Kathryn Carnazza
Senior Honors Thesis

Differential Gene Expression in a Mouse Model of Huntington's Disease

Abstract

Huntington's disease (HD) is a fatal genetic neurodegenerative disorder. It is inherited in an autosomal dominant fashion. HD typically manifests itself in middle age, though in some cases, symptoms appear earlier. Patients with HD suffer from a gross motor dysfunction, neuronal atrophy, cognitive impairments, psychiatric disturbances, and metabolic disruption, among other symptoms. Death follows anywhere from 10-25 years after the onset of symptoms. The causative mutation of this disease, a CAG_n expansion in the *HTT* gene, was discovered more than 17 years ago. This gene codes for huntingtin, which is essential for both development and survival of many organisms. However, the exact molecular workings of this disease are still unknown. Partially due to the paucity of this information, there is no cure for the disease, and treatment consists only of controlling certain symptoms to improve quality of life. Research is focused on elucidating the pathways and molecules that contribute to disease progression. Much of this research utilizes animal models of HD, including a mouse model with a CAG_n expansion inserted into the murine homologue of *HTT*.

A collaborative study with the German Mouse Clinic with the aim of better understanding Huntington's disease pathogenesis led to the identification of a number of differentially expressed genes in brain and liver tissues of a knock-in mouse model of the disorder. In this study, we sought to further characterize the expression of a selection of these genes both knock-in and wild-type mice. We employed TaqMan real-time PCR assays to investigate the previously observed changes in expression. The data we collected for a majority of the genes are inconclusive. However, our preliminary results are encouraging for two of the genes, *Prok1* and *Nfkbia*. We found *Prok1* to be significantly downregulated in the knock-in mice, and observed downregulation in *Nfkbia* in the knock-in mice as well. These two genes provide potential avenues for new hypotheses about the disease process in HD.

Introduction

Background and Human Phenotype

Huntington's disease (HD) is an autosomal dominant genetic disorder that is characterized by progressively worsening psychiatric symptoms, cognitive decline, and involuntary muscle movements followed by premature death [1]. Approximately 5-7 in 100,000 are affected with HD [5]. The disease is caused by an expanded CAG_n repeat in the first exon of the *IT15* (now called the *HTT*) gene, which results in an abnormally long polyglutamine tract near the N-terminus in the protein product huntingtin [2]. Symptoms typically appear in middle age, but there are rare cases of juvenile onset HD. Individuals with 40 or more repeats will develop Huntington's disease at some point in their lifespan, whereas any individual with fewer than 36 repeats will not. A repeat count of 36-39 triplet repeats results in incomplete penetrance. Repeat count is strongly inversely correlated with age of onset, with repeat length accounting for approximately 70% of variation [3]. Repeat instability is also seen somatically and in transmission from parent to child, particularly from father to child [13]. Treatment options are limited, and there is no cure.

A few medical professionals, such as Charles Waters, mentioned the condition in the 1800s [3]. In 1872, George Huntington described a hereditary chorea that presented itself in middle age and was accompanied by insanity in a lecture[4]. The chorea was named after Huntington, though the official name was changed from Huntington's chorea to Huntington's disease in the 1980s [3]. In 1983 the disorder was linked to chromosome 4, and ten yeas later, the causative gene was discovered [5]. As new laboratory techniques were invented and more interest was taken in the disease, the list of physical manifestations of the disease grew. Patients with HD display massive selective neuronal death with particular loss of medium spiny neurons in the caudate and putamen (striatum). Other brains regions show atrophy, and surviving cells show nuclear aggregation and neuronal dysfunction [5]. The chorea and neurological symptoms are accompanied by other symptoms, such as circadian rhythm dysregulation and metabolic dysfunction. Metabolic symptoms include muscle wasting and weight loss that does not stem from the chorea, decreased caloric intake, or other overt symptoms of HD [3,6,28].

Evidence of possible energy deficits in the brain and other tissues has been found in both patients with Huntington's disease and in cell line models of the disease[28].

Many studies have explored early symptoms of Huntington's disease or attempted to find differences between presymptomatic patients and control groups. These inquiries seek not only to better characterize the disorder, but also to pinpoint pathways that may play a role in later pathogenesis. Ideally, these early dysfunctions (if they exist) could lead to treatment before the disease begins to ravage the patient. Experiments have shown that 'presymptomatic' human individuals have various subtle deficiencies, such as cardiovascular autonomic dysregulation, weight loss despite higher caloric intake, decrease in white-matter brain tissue volume, and synaptic and cytoskeletal changes in neuronal cells [34,35,36,37,38]. Some of these alterations can be observed years before clinical onset of symptoms. For instance, Interleukin-6 (IL-6), an inflammatory molecule, has been detected at significantly higher levels in the plasma of presymptomatic patients up to 16 years before they show overt signs of HD as compared to control subjects [39]. It is not clear if these differences contribute to disease progress or if they are results of other dysfunction; however, some may prove to be attractive therapy targets.

Molecular Facets of Huntington's Disease

HD is one of a family of polyglutamine repeat diseases. All of the CAG_n repeat expansion disorders result in selective neurodegeneration after a certain age based primarily on repeat length. However, the polyglutamine tract is found in a different gene and affects dissimilar brain regions in each disorder [8]. This implies that it is not just the repeat expansion that causes the morphology and symptoms, but also the protein context. If the CAG_n expansion alone were sufficient for the disease state, each polyglutamine disorder would have a more similar symptomatic state. This has important ramifications for research direction and for model systems of the disease.

Though the *HTT* gene has been discovered, the exact biological role of huntingtin is still unknown, despite having been the focus of intense research. The protein has been implicated in transcriptional regulation and axonal transport, among other cell activities. Huntingtin also appears to be shuttled in and out of the nucleus [6,7]. It is known that wild-type huntingtin is present in all cell types and is essential for embryonic and post-

embryonic brain development and function [3,5]. However, humans with only one copy of *HTT* develop normally. This supports the theory that mutant huntingtin causes disease symptoms primarily through toxic gain of function rather than a loss of function. Additionally, transgenic expression of mutant huntingtin can rescue otherwise fatal huntingtin knock-out mice, which further substantiates this theory[6]. The mutant protein may also enhance the normal activities of wild-type huntingtin, which could contribute to the disease state [3,6,29]. It should be noted that certain studies have implicated a loss of wild-type function of *HTT* in the presence of the expanded repeat [28].

As with the wild-type function, the precise role of mutant huntingtin and the polyglutamine expansion is not fully understood. The abnormal number of CAG_n repeats in the mutant huntingtin leads to protein misfolding. Both the wild-type and the mutant proteins have a number of cleavage sites near the N-terminus end, and certain fragments from the mutant protein may be toxic [8]. The fragments and the abnormal protein conformations could affect the cell in a number of ways. The misfolded protein may interact inappropriately with other molecules, which interferes with normal cell processes, such as metabolism in the mitochondria and vesicle transport [8]. Certain fragments form aggregate inclusions in both the cytoplasm and the nucleus, whereas other intermediate aggregate formations remain soluble. These inclusions may also sequester other proteins that are important for cell functioning. Furthermore, both the full-length protein and the fragments are shuttled into the nucleus, where they affect gene expression by interacting with transcription factors and perhaps even by directly binding to the DNA. A variety of other factors, such as increased metabolic stress placed on the protein degradation pathways, excitotoxic effects, and inflammation have also been postulated to play a role in HD [6,34]. It is unclear which effects result in neurodegeneration and which are molecular bystanders or the cell's attempt to handle the abnormal protein products. For instance, inclusions are a hallmark of HD pathogenesis, and aggregates may sequester important cellular proteins and disrupt cell activities [7]. However, other data suggests that they are not responsible for the neuronal dysfunction; cells with inclusions tend to have better survival rates than cells without inclusions, and the soluble intermediates are in fact more toxic than the large inclusions [6]. Despite the sometimes conflicting data and multiple theories, a predominant hypothesis is that the protein

interactions and the transcriptional dysregulation lead to a cascade of events that damage the cell even as it attempts to correct the problem, eventually leading to neuronal death [7].

Genetic Models of Huntington's Disease

The uncertainty surrounding both the function of mutant and wild-type huntingtin and how these lead to the Huntington's disease state calls for extensive research. However, it is difficult, if not impossible, to study many of these molecular phenomena in humans. To overcome this problem, different cell lines and models were created. Cell lines from human and different model organism neuronal cells were established in order to better explore the cell biology of Huntington's disease. These neuronal cultures are good models of cell activity, though they do not replicate neuronal circuitry. Invertebrate models such as *C. elegans* and *Drosophila* have also been used. These are useful for the speed with which they reproduce and the number of experiments that can be done with them. This, along with their short lifespans, also makes them extremely attractive for screening potential treatments. However, their evolutionary distance from humans calls into question how well they replicate the disease state [6,30].

A variety of mouse models were generated to get a more accurate reproduction of HD. These mice fall into two main categories: transgenic and knock-in. These categories can be further broken down by the size of the transgene or the repeat count of the expansion that was inserted into the murine gene. The first, but not only, transgenic mouse line was the R6/2 strain, which has a fragment of exon 1 of the human gene, as well as the two murine copies of *Hdh,* the mouse homologue of *HTT*. Similar to HD patients, this mouse strain exhibits a phenotype that includes motor control problems, weight loss, premature death, and huntingtin aggregates in the neurons. R6/2 mice also show brain atrophy and aggressive behavior [7]. The YAC and BAC transgenic strains have a full-length mutant human gene inserted via a yeast or bacterial artificial chromosome into their genomes, respectively. Mice of these strains exhibit behavioral abnormalities and neuronal loss but live much longer than R6/2 mice [7, 10]. These YAC and BAC strains are more genetically accurate than the R6/2 mice in that they express a full length *HTT* gene. As noted, HD is not the only polyglutamine disease. If the expansion alone were sufficient to cause Huntington's disease, all disorders with a CAG_n expansion would have

very similar symptoms, which is not the case. The protein in which the expansion is contained is thus very important, and mice with only a fragment of *HTT* may show phenotypes that are not relevant to HD specifically. However, both the full-length and fragment transgenic mice are expressing a human gene as well as their own two *Hdh* genes. *HTT* is also being controlled by the human promoter as opposed to the mouse promoter, and so may not be expressed at levels typical for a mouse. Additionally, the transgene inserts randomly into the genome. Therefore, it is possible that the insert could interrupt a murine gene, which could affect the behavior or physiology of the mouse. Due to these factors, it can be difficult to separate which aspects of the murine phenotype are relevant to HD.

With this in mind, knock-in models of HD were created. These mice have an expanded CAG_n repeat inserted into their *Hdh* gene at the same location as in human HD patients. Therefore, they have a full-length gene, and the endogenous promoter controls the gene. This means that any changes to the mouse should result only from the expansion, making it a reliable model [7]. They also are heterozygous, and thus mimic a majority of human patients on a genetic level. At first the knock-in mice did not live up to expectations, as they lacked the more obvious phenotype of the transgenic models [7,11]. However, as further studies were done on mice with a repeat count over 90, behavioral, neuroanatomical, and motor abnormalities were found [10,11]. Knock-in mice have periods of hypo- and hyperactivity based on their age, gait abnormalities, and rotarod difficulties, and some strains have aggressive tendencies [7,10,11]. They do not show any neuronal loss, but they do have selective aggregate formation and neuronal abnormalities, such as axonal degradation and atypical nuclear staining. This selectivity follows the pattern of human HD [11]. Knock-in mice have a variety of backgrounds and typical repeat lengths, though the exact repeat count also varies mouse to mouse because of the inter-generational instability seen in CAG repeat number. This mirrors instability seen in human pedigrees [12]. Both of these traits highlight the usefulness of this model. Because the phenotype is milder and has a slower progression than that of the transgenic strains, the knock-ins may shed more light on earlier stages of the Huntington's disease and better illuminate the stages of progression. For example, brain tissues of knock-in mice were assayed at different ages for aggregate formation, and this timeline was compared to the

onset of behavioral abnormalities. The behavioral abnormalities presented themselves before aggregates were found, providing further evidence that these inclusions are not the only factor in the pathogenesis of HD [11]. The knock-in mice are the most genetically accurate models, and are therefore widely studied.

