGAGCCACCTAACAAAGTTCTTCAGGAAAACAGAGAAAGATTC
TCAAGGATCATGACCTCATCTTTCAGTGCCATGGAGGTCAAAATTAAGGATGTTCTGAAAACCCACTGTGAAGAGAG
GTAGGGTTGTTGCTTCTTGTCACTTTGATCCCAGAGCTGTTGACAGTTGGTTTGTGAGCACAATGAAGGGGGAATAG
CCAGTGGCCCCACTGGCTAGAAGCTTTGTTGGCAAGTGAACTTGCCTTGATTTTTCTGTAGGACAGACTTTGTTTTG
TCTTATACCCTCCCGTTGTACTGAAGTCAGTCAACTGCGTTTCCTAAGGATGAACTAAGATGTTTTGACACAGGGCA
GTTGAATCAGTGTAACAGTAAGTAAGGTGTGCGCATTGTGTGAAAGTTCTGGGTACATTGAGCTTGAAATCCTCAAA
TCATCCATCCTGGAGTCCTTCATCTCTTCAGTGTGAAATGAAATGAAAATCTCTAGCCTCCAATGTTGTACAGGTTT
CCGGAGG**AGCCCAGTTCACCATACAGG**TATGCTGAAATCCTCACAGGTACAGTGATGAAGAGTGGGTAGGCGTCCAC
AGAATTCTCTGATGGTTGTGGGTGTGCATCCTCATTGGACCTATCTTTATAATGGGAAA**ATTAAGCAGGGGTTGTGC**
TCGTCAGAGAGACTAATACTTGTGGAATGCTACTAAGACAATCCTGAGTTTTCTAGCACTAAGAGCTGTGAAGTGAC
TCCTGTCTCGGCCAAGGAGGAACACATGCTATTAGCCTGGTAGGAACTAATCCACCTCATTTTCCATTTCCCTGTCA
GAATATTTTAGATCTCCTTTCCCAGTCTTGCATTTAGTGAAGGAAAGTCAGTCGTGCTGTGTTTGAATGGGCTTTCT
TCATGTCGAGAAGGAATATCACAGTGTTCACGTTGTCTTTTACAGGCAGAAACTTTGTCAGGACTATTCTCTTCAGT
TTACAAATTTGAATAGGAAGTTAACTTCGGATGCATACAAACTCAAGAAACATGCCGAAACACTCTCTGTGAGTATC
AGGCTTAGATTTAAAGGAATCCATTTCCTCAGCATCAGAGGTAGCAACACAGGCTTCCTGCAGCAGGTCTGGAGGGA
AGGGACAGGACATGAGCTCTGGGAGACAAGGAGGGTGGTTGATGTGGAGGGGAGAGACTGGTCATCCTCCCTGACTT
GGTTTTTTATCCCCCTTCTGTGGGAAGATGACAAGGGTAATGATGGGTTAGGAGTCTGCATGGTACAGGCCTGGGGG
GATGAGCAGGTACTGTCTGCAGAGTGCAGGAGGAGGTATGGGATTGTATCCTGAGGGGGGACCCTTCAGGATCGACA

CAGTGCAGATGCCAAAGAGTAGGAGGGCCCCACGTAGCCTTCTGACACAGTGATCAGGGATGCGGAGTGAGAGCAGT
TCATGGGGCTGTGGAGAAGCAGGTGGAGACTTAGCTGCCAGTGTGGTTGGAGTTCAGACCGGCCCTGTCCCAGAGGA
GAAGAGGGAAAGGTTCCCATGCAGTGATGGGTGGACCTGAGATGGTGACTGAGTCAACTTTAGCCTGGAGATGAGAA
TGGGCTTTGGGGGACTGGTAACTGGACTCAGTGCTGTCTCTGGGGGCTCTTTCTTCCTGGTGGAGGATGGAGCTGCT
CCCTGGGGATTCTGACTTCCAAAGGTGGGGAAGCAGAAGGGGAACTACTGAGGGGTATGGAGATACTGTCCAGCCCA
GCTCAGGTCAGGAGCGTGGACTCTGGAATCCCACCAAGTGGATCACAGTCACCGTCTTTCTTTTGCTAGTTATGCTA
ACTCAGAAAAATAATTTACATCTGTGCCTTAGTTTCTTGGGAGTGGAAGGAGAGAAGAAGTCCATACGTATCTGAAA
GAATATTGTCAAAGATTTGTTTTAGATGTGTATCCTTATATTGGTACAGGGCAGTATCTGAAGTCAGTCTTAAAGAA
TATAACCCTGTTAGTGTGGCACTGTGTGCTCTACTCTTACTGGAAAGGCCAAAGTGGAGTGGGTCCCTATGTGGAGG
AGTGACAAATGATGGGGATAACTCAAGGTAGGTTCTAAAAATACAGAGAGTCCAGCAGATTTGAAGGCCTCTAGTAG
GTAGGTTTTCAAGACTACAGGAACTACTAGCAGTGTTTTGAGTACTGTTAAATAGGAACCCTCCCCATTCAATGAGT
GAGCTTCACATAATGAAGGTTATTTTTCTACTATTAGGTTACTCCTGTAAATGTGGTCACTGTTTTAATTGAAAACA
