Genetic modifiers of cerebellar hypoplasia in Zfp423-deficient mice
Genetic Modifiers of Cerebellar Hypoplasia in Zfp423-Deficient Mice

A thesis submitted in partial satisfaction of the
requirements for the degree of Master of Science
in
Biology

by
Edward Chen

Committee in Charge:
Professor Bruce A. Hamilton, Chair
Professor Yishi Jin
Professor Immo E. Scheffler

2009
The thesis of Edward Chen is approved and it is acceptable in
quality and form for publication on microfilm and electronically:

____________________________________________________________________
____________________________________________________________________
____________________________________________________________________

Chair

University of California, San Diego

2009
DEDICATION

To my mother, Jennifer Cheng, for her love and unwavering support especially through the toughest of times.
# TABLE OF CONTENTS

Signature Page .......................................................... iii  
Dedication ..................................................................... iv  
Table of Contents ........................................................... v  
List of Figures .............................................................. vi  
Abstract of the Thesis .................................................... vii  
Introduction .................................................................... 1  
Materials and Methods .................................................. 4  
Results ........................................................................... 8  
Discussion ...................................................................... 23  
References ................................................................. 25
LIST OF FIGURES

Figure 1: Arranged model samples of *nur12* mutant phenotypes ..................11
Figure 2: 129xBALB *nur12* mutants genome scan, half plates 1 and 2 ..............12
Figure 3: Genetic map of 129xBALB, half plates 1 and 2 .............................13
Figure 4: B6xBALB *nur12* mutants genome scan, half plate 1 .......................14
Figure 5: 129xBALB *nur12* mutants genome scan, half plate 3 .....................15
Figure 6: 129xBALB *nur12* mutants genome scan, half plate 1, 2, and 3 ........16
Figure 7: Genetic map of 129xBALB, half plates 1, 2, and 3 ..........................17
Figure 8: 129xBALB Permutation map, half plates 1, 2, and 3 .......................18
Figure 9: B6xBALB *nur12* mutants genome scan, half plate 1, 2 and 3 ..........19
Figure 10: 129xBALB Permutation map, half plates 1, 2, 3, 4, and 5 .............20
Figure 11: Genetic map of 129xBALB, half plates 1, 2, 3, 4, and 5 ...............21
Figure 12: 129xBALB Permutation map, half plates 1, 2, 3, 4, and 5 .............22
ABSTRACT OF THE THESIS

Mapping a Modifier of Zfp423 in Mouse with Varying Degrees of Cerebellar Hypoplasia

by

Edward Chen

Master of Science in Biology

University of California, San Diego, 2009

Professor Bruce A. Hamilton, Chair

The cerebellum plays a critical role in the neural functions of both humans and mice, sustaining a significant control over motor, coordination, and proprioception. Despite its small size, the cerebellum comprises over 50% of all neurons in the brain. Therefore, malformation or lesion of the cerebellum, although not fatal, may cause a severe handicap to both physical and mental functions. This thesis describes the design, implementation, and results of two genome-wide linkage studies to detect genetic modifiers of cerebellum defects in mice that lack the zinc finger transcription factor Zfp423. We find significant linkage peaks at chromosome 3 (LOD = 4.74) and 17 (LOD = 4.67), along with a suggestive linkage on chromosome 15 (LOD = 3.23) in
a scan of 222 BALBx129 F2 Zfp323$^{nur12}$ mutant mice, showing a clear genetic basis for phenotypic differences among mutant mice. In a second study of 132 BALBxB6 F2 Zfp323$^{nur12}$ mutants, chromosome 4 (LOD = 3.69) shows nominally significant results that may embed one or more modifiers. This thesis also includes the breeding of congenic mice toward an attempt to narrow and isolate the possible candidate genes in the 129 mouse strain.
INTRODUCTION

Laboratory mice are well established model organisms to study the mammalian genetic complement, not only because of their relatively short generation time but also because the sequencing of the mouse genome is complete and found to be very similar to that of humans. The high degree of homology between the two species and other reasons for cross-species comparisons have been reviewed elsewhere. (Hamilton, 2001). This thesis focuses on using this genetic model to test the genetic basis and to identify specific gene loci that may affect the development of cerebellum in a specific brain malformation model, Zfp323\textsuperscript{nur12}.