The German Mouse Clinic and Aim of This Project

A group of HdhQ111 knock-in mice were sent from Massachusetts General Hospital to the German Mouse Clinic (GMC; http://www.mouseclinic.de/) to undergo rigorous behavioral, anatomical, and molecular analysis. The GMC characterizes different mouse models using unbiased behavioral and anatomical standard screens in order to examine all aspects of the model. This allows subtle alterations between the wild-type and knock-in mouse to be observed that may otherwise have not been noted. The molecular component investigated two tissues of choice; this involved the full organ, rather than individual regions of that organ. We were interested in the brain and the liver because of the metabolic abnormalities seen in HD. Microarray analysis was therefore performed on the brain and liver tissue from 4 wild-type males and 4 heterozygote male knock-in mice that were 18 weeks old. Seventeen genes in the brain and 26 genes in the liver were found to be significantly differentially expressed.

As the exact pathogenic mechanism of HD is not known, any upregulated or downregulated genes in the knock-in mice as compared to the wild-type mice may play a role in the disease state. On the other hand, altered gene expression may be a result rather than a causative agent in HD. In either case, we wanted to first look into the robustness of the results found by microarray analysis. In order to carry this out, 3 genes that were found to be differentially expressed in the liver and 3 genes found to be differentially expressed in the brain were chosen to be assayed using qRT-PCR (quantitative real-time PCR). Two more genes in the brain were later added to the 6. The original 6 genes were picked based on possible biological significance and the fold change seen by the GMC. The two genes added later were selected for their high expression levels in the brain based on data from the Allen Brain Atlas and the Gene Expression Database (GXD). Prolactin (*Prl*), prokineticin 1 (*Prok1*), and pro-opiomelanocortin (*Pomc*) were selected as the original brain genes. PR domain containing 16 (*Prdm16*) and Down syndrome cell

adhesion molecule-like 1 (*Dscaml1*) were the two genes added later in the project. Their liver counterparts were hedgehog (*Shh*), nuclear factor of kappa light polypeptide enhancer (*Nfkbia*), and angiotensin converting enzyme (*Ace)* (See Table 2 for summary).

Differentially Expression Genes in the Brain

The GMC found *Prl* to be downregulated 23 fold in knock-in mice, which was a startling result. We did not expect to see such drastic alterations in such a subtle model, so this fold change was interesting. Prolactin has been explored previously in HD patients. Some studies found lower levels of *Prl* in the blood serum of patients as compared to controls, and some found higher levels [13,14]. Another study found that baseline *Prl* levels decreased in the patients over time [15]. Because of the previous studies and the huge fold change, *Prl* appeared to be a good gene to verify. *Pomc* was downregulated by 6.99 fold in the mutant mice and the gene is known to play a role in body weight, ingestive behaviors, and in energy homeostasis [17]. As HD patients have metabolic symptoms and the fold change was on the higher side of what we expected to see in a knock-in model, *Pomc* was an interesting gene to investigate. The other genes were not up or downregulated to such drastic extents. *Prok1* had a fold change of -1.77 from the wild-type to the knock-in. Prokineticins play a role in a variety of processes, including circadian rhythms, immune response, and the development of nervous system control over the gastrointestinal tract [18]. The two genes selected later were picked primarily because of their high expression levels in the brain and not because of their biological function. Still, the genes have been implicated in pathways that make them potentially relevant. *Prdm16* may play a role in energy homeostasis and neuronal development [20,21]. *Dscaml1* is a cell adhesion molecule. It belongs to a family of adhesion molecules that regulate neuronal networks and act in axonal guidance. *Dscaml1* appears to be involved in creating and maintaining networks of neurons [22].

Differentially Expressed Genes in the Liver

The genes in the brain were of the most interest to us, as the dementia and motor control issue of Huntington's patients are the most striking symptoms. However, the

metabolic problems that are also seen in HD make the liver a tissue of interest. *Shh* was downregulated in knock-in mice by a 3-fold change. It is typically, but not solely, involved in embryogenesis; it also plays a role in maintaining tissue in addition to regulating development. Studies have implicated Liver X Receptors (*LXR*) as possible negative regulators of *Shh*. These receptors are involved in cholesterol and lipid metabolism regulation, and their transcription may be regulated by wild-type (but not mutant) *HTT* in model organisms [31,32]. These studies were not performed in liver tissue of mice, but they nonetheless make the downregulation of *Shh* observed by the GMC intriguing. *Shh* has also been shown to be activated by brain-derived neurotrophic factor (BDNF) and acts as a neuroprotective agent against mitochondrial dysfunction. BDNF is downregulated in HD, and this lack of BDNF has been implicated in pathenogensis of HD [23,24]. The GMC did not find *Shh* to be differentially expressed in the brain, but this link contributes to the possibility that this gene may play a role in HD. *Nfkbia* is an inhibitor of nuclear factor kappa-beta (NF-κβ). NF-κβ is involved in inflammation, apoptosis, and cell differentiation and proliferation in various roles. The inhibition of NF-κβ is associated with apoptosis. *Nfkbia* was upregulated in the knock-in mice by a fold change of 1.38 [25]. Lastly, *Ace* was upregulated in knock-in mice by a 1.64 fold change. *Ace* is involved in regulating blood pressure, salt homeostasis, and in degradation of amyloid-beta, which is a protein that is a key component of Alzheimer's Disease (AD). Higher activation of *Ace* has been linked to AD in several studies, and *Ace* inhibitors have been suggested as a treatment for AD and other dementias [26]. Polymorphisms of *Ace* have also been suggested as a contributor to depression [27]. These prior linkages made *Ace* a gene whose differential expression would be intriguing to replicate. By investigating the results of the GMC, we could not only verify their data, but also delve further into the genes that were differentially expressed and had biological activities that may suggest new hypotheses for pathways of HD pathogenesis.

Methods and Materials

Laboratory Animals

Knock-in heterozygote HdhQ111 and related wild-type mice from the C57BL/6J background were used in this experiment. The mice were housed 5 to a cage by gender and given food and water ad libitum. Normal temperature and light conditions were used. Genotyping was performed by Southern blot and PCR assay. The four knock-in mice used in this project had repeat counts of 113, 123, 126, and 130. The CAG count in the wild-type allele is 7 for all mice used. See Appendix Figure 5 for summary of mouse information.

Tissue Dissection

At 18 weeks of age the mice were sacrificed by spine dislocation. One mouse (M1106) was 20 weeks old at the time of sacrifice. One hemisphere of the brain and a piece of the liver were dissected and snap-frozen in liquid nitrogen. The other brain hemisphere and another piece of the liver were frozen in dry ice. A section of the tail was also taken to check the genotype of the mouse.

Tissue Homogenization and RNA Extraction

Tissue homogenization and RNA extraction were performed using TRIzol reagent (Invitrogen) on the snap-frozen samples. One mL of TRIzol Reagent was used for each tissue sample. Samples were incubated for 5 minutes at room temperature before 0.2 mL of chloroform was added. The tubes were shaken and then incubated for another 3 minutes at room temperature. The samples were then centrifuged at 12,000 rpm for 15 minutes at 4° C. The colorless aqueous phase at top containing the RNA was drawn off and placed in fresh tubes. RNA was then precipitated by adding 0.5 mL isopropyl alcohol. The samples were then incubated at room temperature for 10 minutes before being centrifuged at 12,000 rpm for 10 minutes at 4° C. The supernatant was decanted, and the pellet was washed with 1 mL of 75% ethanol. The samples were mixed by vortexing and then centrifuged at 7,500 rpm for 5 minutes at 4° C. The sample was then air dried for no more than 10 minutes before the RNA was redissolved in DEPC treated water and incubated for 10 minutes at 55-60° C before being stored at -80° C. 50µL of

DEPC water was originally used to redissolve the RNA; more water was added as necessary.

RNA Quantification

The RNA was quantified using a NanoDrop Spectrophotometer. Two µL of the original RNA dilution were quantified, and then the samples were diluted to approximately 1 µg / µl in preparation for cDNA synthesis. The dilution was also quantified to ensure that the dilution had produced the desired concentration. If the samples originally were less concentrated than 1 µg / µl, a greater volume of RNA was used in cDNA synthesis

DNase Treatment

Each sample was DNase treated directly before cDNA synthesis. DNase treatment was performed using Invitrogen Deoxyribnuclease I, Amplification Grade. One µg of RNA, 1 µL of 10X DNase I Reaction Buffer, 1µL DNase I, Amp Grade, 1 U/µL, and enough DEPC-treated water to bring the final reaction volume to 10µL were mixed and incubated at room temperature for 15 minutes. One µL of 25 mM EDTA was then added to inactivate the DNase I. The samples were then incubated at 65° C for 10 minutes. This treatment was sometimes used for 2 µg of RNA by scaling up the reaction

cDNA Synthesis

Invitrogen's Superscript III First-Strand Synthesis SuperMix for qRT-PCR was used to make the cDNA from the RNA template. The 11 µL from the DNase treatment were mixed with 12 µL of 2X RT Reaction Mix and 2 µL of RT Enzyme Mix. The samples were mixed and incubated for 10 minutes at 25° C and then for 30 minutes at 50° C. One µL of *E. coli* RNase H was then added and the tubes were incubated for 20 minutes at 37° C.

Real-Time PCR

Real-Time PCR was carried out using TaqMan assays (Applied Biosystems) and a LightCycler480 real-time PCR machine (Roche). Eight different TaqMan gene expression assays were obtained from Applied Biosystems (See Table 1). Actin beta was used as an

internal control for all experiments. –RT controls were included. One –RT control underwent DNase treatment and then cDNA synthesis, but the RT Enzyme Mix was not added. The other –RT control was neither DNase treated nor mixed with RT Enzyme Mix. These –RT controls were both made from samples of L1777 (wild-type). The TaqMan cycling conditions that were used were: Pre-Incubation (50° for 2 minutes, 95° for 10 minutes), Amplification (45 Cycles of: 95° for 15 seconds, 60° for 1 minute, 72° for 1 second), and Cooling (40° for 30 seconds).

Gene Symbol	Assay ID	Amplicon Length	Primer-Probe Set Hybridization Location
Prl	Mm00599950_m1	139	Exon 2 and 3
Pomc	Mm00435874_m1	60	Exon 2 and 3
Prok1	Mm01204733_m1	90	Exon 2 and 3
Shh	Mm00436528_m1	62	Exon 2 and 3
Nfkbia	Mm00477798_m1	70	Exon 2 and 3
Ace	Mm00802048_m1	79	Exon 17 and 18
Prdm16	Mm00712556_m1	62	Exon 14 and 15
Dscaml1	Mm01174253_m1	57	Exon 1 and 2
Actin-β	Mm00607939_s1	115	Exon 6

Table 1: Gene name, ID, amplicon length, and location primer-probe bonding location. The primers that spanned introns had the forward primer hybridize to one exon and the reverse primer hybridize to the other exon, while the probe hybridized to one of the exons. Information obtained from Applied Biosystems website

Data Analysis:

Data analysis was performed in Excel and the LightCycler Software. The target gene expression level in each well was normalized to the internal actin control. All results were then normalized to one of the wild-type samples (L1776); this sample was used as the calibrator for all experiments. Student's two sample t-test assuming equal variances was performed on the normalized data. Normalized standard error between triplicates for each sample was calculated in the LightCycler Software. The program used the following equations to calculate relative gene expression:

$$R = 2^{-[\Delta CP\ sample - \Delta CP\ control]}$$

$$R = 2^{-\Delta\Delta CP}$$

This is termed the 'delta-delta cp' method for determining gene expression levels, where CP is the crossing point. Here, ΔCP Sample is the difference between the target gene and

reference gene CPs, and ΔCP Control is the calibrator sample [33].