ATTGTAAATGTTTTTCACTCCATCTCTTTTGCATAAATAAAATAAAAAATTGCCTTTTATGAGTTTCCTGTCCCCCC
CCCCTCACTTGTTCTTACATCTTACACTGCCTTTGTTTGTAGAATATGTTTATGGAGCAACAGAAGTTTATTCATGA
AAGTCTCACTCTTCAGAAGAACAGAATGAAGGAATTTAAATCACTGTGTGAAAAATACTTGGAGGTGTGTTGAACAA
CCTCTACAGGGGACACTGAATATCCTGGTATTCTTTCTGTGCTTTGGATAACAGCTTATGAAAGTTACCTGCAAACA
CCTCCCACCATTTCCTCATGCTCCATGTAATCCTAAGTACAAAGTCAGTTTAATTTTAGCATTCACTCATGGTTAGG
AGAATCATTCAAAGCAGGACTTGCTGCCTTGGTCATTTCATCTGGAGCTGTTCATATTGTACAGGGCACTAATGCTC
TTCAGCTCATCCCATCTTGACTCACAGTACATTTTTTCATTGGGTAAGTATGTGTACACATGCACACATGCGTACTT
GTCTGTGCATGGGTGTGTAGATGTCGGAAGACAACTCTTACAGACACACTTCATCATATCCTTTTTTTCTTACCTCT
AGCTTTTTGTTTGTTTGGTTTTAATTTTGACACAGAGTCTCACTGTGTAGCTCTGGCTGTCCTGAACTCACTATGTC
GACCATTCTGGACTCAAACTCACAGAGATCTACCTGCCTGCCTCTGCTTCCTGAGTGCTGGGATTACAGGTGTGCAC
CACTGTGCGGGACTCCCATGATACCCTTTTGAAACAGGATGACTCCCTGGACCTGGAACTATGGCTAGGCTGGCTGG
CTGCTCAGCAGTTCCTTTGCTGCCACAGGTCGATGTGTCACCATGCCTGACTTTTAGTATTTATGCTGAAGGGCTTA
CGTTGGCCTTTTATGCTTATGGAGCAAGCTCCTCCACTACTGCTGCATTCTTACCCCCCAAGAATTACTTTTACACA
CTGCAGCTGTCGTCCTAGACCCTCTCGTTCCATCTTTAAGCAGACTAGTTACAAGAATATTTGCACTTACTTCCCTT
GCTCAGTCCATTTACATTTCCTATCATTGTAGGACTGAAACATGGCTTCATATGGCCAATTAGAACCATAAGACAGC
TGCCAGGTATAGCCCTCAGTAAGAAAGTAAGTGGAAGACAGAGAGATTGTGAATTGAGAGTTGTCCTCCATAATAGT
GTCTGTAAGTTCTTGCTAGGCATGTTGATTGTCCTATGCTGTCTTCAGGTTTGTAGCTCTCTACAAGACAATTAAAT
ACTGGACGCTCTACTTTTGTTTCAGACTCAATGTGAA**TTGGCTTGTTCAAGTTGCAG**TTAAAATTTTGTTCATCCAG
AGTAATTCAGATGACCCAAAACAACCAAGTTGACAGAATTCATCCTTGCTTTCTTAATGCCTGTCTCCTACAAAAAA
ACCCTACCCCCGCACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACGGGGCAGGG**GTTAG**
GGGCCAGTGTGTGTTCCTGATGAAAGATCATTCTTTAAGTACTCAGTAGCCTAGAAACTTGCTCAGCCATTCCAATG
TTAAGATCTCTGTGTTGTTCCAGTTTCTGTGGTCACTGAACAAGATTATGGAAGAGGTCCTGCCAAGTTTTATAATG
ACTCCACAGTGTTTAGCAAGACTGTAGATGTAGTGTGTGCTGTTTCACTCAGTGCCAAAGGGAATGTGGAGGTGGAA
ACCTTGAGTTTTAAGTTCCTGGTGCACAGTGGTCATAGTACGGGACAAGCTATTCCATCTTGCTAAGAAGTGGGCGA
GTGGCAATGAATGGGCTCATGCTTCATGCTCTGTTTGCATTTGGCTTCATGGGCTTTCATTTGTGCTGTGCATGGAG
AGTTATCCCAGACTCTCAGACATGGTAGGCAAGCACCCTTCCCCTGGGCCACCCCTATAGCTGATCCCTCACAATTT
GACCATTCAGCAGTGGGTTGAGAAGTGGCAAAGTCGGTGTCACTTTGGTGGATGTTTTGATCTTCTTCTTGTACAGG
GTGCGTGCTCGCTGCTTTTGGAGTAATTTTTGGAGTAATTAAACCATGTTTTTTTTACCACACTTGCTCCCTTTGAA
ATGTTGATGTTCCCATTTGAAAGACAAGATGACCAGGGCTGATTTCCTCACATGTTTTTTGCTTCAGAAACTGGAGG