It is well understood that specific genotypes of a particular animal would produce certain phenotypes, such as traits or appearances. This thesis encompasses the work in finding new genetic differences among different strains of mice based on the \textit{nur12} mutation. Previously, mutation of the transcription factor \textit{Zfp423} on chromosome 8 of mice results in loss of the corpus callosum, reduction of hippocampus, and a malformation of the cerebellum due to the reduced proliferation of neural precursors (Alcaraz et al., 2006). My research is directed toward finding genes that might modify this mutation, thus providing insight on genetic networks important for cerebellar development. To test for possible linkage of a modifier gene to a particular locus in the genome, I conducted genome-wide scans using > 100 well-spaced markers on several hundred animal subjects, comprising more than 30,000 single-point genotypes in all. I used well-documented, non-parametric multi-point linkage methods in the R/QTL package to assess the statistical significance of the results.
My work builds on previous studies by others. The most relevant and important study on this topic was done by Wendy Alcaraz, who determined the physiological effects of the loss of transcription factor Zfp423, including ataxia and brain malformations (Alcaraz, 2006). Tsai and Reed’s earlier work (1997, 1998) provides important biochemical background to this research. They showed that Ebf1 (Olf-1) transcription factors, responsible for the olfactory receptor neuron cell fate determination, becomes reduced in transcriptional activity in the presence of Zfp423 (Roaz) protein (Tsai, 1997). Thus, Ebf genes could be candidate modifiers for the transcription factor Zfp423. Furthermore, the bone morphogenetic proteins (BMPs) may also play a role in the research. BMPs affect many aspects of mice growth, from skeletal to kidney, lung and gut to teeth developments through the Smad transcription factors. Zfp423 associates with Smads in response to BMP2 and keeps neuronal precursor cells in proliferation (Hata et al., 2000), and the modifiers we look for may affect this signaling pathway.

The nur12 mouse, with an ataxic gait and cerebellar hypoplasia, was found to have a nonsense mutation for the 30 zinc-finger protein Zfp423. This nonsense mutation leads to an early termination of protein translation, causing an effect similar to that of a null allele of the Zfp423 gene. To carry out my research, I am using mice with this mutation in the Zfp423 gene, which is a gene involved in maintaining cells in a proliferative state (Alcaraz, 2006).

However, among the mutants of nur12/nur12, there is a variety of phenotypes observed, ranging from a complete lack of cerebellum to a cerebellum with limited, joined hemispheres (Figure 1). Therefore, we proposed that these various strains of
mice have genetic variations that led to their alternate phenotypes (Alcaraz, 2006). Since then, efforts have been devoted to identifying the location of the modifier genes.

My thesis work has tested the hypothesis that modifier genes present in common, inbred strains alter the phenotypic expression of $Zfp323^{nur12}$ by implementing complete genome scans of intercross progeny. For the scan, we use three separate and commonly known strains of mice to create two intercrosses: 129S1/SvImJ (129) x BALB/cHAD-\textit{nur12} (BALB) and C57BL/6J (B6) x BALB/cHSD-\textit{nur12}. With the availability of the informative simple sequence length polymorphisms (SSLPs) information for the mice (Dietrich et al., 1996), the genome scan can be accomplished through the use of numerous SSLP primers to determine a rough location of each recombination event in each mouse. In order to find out the chromosome(s) of interest, a wide array of primers is selected such that each marker would be close enough to determine recombination event occurrences.