Results

TaqMan Background

The ideal TaqMan assay, like other PCR assays, involves one double stranded target sequence separating, binding to two primers, and replicating to create two new copies in each cycle of the reaction. Unlike other kinds of PCR, the TaqMan assay uses a primer-probe set to increase specificity of the results. This set consists of two primer sequences and a probe containing a quencher molecule and fluorescent tag. This quencher prevents the tag from fluorescing when the two molecules are in close proximity to each other. Each time the target hybridizes with a primer-probe set and is replicated, the fluorescent probe is cleaved from the quencher and primer by Taq polymerase. This allows the tag to fluoresce; thus, each fluorescent signal ideally represents one new copy of the target (See Figure 1).

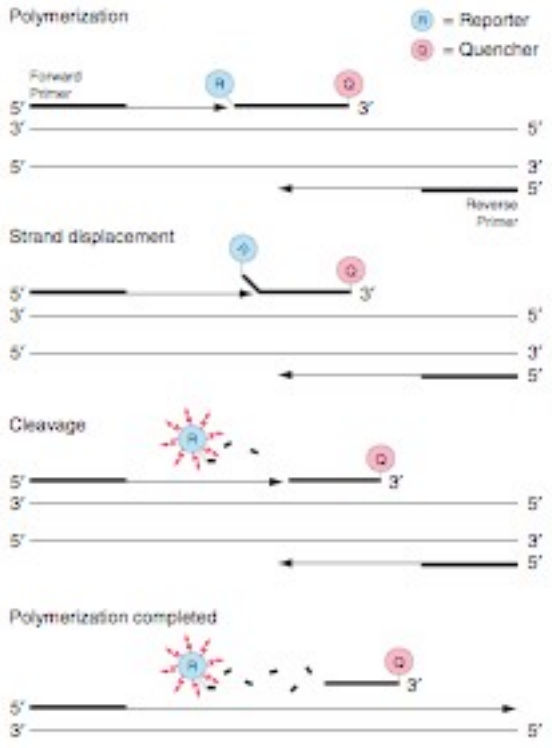

Figure 1: A visual representation of the primer-probe set. From Applied Biosystems TaqMan Universal PCR Master Mix Protocal

The primer-probe design also allows the reference gene to be run in the same well as the target gene; they are each tagged with a molecule that fluoresces at a different wavelength of light. Two fluorescent readings are done each cycle. All target primer-

probe sets in this experiment fluoresce at 465-510 nm, while the reference sets fluoresce at 533-580 nm.

Once the level of fluorescence in a given well reaches a certain point, it crosses the threshold between background noise and actual amplification. This transition can be measured in time and assigned a call time or can be measured in cycles and called a crossing point; they are ultimately different names for the same thing. An earlier call-time is indicative of higher concentrations of the target. If a call time comes after a certain time in the cycling conditions, it is considered to be less reliable because it is more likely to have occurred by chance rather than by actual amplification. Any crossing point after 40 cycles is not considered to be reliable at all.

Optimizing the Assay

Each primer-probe set amplifies at a certain concentration of cDNA particularly well. This concentration is defined by consistent call times between the replicates of that concentration and by a reliable call-time. In order to ascertain the best concentration for each primer-probe set, an efficiency plate was set up with a serial dilution of cDNA for each target gene using one wild-type sample. A standard curve was generated from the qPCR data, and the concentration that amplified the most reliably was selected to be used for the experiment with all of the samples. However, not all the genes tested amplified well in this range of concentrations. Figure 2 shows a successful standard curve for one of the assayed genes. (See Appendix Figures 1 and 2 for an example of standard curves that were unsuccessful and Tables 3 and 4 for a list of optimized and non-optimized genes)

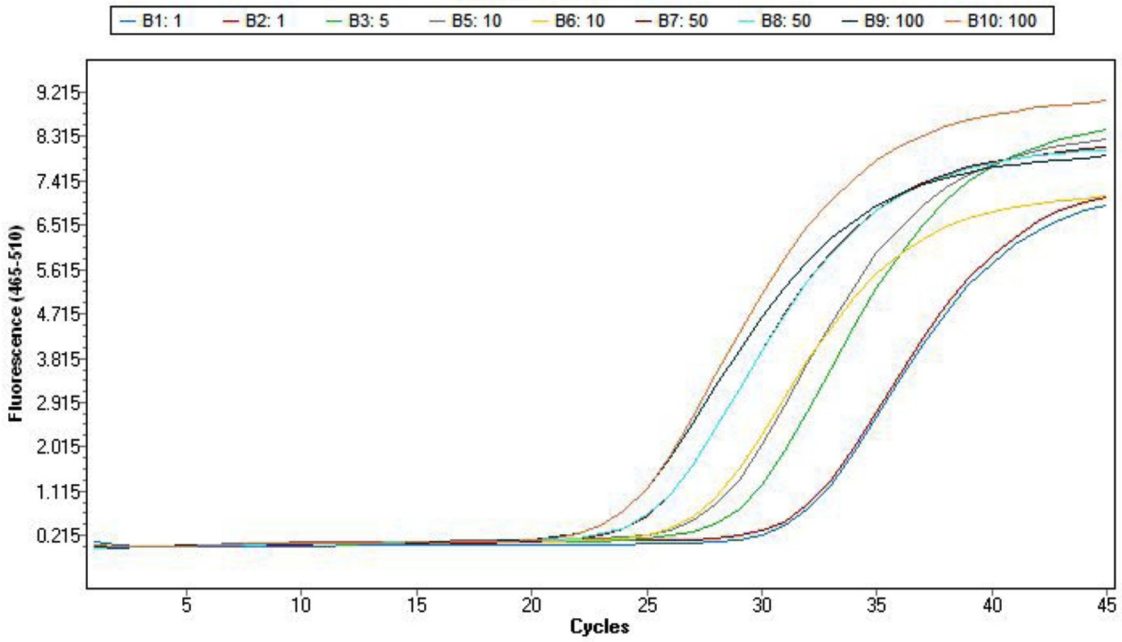

Figure 2: An example of a gene that amplifies well. Each curve represents one well, and the curve right next to it is the replicate well for that concentration. The concentration decreases from left to right. The key at the top represents the range of the amount of cDNA per well (from 1µg to 100 µg per well)

Once the run is completed, the call time or crossing point of all of the wells can be plotted against the log of the concentrations to generate a standard curve. The standard curve for Figure 2 is shown below in Figure 3.

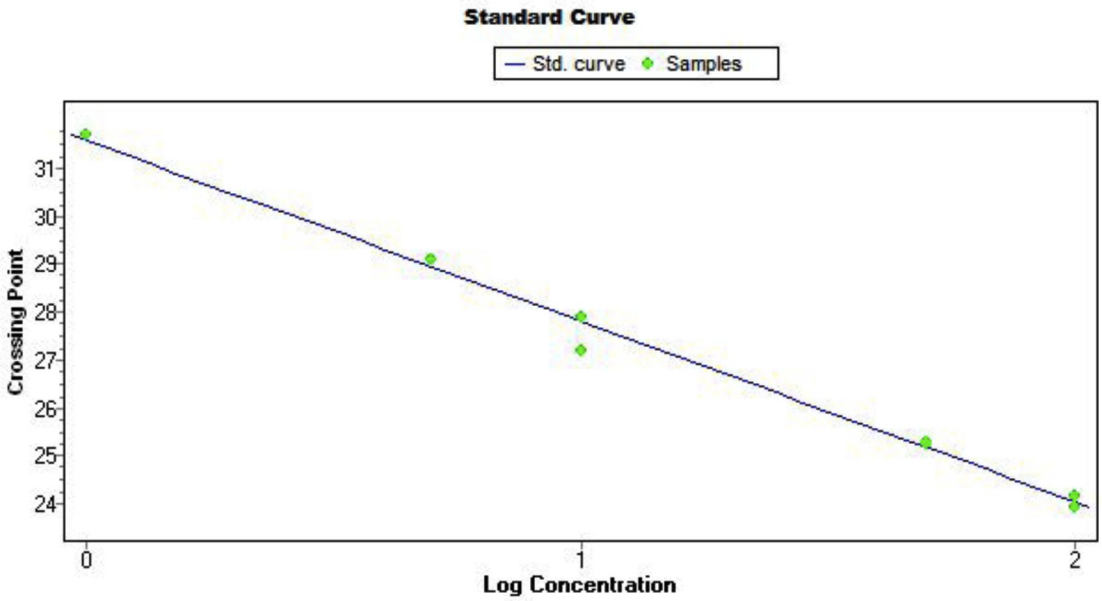

Figure 3: The log of the concentration versus the crossing point of the wells.

There are two important factors in choosing an optimal concentration. First, the well at the ideal concentration should be close to the standard curve line. Second, this amplification should be replicable; therefore, the two replicate wells should be close to each other. Here, the second highest concentration best meets these two criteria. Theoretically, if two concentrations met these two criteria equally well, either could be used; however, we did not encounter this in our experiment.

Genes Chosen and Summary of Runs

In order to determine the robustness of the fold changes observed by the GMC, eight total genes were selected to be assayed with TaqMan probes. See Table 2 for a summary of the genes and the fold changes originally observed

	Gene	Fold Change Seen by GMC (KI vs. WT)
Brain	Prolactin (*Prl)*	-23.29
	Pomc	-6.99
	Prokineticin *(Prok1*)	-1.77
	Prdm16	-1.89
	Dscaml1	-2.77
Liver	Hedgehog (*Shh)*	-3.05
	Nfkbia	1.38
	Ace	1.64

Table 2: Summary of the genes chosen to assay along with the fold changes originally seen by the German Mouse Clinic

Originally 6 male mice (2 wild-type and 4 mutant) and 2 female wild-type mice were sacrificed, and their tissues were used to extract RNA and synthesize cDNA. Female mice were used because there were no wild-type males of the right age at the time. The female mice were to be used until male mice could be obtained. However, when we assayed the brain genes in these eight samples, the female data was very different than the two wild-type males, and experiments were halted until we could acquire the male mice (Data from female mice not shown). Once we had a full set of 4 WT and 4 KI males, we repeated the experiments that we had previously performed (See Appendix Figure 5 for a list of all mice sacrificed and their gender, genotype, and date of death).

A benefit of these TaqMan assays is that the internal control is run in the same well as the target gene, thus minimizing the effects of experimental error. However, if the concentration of the internal control primer-probe set is too high, it can compete with the target for materials, such as polymerase and nucleotides, and ultimately affect the results. This can be especially problematic for lower expressed target genes. Using a lower concentration of the actin primer-probe set, called a primer-limited set, reduces the probability of this happening. In our experiment, this primer-limited set was originally ordered, but the regular concentration arrived instead. This problem was not identified until after the optimized genes had been run at least once, and as we saw variability in our initial results, we redid the assays. However, switching to the low concentration primer did not noticeably affect the patterns of expression (high concentration primer run vs. low concentration primer run not shown). See Tables 3 and 4 for summary of runs completed with different primers.