TACTGAGGGATTCTCGGGGAAATTCCATCGCTGAAGAGCTGAGACGTCTCATAGCCACCTTGGAAATCAAACTTCTG
ATGCTGCATGTAAGTACCTAGCCTTTCCCATAGAAAATGGGGATAAGGCCAACCTTGAGTGCTATTTATTTTATGCC
CATTTAATATTTTCATTACTTAACAGAACCAGCAAAAGACTGCTGCTCCTCCACAGTCTCTCCTGGATGTGTTATTC
TCATAAAACTTTGGAGCACAAAAGCCGATATGAGGGGAGCAGTATAATCATCTGGGTGAAGCTGCTAGTGCTGATGA
TTGCAGCTTCGGTGCCTGGATCTTCTCCTGTCTTCCTCCTGTTCCTGAAGATGCCCCGTTCTGTTAAATCCAAATAA
ACCCTGGGATTTGTGTGACAGCAGTGCAAATATGTCTGTAGGGAACTCCAGCCCTCGGGTAGAGATGGCAGGCTTTC
CTTGGTAGCATCGCAATTACTGGTTAGCACACATTGGTGTTGATACAGTAGATTATCATACCTGAACATCATCTCCC
CAAATTGGAAGGCACACAAGTGGGCTATTGTTGTCGTGTCCAGATAAATAATTAGAGAAGATACAGGCTTTCTTTGA
CATTCTGCTCTGAAATGTGTTCAGTGATGGAGCCTATGAAA**CTGACCTTCAGCCTCCAAAG**GCAGTGTGCTCTTAAT
AGTTAAAGTAGTAGAATTCATCAAACCTCCCGAGTAAGAAAGATGTGAGTAACGTGGCTTGAATCAGTGGAAATGCC
ATTTGCAGACAACATTTACAAGGTCTTTCAATTAAATGAATCTGTCACACTAAGAGGTGATTCTGAACATGAATACA
GCTTAATTCACCATGGTTTATTAATCCCTCACTTGTATGGATAGGAGTAACTTAGAATGTTTTGTTTGTGGAATTTA
TCGCACATGTAGTATGGGATCATCAGGGATATTATATATGAGGTCTCGTGTACAGTTAAAGCTCGCGTCCTGTGATT
TCCTGTATGGATACCCTGGGAAATCCTTACTTGGGAACGTTTGATTTAATAAACATGAAATAGTTTCAGACTTGTTG
GCATGAGTCATTTAATTGCATTCTCTTTCTGACATCCCTGACGTTTCACCCCTTCTCAGACAGACATGGCCCCATGC
CATGTGGAGGTGTGCTGTCACCTGTGTTCCTGCATGGGGCCAGTGTAGTCCAGGCCACTTGTCTCATTGTGTCCCGT

TGTGTAGAAGCATTGAAAACAGGCTGTGTTCAGCCTCCCCCATTTTCTCCAAGTTGAATTGCATTCCTCCAGACTGA
ATGAGGGCTTCCCTGCTGGGATCATTTTACTGTCTAAAGTTTGCATTATCTAGAAGAACTTCCCCCTTTGTGCAAAC
TACCTGT**AACCCTGTGGGCCTTTAGTA**TTATTTCTTTAATAATCATTAATCATAATAAAATGTTGTATTATTAGCAG
CTTTGAATGTCTCCTACTTGAAATATTCTGCCTATCTTTACCATATATGGATACTGTCCATAGTGTTTCTCCAGTGT
TGAATTCTATTTTATTTCTCTGTCTCTCTGTCTCTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
GTGTGTGTGTATAAATGTTCGTTTATGTGAGTTTTGATTTTACCTGATTACTGGGATTTGTGACAGCTAACCAGGGG
GTGATGGCTGGGAATAGGAGCAGAATCCACAGCAAGAGATGTTGGGTGGTCGATGCAGGGGGACCATGGAGACCTTG
GGGGCAGGGAAACGGTGTTACTAGAACACAACCAGCGCTAGTGTGGGTTGCTTTCACGCAACAGAACAACAACAACA
ACAACAACAACAACTAAGACTTCTCACCTCCCTGAGCACAGGACTTGGCTCTAACATTAAATTTCTTGGTGATCTTT
TTCTCCTCAAACTGTACATTTTATGTCTCTT

Xlr3b_4cF1	GTGGGGCACACTCTACTTCC
Xlr3b_4cR1	CTGCTGCTCTCCTACCCCTA
Xlr3b_4cF2	AGCCCAGTTCACCATACAGG
Xlr3b_4cR2	GTGACAACTCCCTTGCCTTG
Xlr3b_4cF3	TTGGCTTGTTCAAGTTGCAG
Xlr3b_4cR3	GAGCACAACCCCTGCTTAAT
Xlr3b_4cF4	CTGACCTTCAGCCTCCAAAG
Xlr3b_4cR4	AACACACACTGGCCCCTAAC
Xlr3b_4cF5	TTAGCAGCTTTGAATGTCTCC
Xlr3b_4cR5	TGGCATTTCCACTGATTCAA