To make the data collected meaningful, the R/QTL program is used to statistically calculate the logarithm of odds (LOD) score of each locus based on the genotypes gathered from PCR and the phenotypes gathered from empirical dissection. A LOD score (logarithm of the odds ratio) statistic is used to indicate the ratio of the probability of a particular locus having true linkage to phenotype to the probability of that locus having a random linkage to phenotype. Our genome scans, based on the selected SSLP markers, identify three loci of interest that meet the empirical thresholds for genome-wide significance (Broman 2003).
MATERIALS AND METHODS

This research puts the primary focus on polymerase chain reaction (PCR) of the mouse genome. The purpose of the first stage is to create genome scan of the two F2 mouse intercross strains of 129xBALB and of B6xBALB. Using two pure strains of mice, one of C57BL/6J (Black 6) and one of 129/SVImJ (129), and crossing each of them with a BALB/cHAD-nur12 (BALB) mouse, we produced two separate heterozygous F1 progeny. By crossing each of the progeny F1 together, we produced the two F2 intercrosses that are used for the genome scan. Upon confirmation of nur12 mutants in the F2 generation through the use of PCR on nur12 markers, we collected the DNA using tail samples during P7-P10. The DNA samples were prepared with 600 ul of 50nM NaOH heated at 80°C overnight and with 50 ul 8.0 pH Tris inserted after the heating and proper mixing. The samples were arrayed on 96 well PCR plates to facilitate greater throughput.

For the completion of this project, there are two types of PCR procedures for discerning the genotypes of the mice. First is the Nur12 process which allows for differentiation of the nur12 mutants from the heterozygotes, and the second is the general PCR which distinguishes the strain genotype of the nur12 mutants in the genome scan. The main difference is the annealing temperature of the procedures, especially for Nur12 as there is only a slight difference in the primer sequences. For both reactions, the reagents required 100 uM of both ends of primers (DNA oligonucleotides), 10 mM of each of the four dNTP (deoxynucleoside triphosphates), 10 mg/ml BSA (bovine serum albumin), and PCR buffer consisting of MgCl2.
Microsatellite markers were amplified by PCR with primers from public databases (UCSC genome). The Nur12 PCR is conducted with 35 repeated cycles of denaturation, annealing, elongation, finishing with a 6 minute period of incubation. Its annealing temperature is at 66°C, higher than general PCR reactions as a high specificity is required for the annealing process due to the minute primer sequence differences. The general genome scan PCR was set at 40 cycles with annealing temperature of 55°C. PCR products were electrophoresed on a 3% superfine agarose gel. Suspicious genotypes such as those that created double recombinants or fake bands were reassayed until clear.

A list of 88 specific microsatellite markers that spread out throughout the chromosomes was initially complied such that we could detect the SSLP (simple sequence length polymorphism) at various loci. The markers were specifically chosen such that they were spread out ≤ 20cM apart from each other. The selection of primers is especially difficult for the BALB-nur12/+ x 129 intercross because there exists no comprehensive data for microsatellite alleles in 129. Each primer has to be tested ahead of time to ensure the bands will separate upon PCR and would be effective in discerning the genotypes. The estimated 20 cM distance between markers ensures that the positions of a gene would be no more than 10cM from at least one marker to provide high information content across the genome. As the experiment progressed, further markers were selected to narrow the distances between markers as well as to focus in on the chromosomes of particular interests. The current number of primers employed stands at 105.
The raw data is first obtained through the gel pictures taken under UV light at the Variable Intensity Transilluminator. These pictures are then analyzed to determine the genotypes of each mouse at the specific locus, and the results are tabulated. When a collection of them has been verified and tabulated, the data is processed to calculate the linkage probability using programs such as R/QTL in R.

After a statistical comparison between the genotypes obtained through PCR and the phenotypes obtained through dissections of nur12 mutants, we calculate a LOD score value for each of the markers as a statistic of linkage from the QTL program. This QTL program does quantitative trait loci analysis as well as genome wide permutation tests to calculate empirical p values for conventional significance thresholds. In our experiments, a LOD score of ~3.3 or greater and a p value less than 0.05 signifies significant linkage.

As a second step, chromosomes with a high LOD score values are especially targeted for review, and additional SSLP markers in that area are used to confirm and refine linkage in these areas of interest. In these tests, the entire linkage region would be covered and examined for genotypic changes in relationship to phenotypes. From this result, we can tell which segment(s) of the gene would likely alter the phenotypic appearance and trait of the mice in the cerebellum. These gathered data will facilitate genetic mapping studies and identification of novel candidate genes for the trait loci mapped to this region.