Gene	Using Full Concentration Actin Primer-Probe (High Concentration)				
	Standard Curve	Optimized?	Optimal Concentration	Number of Runs (With Females)	Number of Runs (With Males)
Pomc	X	Yes	10 µg/well	X	

Prok1	X	Yes	50 µg/well	X	
Prl	X	No	N/A		
Ace	X	No	N/A		X
Shh	X	No	N/A		X
Nfkbia	X	Yes	50 µg/well		X

Table 3: A summary of the number of qPCR runs done with the high concentration of the Actin internal control for each gene including standard curve generation, whether or not the standard curve yielded reliable results, runs with the female mice, and then runs with all male mice. Each X represents one run. Female mice were not used once 8 male mice were obtained

Gene	Using Primer- Limited Actin Primer-Probe (Low Concentration)			
	Standard Curve	Optimized?	Optimal Concentration	Number of Full Runs (With Males)
Pomc	X	Yes	10 µg/well	X
Prok1	X	Yes	50 µg/well	XX
Prl	X	No	N/A	
Prdm16	X	Yes	50 µg/well	XX
Dscaml1	X	Yes	50 µg/well	XX
Ace	X	No		
Shh	X	No		
Nfkbia	X	Yes	50 µg/well	

Table 4: A summary of the number of qPCR runs done with the low concentration of the Actin internal control for each gene including standard curve generation, whether or not the standard curve yielded reliable results, and total number of runs. Each X represents one run. Female mice were not used once 8 male mice were obtained

All data in this paper represents the results of a single run. Some genes were run multiple times. The patterns of expression were consistent between runs, and call times were similar (Data not shown).

RNA Quality: Degradation and Assaying for Genomic DNA

After assaying genes in the full set of eight male mice, we observed a fairly consistent down regulation in the two wild-type mice that were sacrificed later than the original six (See Appendix Figures 6-9). We were initially unsure of what was causing this variability in the mice, and attempted to find the cause of, or diminish, the variation through a number of means. This included checking for genomic DNA contamination, changing primer concentration, assaying genes with higher biological levels of

expression, and assessing RNA quality. This last factor was the explanation for a majority of our variation.

A gel was run to ascertain whether or not there was RNase contamination in the RNA that had been extracted from the liver or the brain of any of the mice. We had checked RNA quality when RNA from the first batch of mice was extracted, and it showed no degradation (Data not shown). The call times of the actin internal control were not very different between samples throughout our experiments, and so because of these two facts, we had assumed that RNA quality was not an issue. When we ran the gel, we expected to see the characteristic 28s and 18s bands that indicate the presence of ribosomal RNA, as we extracted total RNA, not specifically mRNA. However, the gels showed that there was degradation in the RNA extracted from the brain of the two newer mice, M1285 and M1106 (See Figure 4). The RNA from the liver tissue was intact, though not of as good quality as we had hoped. Stronger or weaker banding patterns do not correlate with expression levels (See Figure 5, Figure 11). However, these results highlight the need for more careful sample preparation and storage in the future.

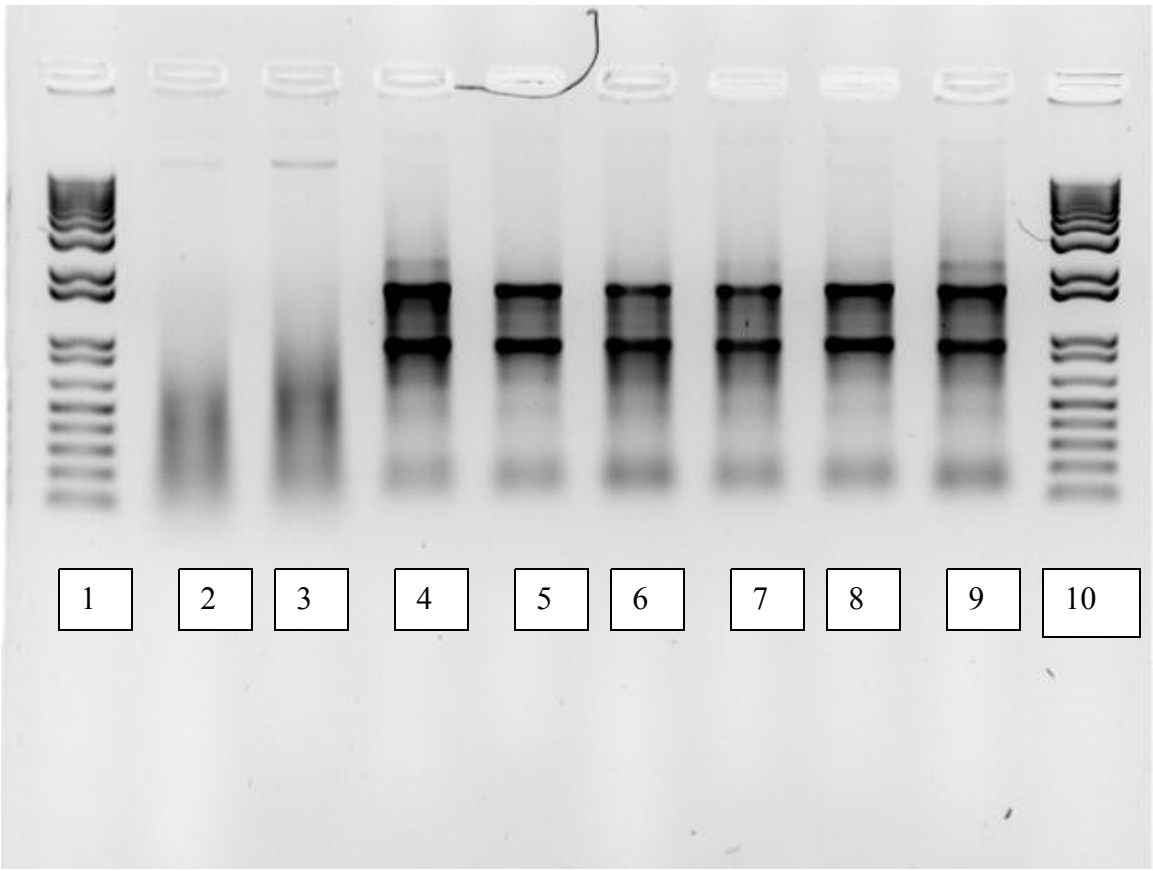

Figure 4: A 1% agarose gel of 1µg of RNA extracted from the brains of eight male mice. Lanes 1 and 10 are 100bp ladders. Lanes 2 through 9 are (in order): M1825, M1106, L1776, L1777, L1775, L1778, L1784, and L1785. In other mice, the 28s and 18s ribosomal bands can be seen. Staining was done with ethidium bromide

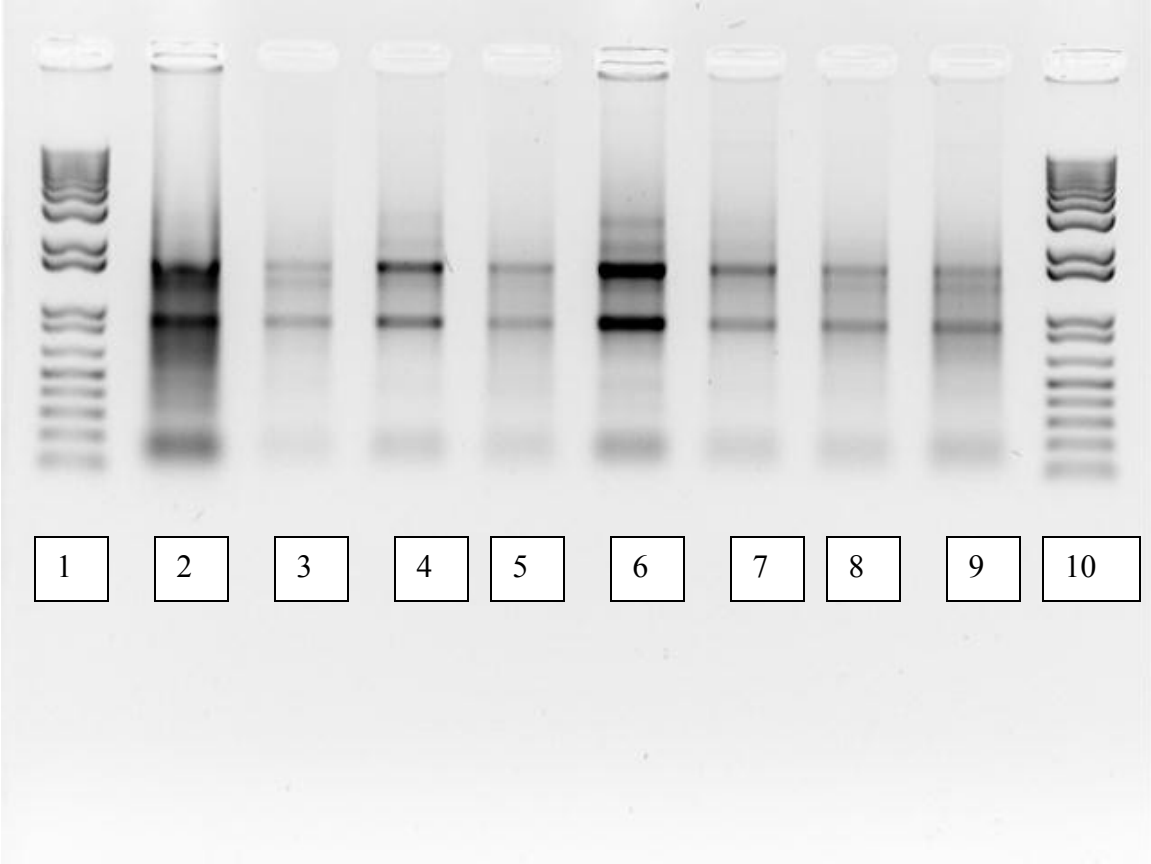

Figure 5: A 1% agarose gel of 1μg of RNA extracted from the livers of the eight male mice. Lanes 1 and 10 are 100bp ladders. Lanes 2 through 9 are (in order): M1825, M1106, L1776, L1777, L1775, L1778, L1784, and L1785. Staining was done with ethidium bromide. In all samples, the 28s and 18s ribosomal RNA bands can be seen.

Another gel was run to re-check these results and also to better ascertain when the contamination had occurred. Each RNA extraction from the brain resulted in a stock solution that needed to be diluted. This dilution was used to make the cDNA that was assayed. These dilutions and the stock solutions of all four wild-type mice were run on the gel. It showed that the RNA in both the dilution and the stock solutions of the two new mice was degraded, indicating the contamination had likely occurred during extraction or between extraction and the creation of the dilution (See Figure 6).

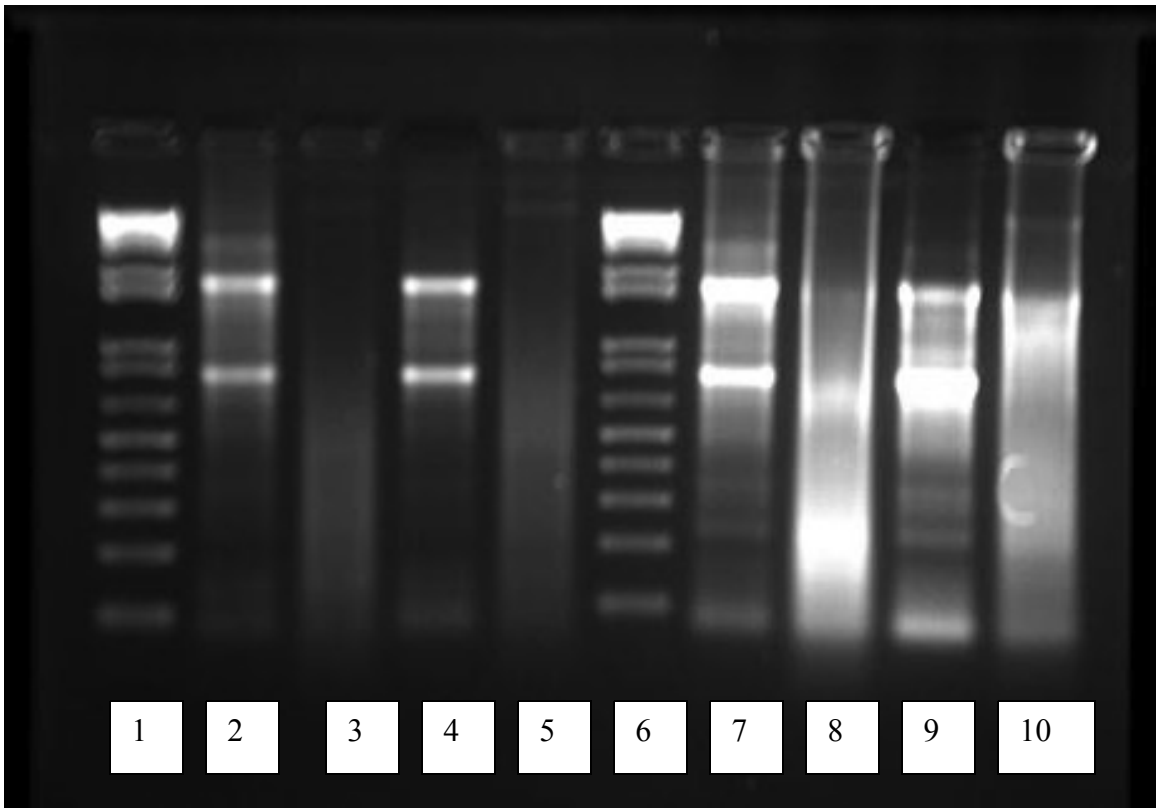

Figure 6: A 1% agarose gel was run with the dilutions and stock solutions of the four wild type mice. 1µg of RNA was used of the dilutions, and 1µl of RNA was used for the stock solutions. Lanes 1 and 6 are 100bp ladders. Lanes 2 to 5 are the dilutions, and lanes 7 to 10 are the stock solutions. Lanes 2 and 7 are from the mouse L1776; Lanes 3 and 8 from M1285; Lanes 4 and 9 from L1777 and Lanes 5 and 10 are from M1106.

We did not use the data from M1285 or M1106 in our data analysis of the genes expressed in the brain because of the degradation we saw. We did have the other hemisphere of the brain for these two mice, but we did not have enough time to go from the intact tissue to the real-time PCR assays, and likewise not enough time to sacrifice new mice. This did decrease our sample size significantly, and therefore our results are only preliminary. As the RNA extracted from the liver was intact, we included all samples in the analysis of the liver genes.

In addition to RNase contamination, possible genomic DNA contamination was a concern, though the samples were DNAse treated. If there was genomic DNA present and it was amplified by the assay, it could cause variable and confusing results. The primer-probe sets were also designed around large introns, as the two primers hybridize to

different exons, and the fluorescent probe and quencher set bind in between the forward and reverse primer binding sites (See Table 1 in Methods). Because of the short time span of the cycles of denaturation and renaturation, amplification of large products is very difficult. Even if genomic DNA were present, the Taq polymerase would not be able to replicate the DNA sequence quickly enough to reach and cleave the probe, thus minimizing the effect genomic DNA could feasibly have on results. All of the target genes had primers that bound to two different exons; only the actin primers were both placed inside one exon. To ensure that we were not amplifying genomic DNA, no reverse transcriptase (-RT) controls were run. One –RT control was DNase treated and did not undergo reverse transcription, whereas the other control was not DNase treated and was not reverse transcribed. No amplification was seen of the target in the –RT control wells in either control. The well that was not DNase treated did have some amplification of actin, but the call time was too late to be reliable. (See Appendix, Figure 3 and 4). Therefore, even if the DNase treatment failed, the genomic DNA would not have affected results.

Expression of *Pomc, Prok1,* and *Prl* in Brain

Of the three original brain genes, *Pomc* and *Prok1* were optimized, whereas *Prl* was not (See Tables 3 and 4). Data from six samples from male mice (two wild-type and two knock-in) are shown for *Pomc* (see Figures 7 and 8) and *Prok1* (see Figures 9 and 10). Repeat count is also included in the legends. If these genes did play a role in HD, we might see a correlation between increasing repeat count and expression level of the gene. Therefore, all knock-in mice are arranged in ascending order, from smallest repeat count to largest.

We saw no distinct pattern in *Pomc* expression and were unable to reach any conclusions due to the small sample size. *Prok1*, on the other hand, was found to be significantly downregulated in knock-in mice. Also, a preliminary correlation can be seen with expression of the gene and increasing repeat count. It is not a statistic correlation, but is nonetheless a promising finding.

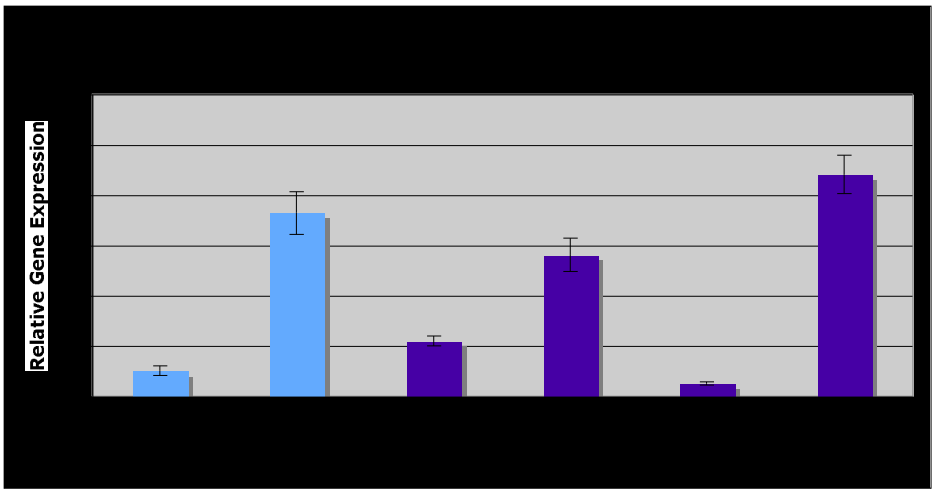

Figure 7: Relative expression of *Pomc* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. This figure has data from one run using the low concentration actin primer-probe. The blue bars represent wild-type mice, and the purple bars represent knock-in mice. Error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample.

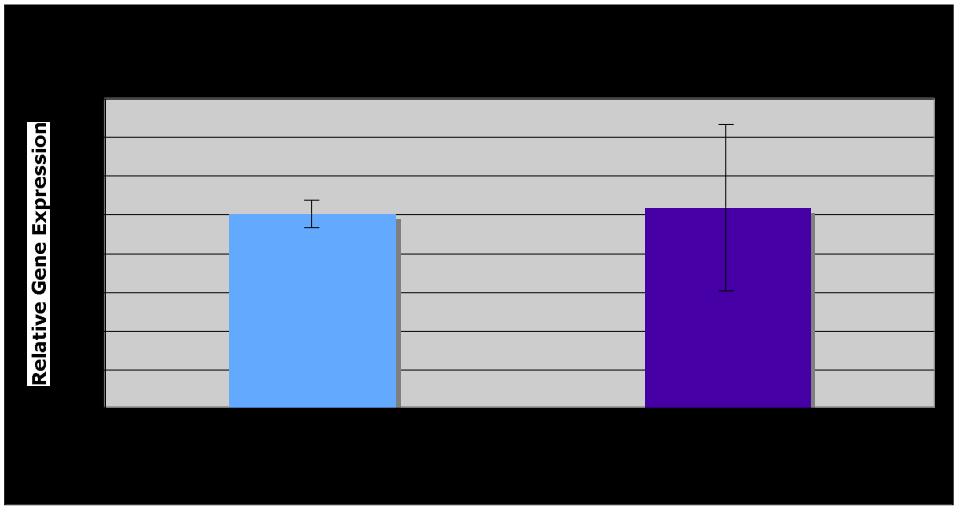

Figure 8: Relative expression of *Pomc* between the wild-type and knock-in mice normalized to the average gene expression level of the wild-type mice. Error bars represent the standard error within genotypes. No statistically significant difference was observed (p=0.97)

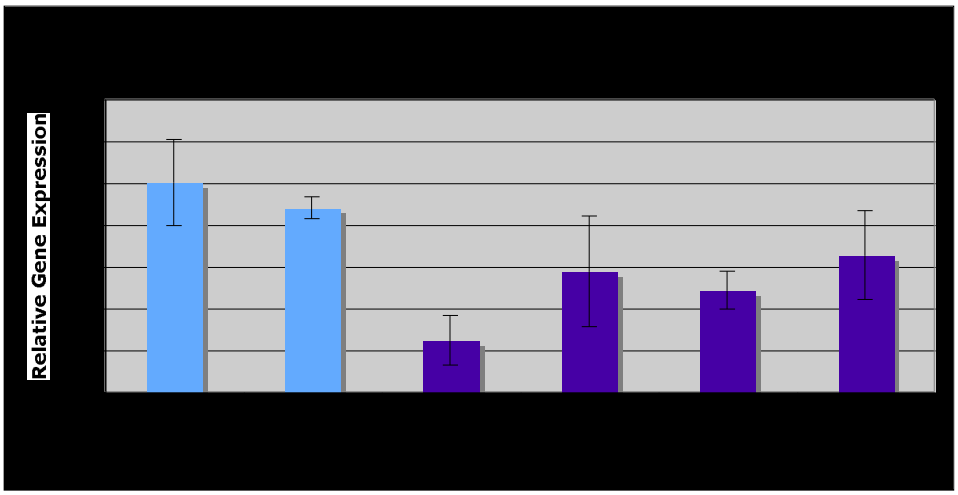

Figure 9: Relative expression of *Prok1* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. This figure has data from one run using the low concentration actin primer-probe. The blue bars represent wild-type mice, and the purple bars represent knock-in mice. The error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample.