In addition to accomplishing the genome scans, the development of congenic strains of mice is also attempted. After confirming with statistical analysis that specific chromosomes in the F2 intercross strains of mice have significant
probabilities of containing modifier genes, the creation of congenic mice specifically with the effective strain isolated while with the BALB strain in other areas of chromosome is performed to validate the effects of the modifier genes. It usually takes ten generations of breeding to create conventional congenic mice, all the while taking care to preserve the areas of interest and constantly testing of the genotypes of mice of each generation to perform the breeding. Congenic mice are weaned at the standard period of 3 weeks after birth, while tails are collected at P10-14 for genotype testing and selection. Due to time limitations, mice after four to five generations of breeding (N4 and N5) are introduced to nur12 and its mutants are dissected for phenotypical analysis.

The dissection protocols are the same as the phenotypical analysis of the genomic scan mice. Each nur12 mutant mice of the N4 and N5 congenic are selected along with one of their wildtype littermates as a control. The dissection is usually performed at P14 with 1-2 days variance. The mice are first measured in weight to establish a standard, and then they are rendered unconscious with avertin with the ratio of 100ul per 5g mass. Several minutes after injection, they are tested for response to pain. When each is completely unresponsive the procedure can begin. The dissection consists of opening up the rib cage and cutting open its main artery to allow the circulation of the injected 10ml PBS first then 10ml of 4% PFA (paraformaldehyde). After the flow of the fluids is complete the cerebellum is collected for observation. They are also stored at 4% PFA to maintain its form and prevent deterioration. The phenotypes of each mouse are collected along with their genome scan data from the list of primers for further statistical analysis.
RESULTS

To test the hypothesis that there exist modifier genes among the three separate strains of mice, we created the two F2 intercrosses and collected a substantial number of mutants and compared their genotypes to phenotypes in a R/QTL statistical program. After much work in PCR and gel electrophoresis, we have obtained satisfactory data for several sample populations of \textit{nur12/nur12} mutant mice from the F2 generations of BALB-\textit{nur12/+} x 129 and BALB-\textit{nur12/+} x B6. Before each plate is compiled, each mouse undergoes a phenotypical analysis that judges the shape and size of its cerebellum. Figure 1 shows the standards by which the cerebellums are judged. The leftmost picture shows a wildtype phenotype in which a whole cerebellum is intact. The phenotype categories run from 0 to 3, with 0 being the phenotype with the least amount of cerebellum development with virtually none in appearance and 3 being the phenotype closest to that of a wildtype phenotype, including a joined hemisphere. Once the mice cerebellums are categorized their DNA samples are collected in the methods mentioned previously and the genomic scans are performed. Due to the large number of samples gathered and the relatively long breeding periods of mice required for collection, the results are compiled and organized in three separate stages in time.

Stage 1

With the 1\textsuperscript{st} and 2\textsuperscript{nd} half plates of DNA of the first sample population of 88 \textit{nur12} mutants of BALBx129 collected, the genome scan is performed and the data calculated in R/QTL. This initial scan shows significant linkage peaks at chromosomes 3 and 17 (Figure 2). At the same time, a genetic map is produced with
the data gathered using the same program R/QTL (Figure 3). This allows a direct
observation of the genetic distances between each marker based on the recombination
events that have taken place for the 88 mice collected. At this point, chromosome X
was not taken into account, and there still exist several rather large experimental
genetic distances between the markers chosen in chromosomes 2, 10, and 13. The first
sample population of BALBxB6 mice is also scanned with the same markers, and the
resulting scan showed a singular peak at chromosome 4 (Figure 4).

Stage 2

The 3rd half plate of BALBx129 is collected and genome scanned. Based on
that data alone, a significant peak of genotype effect is located at chromosome 15, but
both peaks discovered at chromosomes 3 and 17 in half plates 1 and 2 in stage 1 are
not repeated (Figure 5). At this point, the gender of the mice is determined as a
separate category of analysis and allows the evaluation of chromosome X. Compiled
overall data from the three half plates of 133 mice now show significant peaks inside
chromosome 15 and 17. Only a suggestive importance is presented in the loci at
chromosome 3 (Figure 6). The compiled data is also used in conjunction with the
R/QTL program to create a new genetic map (Figure 7) and a 1000 sample
permutation map (Figure 8) to determine our confidence level in terms of p value for
each loci. Furthermore, the compilation of half plates 1, 2, and 3 for the BALBxB6 F2
intercross was completed, showing a significant peak at chromosome 11 with a LOD
value of 3.69 (Figure 9). At the same time, congenic mice created based on the result
at the end of stage 1 are progressively breeding at generations N2 and N3, and all their
BALBx129 heterozygous chromosomes are constantly purged to BALB homozygous
genotypes through each generation except their chromosome 3, which is left intact at the BALBx129 heterozygous state.