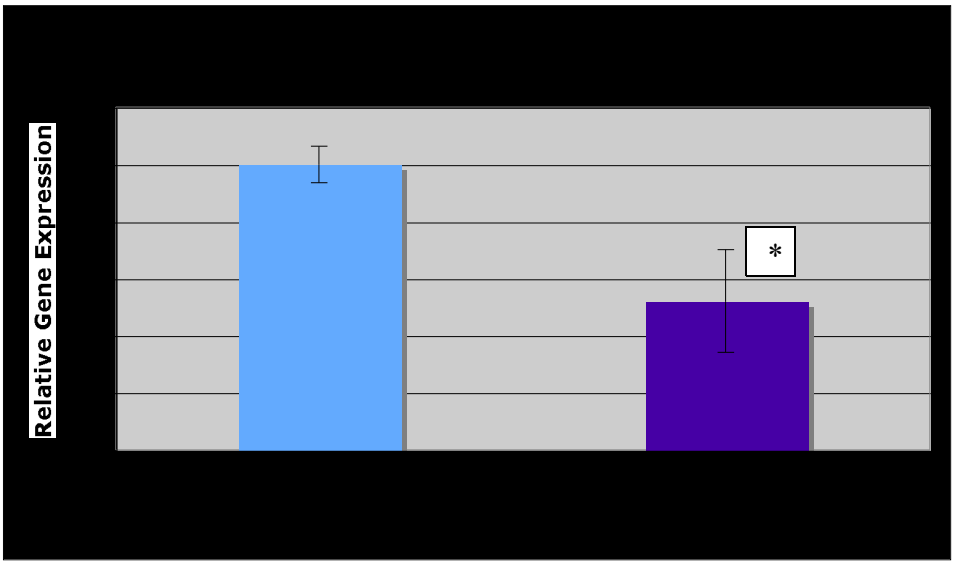

Figure 10: Relative expression of *Prok1* between the wild-type and knock-in mice normalized to the average gene expression level of the wild-type mice. Error bars represent the standard error within genotypes. A statistically significant difference was observed (p=0.03)

Expression of *Shh, Ace,* and *Nfkbia* in Liver

Nfkbia was the only gene assayed in the liver that could be optimized; *Shh* and *Ace* could not be (See Tables 3 and 4). As all the RNA extracted from the liver tissue was

intact, we included data from all samples. Certain samples did have fainter and darker banding patterns, but these do not correlate with decreased or increased expression levels (See Figure 5). *Nfkbia* was not statistically up or downregulated, and we saw variability in the wild-type samples.

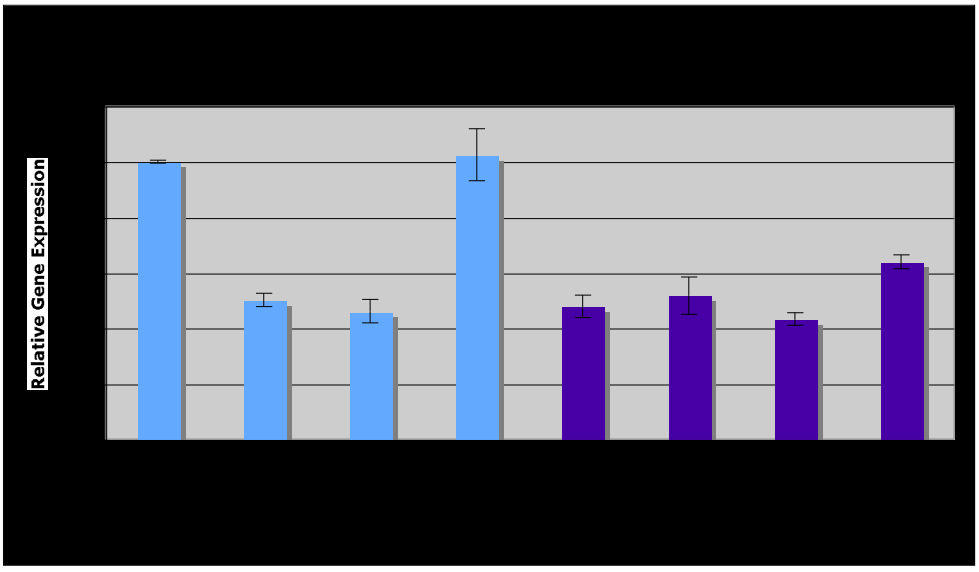

Figure 11: Relative expression of *Nfkbia* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. This figure has data from one run using the high concentration of actin primer-probe. The blue bars represent wild-type mice, and the purple bars represent knock-in mice. The error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample.

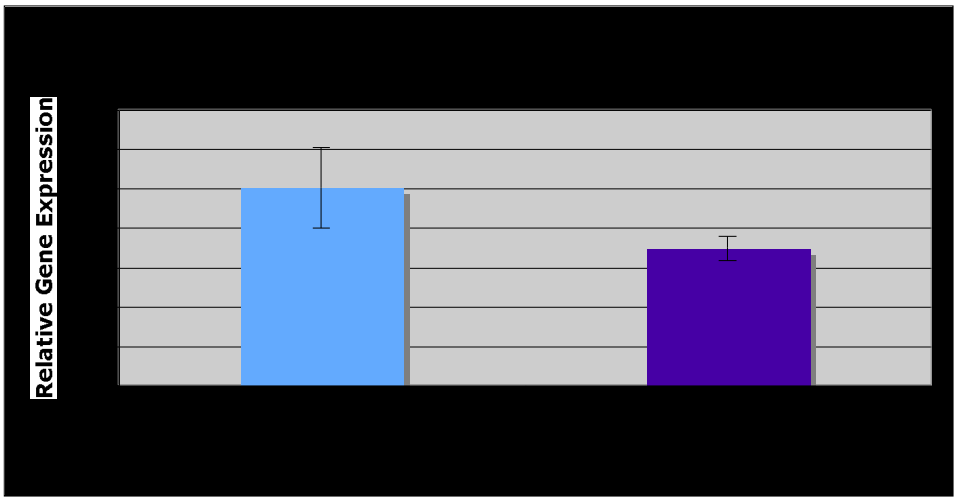

Figure 12: Relative expression of *Nkbia* between the wild-type and knock-in mice normalized to the average gene expression level of the wild-type mice. Error bars represent the standard error

within genotypes. No statistically significant difference was observed. (p=0.2)

Expression of *Prdm16* and *Dscaml1* in the Brain

We also were interested in how biological expression levels could affect our results. Many of the brain genes of interest were expressed at low levels according to the Allen Brain Atlas and GXD, and we theorized that this could influence the data we observed. In order to shed more light on this possibility, two highly expressed brain genes (*Prdm16* and *Dscaml1*) were selected, and their primer-probe sets were ordered. Both of these genes had variable results even without the two samples that experienced RNA degradation. Without a larger sample size, it is hard to draw conclusions from these gene expression data.

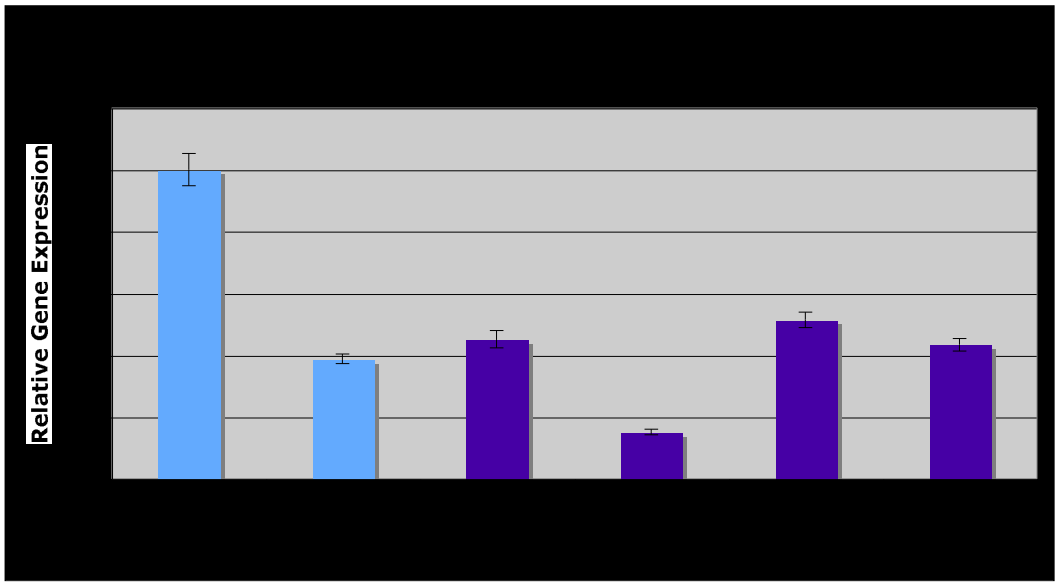

Figure 13: Relative expression of *Prmd16* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. This figure has data from one run using the low concentration actin primer-probe. The blue bars represent wild-type mice, and the purple bars represent knock-in mice. The error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample.

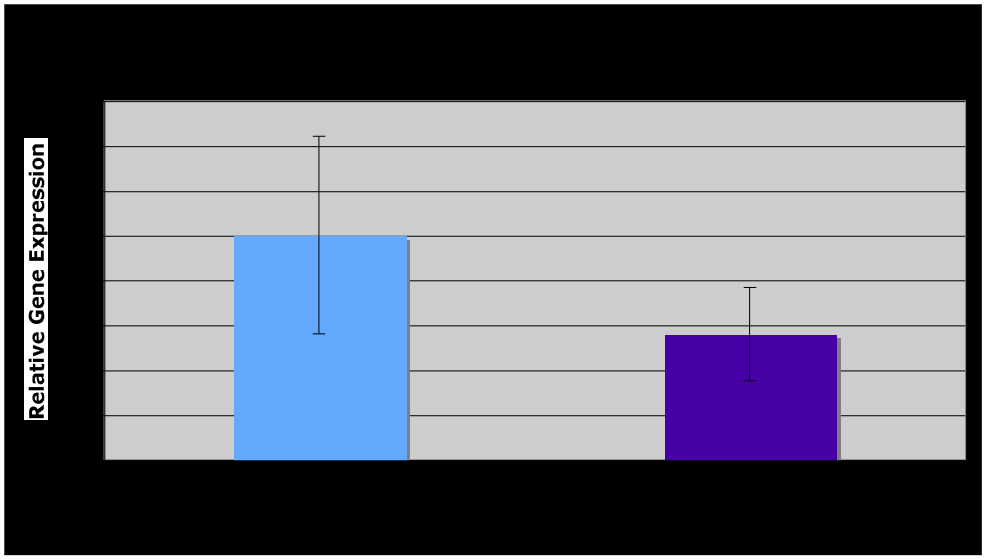

Figure 14: Relative expression of *Prdm16* between the wild-type and knock-in mice normalized to the average gene expression level of the wild-type mice. Error bars represent the standard error within genotypes. No statistically significant difference was observed. (p=0.70)

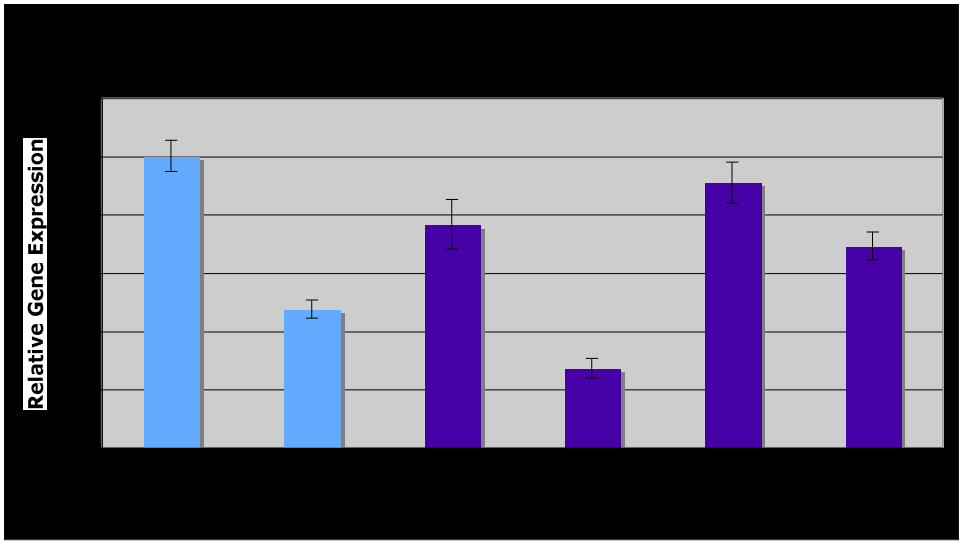

Figure 15: Relative expression of *Dscaml1* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. This figure has data from one run using the low concentration actin primer-probe. The blue bars represent wild-type mice, and the purple bars represent knock-in mice. The error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample.

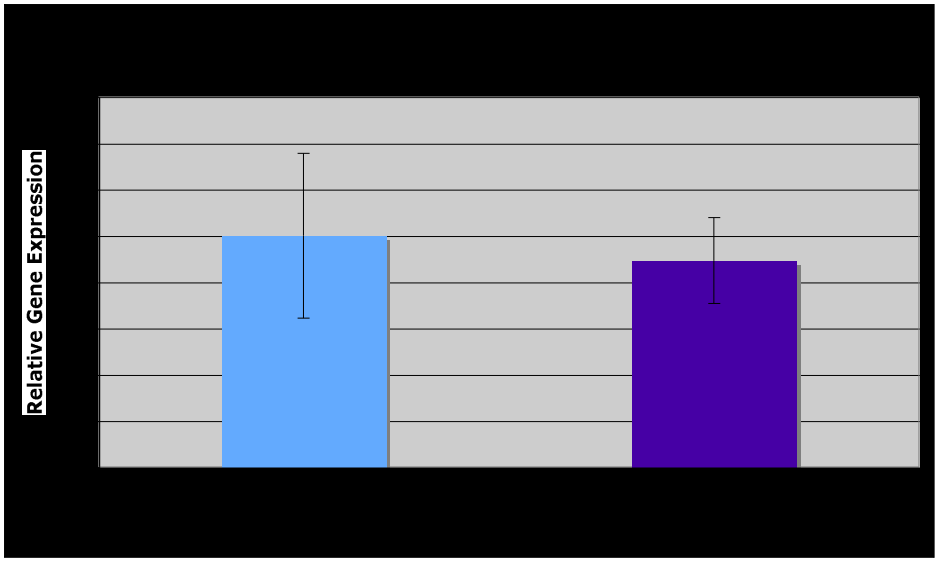

Figure 16: Relative expression of *Dscaml1* between the wild-type and knock-in mice normalized to the average gene expression level of the wild-type mice. Error bars represent the standard error within genotypes. No statistically significant difference was observed. (p=0.78)

Discussion

Goals and Expectations of Thesis Project

As part of a larger study into identifying phenotypes of the HdhQ111 strain of knock in mice, the German Mouse Clinic performed microarray analysis on 4 wild-type and 4 knock-in Huntington mice and found a number of differentially expressed genes. The goal of this project was to determine how robust the fold changes observed by the German Mouse Clinic were using a sensitive qRT-PCR approach, and then to explore whether or not these differentially expressed genes could present novel hypotheses of HD pathogenesis to be tested. Overall, we did not have enough samples to fully test the GMC finding, but our preliminary results do replicate the statistically significant downregulation of *Prok1* in knock-in mice. Due to time constraints, we were unable to look further into any possible role the gene may play in the disease process. However, this does not necessarily mean that the other genes studied in this project are unrelated to Huntington's disease.

While the goal of this project was to replicate the observations of the GMC, we did not expect to obtain exactly the same fold change results as they did. From an experimental standpoint, the GMC used a vastly different procedure than ours. Microarray analysis is excellent for assaying large numbers of genes at once, but does not achieve the same level of specificity as real-time PCR does, particularly the real-time PCR. The GMC also pooled their wild-type samples before hybridizing the cDNA to the microarray, which could mask variability in the wild-type mice. Moreover, these genes have alternative splice variants. It is possible that the microarray probe and our TaqMan assay did not pick up the same mRNA transcripts, which could also explain variations between the two sets of data. While most of the primers used should pick up on all the different splice variants for that gene, some primers do not (see Appendix Figures 10-18). From a sample standpoint, the mice used in the two studies were subtly different. The trip to Germany was undoubtedly a stressor for the mice. The mice were given time to acclimate and the genes studied may not have been affected by the change; however, whether the acclimation time was sufficient, and whether the expression of the genes was truly unaffected, is unknown. The mice in Germany also had different food, bedding, and generally disparate environments in comparison to their American counterparts. Lastly,

the wild-type and knock-in mice in the GMC study were from an earlier generation of the same background of mice as compared to the mice in this project, and thus were not as bred into the background as the mice used for our project. This, too, is unlikely to have caused any large changes. Still, the potential effects of these alterations should not be underestimated. Because of this, we sought to only test the robustness of the gene expression changes, and did not expect to replicate the fold changes or even the direction of the change.

Troubleshooting in this Project

The major obstacle in this project was not that we only replicated the results of one of the seven genes assayed, but rather that we saw variations in the results from the eight males. While this variability may be indicative of the actual biological expression levels, we considered other possible explanations first. We attempted to rule as many out as possible, though some measures were impractical, given the time constraints. First, we saw a large difference between the female and the male mice in the original set of mice. Our results of the original six male mice show a statistically significant difference in the *Prok1* gene. It was therefore hoped that obtaining two more male mice would help improve the significance of the data. Upon acquiring the two new males and assaying all eight male mice, it became apparent that gender was not the only issue. When we checked the quality of the RNA we had extracted, we found that the RNA taken from the brain tissue of the two new wild-type males had been degraded. Therefore, we were unable to use the data obtained for these mice. Additionally, not all genes could be optimized. *Pomc, Prok1, Nfkbia, Prdm16*, and *Dscaml1* were all optimized, whereas *Prl, Ace,* and *Shh* were not. (See Tables 2 and 3). As we did not discover the RNA degradation until the end of this project, we made other attempts to find the source of variation, such as selecting new genes, assaying for genomic DNA contamination, and changing the primer concentration.

As three of the six genes assayed at this point could not be optimized, their typical expression level became a concern. *Prl, Pomc,* and *Prok1* were all expressed at low levels in the brain according to the Allen Brain Atlas and the GXD. We hypothesized that the low gene expression levels might be causing the variability, as any slight changes in the

amount of target detected would drastically alter the ratio of one gene to another across samples and is otherwise subject to experimental artifact. In order to elucidate whether or not this could be affecting our results, we selected two highly expressed brain genes that the GMC found to be differentially expressed: *Prmd16* and *Dscaml1*. We also wanted to see whether or not using the primer-limited actin probe set could influence our results. Therefore, the optimization runs and full runs were re-done with the new actin primer-probe sets for both *Prdm16* and *Dscaml1* and the six original genes. The results showed that both of these measures were unsuccessful. Even with the new primer, *Prl*, *Ace*, and *Shh* could not be optimized and the variability between samples was not decreased to any substantial degree. *Prdm16* and *Dscaml1* also had variable results with no clear pattern, though as the sample size was only 6, it is difficult to judge the extent of variability. The fact that even highly expressed genes showed variation in expression level between samples indicates that biological gene expression level is not the cause of differences between samples, though it does not preclude it from contributing to it.

Another concern was the presence of genomic DNA in the samples. Two measures were already in place to prevent any DNA from affecting the results. First, the assay was designed to amplify only the cDNA version of the target of interest, as the primer-probe sets span large introns. Also, the samples were all DNase treated before undergoing reverse transcription. To see whether or not these two features were sufficient to prevent genomic DNA contamination, we ran two –RT controls; one that was a no reverse transcriptase well, and one that was a no DNase or reverse transcriptase well. There was no amplification in the DNase well. In the well that was not DNase treated, we saw no amplification of the target and only very late (40^{th} cycle) amplification of the reference. These data indicate that both the DNase treatment and the primer-probe set design were successful.

Though the assay itself did not seem to be the source of the variability, the plate or sample preparation could. One of the precautions in place was the fact that target gene expression levels were normalized to the actin levels in the same well. This meant that if a pipetting error occurred that resulted in a lower concentration of cDNA in one well, both the actin and the target would have a later call time. This minimizes the effect of a plate preparation error. Originally, we believed that we did not have a great deal of

sample preparation error. This was because the internal control call times were fairly consistent across genes and samples.

However, when we ran a gel, we found brain RNA degradation of the two new samples (M1285 and M1106). This explained why these two samples had very low expression levels of some genes (See Appendix Figure 6 for example of expression level data with these two mice), and accounted for a large amount of the variation we had observed in the wild-type brain samples. This decreased our sample size and made it very difficult to evaluate certain brain genes. For instance, the two wild-type samples remaining had very different expression levels of *Prdm16* and *Dscaml1*. As the sample size is only two, though, it's hard to classify this as variability or draw conclusions from the results.

The RNA from all other samples was intact, though we would have liked to see better quality RNA with stronger and more consistent banding patterns in the liver. Some samples did have particularly strong or weak bands. However, this does not seem to correlate with the expression data we observed. For instance, M1285 and L1775 had a much more distinct banding pattern that many of the other liver samples. These two samples did not have the highest levels of expression of *Nfkbia* in their respective groups. M1106 had the faintest banding pattern on the gel, but had an expression level similar to M1285 (See Figures 5 and 11). The stock solution of L1777 had a weak 28s band, which raises the question of whether it has been contaminated with RNase. All cDNA was made from the dilution though, which shows no degradation. Still, this highlights the need for further precautions in future experiments, such as more careful RNA extractions and running quality control gels before data collection.

Pomc, Prok1, Prdm16,* and *Dscaml1

These four brain genes assayed were the focus of the project. The genes found to be differentially expressed in the liver were also of interest, but given the nature of HD, these four were of a higher priority than the liver genes when time limits were an issue. Because of this and time constraints we opted to focus on the *Pomc, Prok1, Prdm16*, and *Dscaml1*. At least two assays were run with each sample (only one run is shown in each Figure). We observed inter-run consistency but inter-sample variability. Also, with the

loss of the data from the two newer wild-type males and the resulting decrease in sample size, it is even more difficult to draw any concrete conclusions from the data. However, some genes had promising preliminary results.

Pomc amplified well and gave similar results in all runs. However, the relative expression levels between samples were always quite variable. There was no discernable pattern in the expression level of *Pomc*. Because of the amount of variability seen in *Pomc*, this would not be one of the genes we explored further right away. It would take a much larger sample size to see if there was any change between wild-type and knock-in, and our data do not indicate there would be a difference.

Prok1 had the most promising preliminary results of the genes we investigated. When the data from the original four knock-in and two wild-type male mice were compared, there was a statistically significant downregulation of *Prok1* in the knock-ins as compared to the wild-type mice. Due to the sample size, this significance may not be robust. However, because *Prok1* had a preliminary differential expression that was statistically significant, and because our data agrees with the findings of the GMC, this gene is extremely promising and would be one to further investigate.

Nfkbia

Though the liver genes were not the focus of this experiment due to time constraints, they were important nevertheless, and our findings were interesting. *Nfkbia* had very reliable results, as seen by the small normalized error between triplicates (See Figure 8). It was also slightly less variable compared to the other assays and had the second lowest p value in our data. As of this time, it has not been run with the primer-limited actin control, though switching to the low concentration primer has not proved to alter results very much in the other genes. We visualized the 28s and 18s ribosomal bands on the gel for all samples, but M1285 and M1106 did have lower expression levels of *Nfkbia* than the two original wild-type mice. As the mice were born from different litters and were sacrificed at different times, and as the RNA extraction and cDNA synthesis took place separately, it is impossible to say from the sample size we have whether this variability seen in the wild-type samples is biological or due to a difference in sample preparation. With this in mind, it would be best to start any future studies with a set of mice that were sacrificed at the same time The GMC found *Nfkbia* to be upregulated in

the mutant mice, whereas we have found it to be slightly, but not significantly, downregulated. However, given the small sample size, this should not disqualify *Nfkbia* from further testing. Additionally, gene expression data from immortalized mouse cell lines have show a statistically significant differential expression of *Nfkbia* in HD cells as compared to normal cells. Therefore, this is another one of the more promising genes and one of the best suited for future studies.