Stage 3

This is the most recent stage of the project. Both 4\textsuperscript{th} and 5\textsuperscript{th} half plates of the BALBx129 \textit{nur12} mutants have been collected and have been genome scanned. The resulting mice sample size for this F2 intercross stands at 222. Despite a different founder mouse for the F2 intercross of half plates 4 and 5, the overall compiled data still shows significant peaks at chromosome 3, 15 and 17 with satisfactory LOD scores for each (Figure 10). The genetic map from the same raw data shows acceptable distances between each marker (Figure 11) and the permutation test presented the confidence values of the result at each peak of the three chromosomes (Figure 12). Several pairs of the \textit{nur12} mutants of the BALBx129 chromosome 3 congenic mice of N3 and N4 were also dissected and their phenotypes observed, but the range of phenotypes observed varied from 1 to 3.
Figure 1. An arranged model samples of *nur12* mutant phenotypes. A sample wild type cerebellum is presented on the far left (+/+). From left to right, the cerebellum phenotypes range from categories of virtually no cerebellum (0) to a limited, joined hemispheres (3).
Figure 2. 129xBALB nur12 mutant genome scan, half plates 1 and 2. 88 sample mice gathered in this genome scan with the use of 97 SSLP markers. Significant peaks noticed at chromosome 3 (LOD = 4.27) and 17 (LOD = 6.78).
Figure 3. Genetic map created based on the genotype data collected from half plates 1 and 2 of the 129xBALB nur12 mutants. Several gaps of over 20cM are still noticeable due to lack of SSLP markers at specific positions.
Figure 4. B6xBALB \textit{nur12} mutant genome scan, half plate 1. 45 sample mice gathered in this scan with the use of 88 SSLP markers. No significant peaks are found in this scan.
Figure 5. 129xBALB *nur12* mutant genome scan, half plates 3. This scan contains 45 sample mice bred from the same founder of half plates 1 and 2. It shows a significant peak at chromosome 15 with a LOD score of 3.4.
Figure 6. 129xBALB nur12 mutant genome scan, representing the combined data of half plate samples 1, 2, and 3, with a total of 133 mice. This scan shows two significant peaks in chromosomes 15 (LOD = 4.72) and 17 (LOD = 3.87), with a strong suggestive peak at chromosome 3 (LOD = 3.22) of this intercross strain.
Figure 7. Genetic map complied from data gathered from half plates 1, 2, and 3 of the 129xBALB intercross. Most of the large gaps of 20cM or greater from the previous genetic map has been filled with the new markers to evaluate every area of the genome.
Figure 8. Permutation map of half plates 1, 2 and 3 of the 129xBALB *nur12* mutants. With the genotype raw data collected, the R/QTL program calculates a permutation simulation in which 1000 genome wide samples are randomly gathered. The LOD score of each simulated sample then is tabulated and averaged to find out the likelihood of a high LOD being achieved simply by chance. Above are the three loci with the highest LOD score from the three half plates, with their p-value confidence calculated from this simulation and their positions listed from the R/QTL program. The genetic mapping distance of the highest LOD score is also extrapolated with this calculation.

<table>
<thead>
<tr>
<th>chromosome</th>
<th>position</th>
<th>LOD</th>
<th>pval</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>34.2</td>
<td>3.22</td>
<td>0.070</td>
</tr>
<tr>
<td>15</td>
<td>28.0</td>
<td>4.72</td>
<td>0.000</td>
</tr>
<tr>
<td>17</td>
<td>33.7</td>
<td>3.87</td>
<td>0.015</td>
</tr>
</tbody>
</table>
Figure 9. B6xBALB *nur12* mutant genome scan, using the compiled genotype data from half plates 1, 2 and 3, which has a total of 134 sample mice. This scan shows one significant peak at chromosome 4 (LOD = 3.69).
Figure 10. 129xBALB *nur12* mutant genome scan, representing the entire collected data from half plate samples 1, 2, and 3, 4 and 5 with a total of 222 mice. This scan utilizes a sum of 105 markers across the genome. Two significant peaks in chromosomes 3 (LOD = 4.74) and 17 (LOD = 4.67) are observed, along with a strong suggestive peak at chromosome 15 (LOD = 3.23) of this intercross strain.
Figure 11. Genetic map complied from data gathered from half plates 1, 2, 3, 4 and 5 of the 129xBALB intercross. With the total of 105 markers spanning across the genome, all gaps of 20cM or greater of theoretical genetic distances from the previous genetic map has been filled. Any perceived larger gaps are the result of unexpected recombination events between markers in this sample size of 222 nur12 mutant mice.
chromosome position  LOD    pval  
3         26.2      4.74    0.002  
15        30.9      3.23    0.072  
17        27.7      4.67    0.003  

Figure 12. Permutation map of half plates 1, 2, 3, 4 and 5 of the 129xBALB nur12 mutants. Above are the three loci with the highest LOD scores from the five half plates of 222 sample mice, with their p-value confidence calculated from this simulation and their positions listed from the R/QTL program. The low p values for chromosomes 3 and 17 from this simulation shows strong evidence of true linkage of genotype to phenotypic changes at those genetic distances.
DISCUSSION

At the end of stage 1, based on the data gathered and statistically analyzed, we noticed significant peaks at chromosomes 3 and 17 with LOD scores both over 4.0 (Figure 2). By the convention mentioned previously, a LOD score of 3.3 or higher would represent significant evidence of linkage. These exceedingly high values convinced us to create congenic mice of BALBx129 to isolate these two chromosomes of interest individually. As the congenic mice require several generations to develop, more of the BALBx129 nur12 mutants are collected for further genome scans to improve data and confidence levels. This led to figures 5 and 6, which tell us that chromosome 15 is also an important location that possibly contains modifier genes. However, the same data also presented difficulties as both chromosomes 3 and 17 had diminished LOD score values significantly even though the 45 mice sampled from half plate 3 were from the same founder of half plates 1 and 2, which should have provided a relatively similar genomic scan to the 1st and 2nd half plates due to their close relations. Nevertheless, the congenic mice were already well under way in N3 and N4 generations, so we decided to test the congenic mice and analyze them. Several pairs of mice have been collected from the congenic strain and have been dissected for phenotypical analysis, but the range of phenotypes observed still varied from 1 to 3. Therefore, we decided to include a 4th and 5th half plate sample of the BALBx129 nur12 mutant mice to increase both the population as well as our confidence values. Based on the data from phenotypic analysis of the congenic N3 and N4 mice of chromosome 3, we decided to create a new strain of congenic mice with both chromosomes 3 and 17 isolated to the dominant 129 strain. This strain is
now underway and the phenotypes of these congenic chromosome 3 and 17 *nur12*
mutants will be analyzed in the near future.

Considering the data up to this point, we have determined strong linkages in
chromosomes 3 (LOD = 4.74) and 17 (LOD = 4.67) in the BALBx129 strain, and a
very likely singular linkage in chromosome 11 (LOD = 3.69) in the BALBxB6 cross.
With these high LOD scores, we have great confidence that there exist genes that
encode the modifiers for our phenotype of *nur12* mutants for both intercrosses. With
the development of congenic strains, we hope to further narrow down the loci of
interest and find tightly linked effects between strain differences and phenotypic
changes once this congenic strain is ready.
REFERENCES

Alcaraz, W., Gold D., Raponi E., Gent P., Concepcion D., Hamilton B.  Zfp423 Controls Proliferation and Differentiation of Neural Precursors in Cerebellar Vermis Formation. *PNAS.* 2006; **103**(51): 19424-29.


*UCSC Genome Bioinformatics*. 16 April 2008. University of Santa Cruz. 22 April 2008 <http://genome.ucsc.edu/>