Prl, Ace,* and *Shh

Prl, Ace, and *Shh* were not optimized in these assays. The serial dilution we used included 1, 5, 10, 50, and 100 μg of cDNA per well. With the concentration of cDNA we synthesized, the dilution could go up to 200 μg per well for any of these three genes. We opted to focus on the genes that were optimized instead of finding the ideal amplification concentration of *Prl, Ace,* or *Shh*. As *Prl* is a hormone, stringent controls could be necessarily to help differentiate between hormonal fluctuations and gene expression differences that relate to Huntington's disease. For instance, sacrifice may have to take place at a certain time of day or the mice would have to be awake for a certain amount of time before death; it may be difficult to achieve all of these controls. It would be interesting to see if either *Ace* or *Shh* is differentially expressed in the knock-in mouse because of their previous association with Alzheimer's disease and possible links to HD, respectively.

Overall and Future Studies

Ultimately, we found one statistically significant differential expression: The downregulation of *Prok1* in the brain of knock-in mice. However, our small sample size precluded us from coming to anything but preliminary conclusions. We did see variability between samples, which could stem from several assay or sample related factors, or could be indicative of the actual gene expression levels in each mouse. We attempted to rule out as many factors as we could in the time that we had, and found that RNA degradation had caused a majority of the variability. It is important to note that certain aspects, such as remaining degree of variability, were difficult to weigh because the sample size was so small. With only two or four mice in a category, one cannot conclude which samples are

outliers and which are indicative of the normal wild-type or knock-in population. Sacrificing a new group of mice would be the best way to better tease apart the complicating factors from the root cause (or causes) of any variability that might be encountered. This group of mice should all be sacrificed at the same time and as closely related as possible. While it would be interesting to re-assay all of the genes, not all would be as worthwhile. Because of the huge variability, *Pomc* would not be a good candidate for further study without a much larger sample size. *Prl* would be interesting to study, but may not be as relevant or easy to investigate as the other genes. Therefore, these would likely not be the first choice for future studies. *Prdm16* and *Dscaml1* could similarly be intriguing with a larger sample size, for both their high expression level and their roles in energy homeostasis and involvement in neuronal networks, respectively. *Ace* and *Shh* were not optimized, but may be of interest as a secondary focus in future studies because of their associations with other neurological disorders and potential link to the mutant huntingtin protein.

Of all the genes assayed, *Prok1* and *Nfkbia* are currently the most promising genes that we investigated in this study. They had fairly low levels of inter-sample variability as seen by their p values (See Figures 10 and 12). There are also studies in the literature that indicate that *Prok1* and *Nfkbia* may also play a role in Huntington's disease. *Prok1* has been shown to contribute to the regulation and activation of the NFAT/calcineurin pathway. Among other activities, the NFAT/calcineurin pathway activates various inflammatory molecules, including interleukin-11 (IL-11), a member of the IL-6 family [40]. Therefore, a decrease in *Prok1* expression may result in decreased IL-11 levels. IL-11 has not been shown to be up or downregulated in HD patients; however, this may be because it has not yet been assayed, as it has not been mentioned in the literature. IL-11 has a variety of functions, and a study of Alzheimer's disease patients showed that they had increased IL-11 levels in cerebrospinal fluid. However, it is postulated that this increase is in fact a neuroprotective measure, as IL-11 has been implicated in neuronal survival [41]. It is worth nothing that inflammatory pathways seem to play a larger role in AD than HD [34]. However, this is still an intriguing connection, particularly as our preliminary results and GMC results both found *Prok1* to be significantly downregulated in knock-in mice as compared to wild-type mice.

Nfkbia also appears to have a potential link to Huntington's disease through an inflammatory pathway. The protein product of this gene inhibits NF-κB, which is involved in anti-apoptotic and pro-inflammatory pathways. Typically, NF-κB is sequestered in the nucleus by inhibitors, including *Nfkbia*. Degradation of these inhibitors releases NF-κB, which is transported into the nucleus and acts as a transcription factor [42]. NF-κB is also controlled by negative regulation, as it activates the transcription of *Nfkbia* [43]. Interestingly, activation of NF-κB has been linked to Huntington's disease. It has been shown that mutant huntingtin directly interacts with positive regulators of NF-κB, and elevated NF-κB have been seen in striatal cells from HD transgenic mice as well as cell lines expressing mutant huntingtin. Reducing levels of NF-κB also decreased the HD-mediated toxicity in these cells [44]. Our data does not fit perfectly into this scheme. The GMC data and the expression levels seen in different cell lines both show that *Nfkbia* is upregulated in the HD-state, whereas our data indicates it is downregulated. However, this conflict does not preclude *Nfkbia* from being studied further. The fact that this gene is differentially expressed is interesting regardless of the direction of the change in expression. An upregulation of *Nfkbia* could be a response to the increased NF-κB levels or a compensatory attempt by the cell to control this increase. A downregulation, on the other hand, could contribute to the increased NF-κB levels and/or represent a defect in the negative regulation of NF-κB. Both the GMC data and our data are from liver tissue instead of the brain. Whether or not this change in expression in liver tissue is relevant to HD remains to be seen. It is possible that there is differential expression of *Nfkbia* in particular brain regions, which may not have shown up on an assay of the whole brain. It would be interesting to delve deeper into this gene to tease apart whether it is a side effect or causative effect of Huntington's disease, and whether its expression in the liver plays a role in HD.

Overall, *Prok1* and *Nfkbia* both are intriguing subjects for further study. They may yet offer an insight into Huntington's disease or perhaps even be a target for drug therapy. Using a new set of mice, more cautious preparation, and perhaps a larger sample size, these genes would be well-suited for further exploration.

Conclusion

In this study, we attempted to replicate the differential gene expression findings of a study done by the German Mouse Clinic in the HdhQ111 knock-in mouse model of Huntington's disease. We assayed the expression levels of five genes that the GMC had found to be differentially expressed in the brain, and three gene found to be differentially expressed in the liver. We found that one gene, *Prok1*, was significantly downregulated in the knock-in mice as compared to the wild-type mice. Promising results were also seen for *Nfkbia*. Our sample size was decreased due to RNA degradation, and our results may not be robust. However, as our *Prok1* results agree with the GMC data, the results are very encouraging. We recommend that further studies be done with a new set of mice, and that particular attention be paid to *Prok1* and *Nfkbia*.

Bibliography

[1] Martin, JB, and JF Gusella. "Huntington's disease. Pathogenesis and management." *New England Journal of Medicine* 315 (Nov. 1986): 1267-1276.

[2] HD Consortium. "A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes." *Cell* 72.6 (1993): 971-983.

[3] Roos, Raymond AC. "Huntington's disease: a clinical review." *Orphanet Journal of Rare Diseases* 5 (2010): 40.

[4] Huntington G. "On chorea." *Medical and Surgical Reporter of Philadelphia* 26 (1872): 317-321.

[5] Walker, Francis O. "Huntington's disease." *Lancet* 369.9557 (2007): 218-228.

[6] Ross, Christopher A, and Sarah J Tabrizi. "Huntington's disease: from molecular pathogenesis to clinical treatment." *The Lancet Neurology* 10.1 (2011): 23-98.

[7] Ferrante, Rober J. "Mouse models of Huntington's disease and methodological considerations for therapeutic trials." *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1792.6 (2009): 506-520.

[8] Truant, Ray, Randy Singh Atwal, and Anjee Burtnik. "Nucleocytoplasmic trafficking and transcription effects of huntingtin in Huntington's disease." *Progress in Neurobiology* 83.4 (2007): 211-227.

[9] Hickey, Miriam A, and Marie-Françoise Chesselet. "Apoptosis in Huntington's disease." *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 27.2 (2003): 255-265.

[10] Heng, Mary Y, Peter J Detloff, and Roger L Albin. "Rodent genetic models of Huntington disease." *Neurobiology of Disease* 32.1 (2008): 1-9.

[11] Menalled, Liliana B. “Knock-In Mouse Models of Huntington’s Disease.” *NeuroRX* 2.3 (2005): 465-470.

[12] Dragileva, Ella, et al. “Intergenerational and striatal CAG repeat instability in Huntington’s disease knock-in mice involve different DNA repair genes.” *Neurobiology of Disease* 33.1 (2009): 37-47.

[13] Hayden, M R, et al. “Impaired ProlactinRelease in Huntington’s Chorea : Evidence For Dopaminergic Excess.” *The Lancet* 310.8035 (1977): 423-426. *ScienceDirect*. Web.

[14] Müller, Eugenio E, et al. “Prolactin Control in Huntington’s Chorea.” *The Lancet* 310.8041 (1977): 764-765.

[15] Kremer, H.P.H, et al. “Endocrine functions in Huntington’s disease: A two-and-a-half years follow-up study.” *Journal of the Neurological Studies* 90.3 (1989): 335-344.

[17] Butler, A. "The Melanocortin System and Energy Balance." *Peptides* 27.2 (2006): 281-90.

[18] Ngan, E., and P. Tam. "Prokineticin-signaling Pathway." *The International Journal of Biochemistry & Cell Biology* 40.9 (2008): 1679-684.

[19] Seale, Patrick, et al. "Transcriptional Control of Brown Fat Determination by PRDM16." *Cell Metabolism* 6.1 (2007): 38-54.

[20] Bournat, JC, and CW Brown. "Mitochondrial Dysfunction in Obesity." *Current Opinion in Endocrinology, Diabetes and Obes* 17.5 (2010): 446-52.

[21] Kinameri, Emi, et al. "Prdm Proto-Oncogene Transcription Factor Family Expression and Interaction with the Notch-Hes Pathway in Mouse Neurogenesis." Ed. Michael Hendricks. *PLoS ONE* 3.12 (2008): E3859.

[22] Agarwala KL, et al. "Cloning and Functional Characterization of DSCAML1, a Novel DSCAM-like Cell Adhesion Molecule That Mediates Homophilic Intercellular Adhesion." *Biochemical and Biophysical Research Communications* 285.3 (2001): 760-72.

[23] Wu, Chia-Lin, et al. "Erythropoietin and sonic hedgehog mediate the neuroprotective effects of brain-derived neurotrophic factor against mitochondrial inhibition." *Neurobiology of Disease* 40.1 (2010): 146-154.

[24] Wu, Chia-Lin, et al. "Sonic Hedgehog Mediates BDNF-induced Neuroprotection against Mitochondrial Inhibitor 3-nitropropionic Acid." *Biochemical and Biophysical Research Communications* 385.1 (2009): 112-17.

[25] Chen, F, et al. "New insights into the role of nuclear factor-κB, a ubiquitous transcription factor in the initiation of diseases." *Clinical Chemistry* 45.1 (1999): 7-17.

[26] Edwards, Todd L., et al. "An Association Analysis of Alzheimer Disease Candidate Genes Detects an Ancestral Risk Haplotype Clade In *ACE* and Putative Multilocus Association Between *ACE, A2M* and *LRRTM3*." *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 150B.5 (2009): 721-35.

[27] Stewart, Juhani Akseli, et al"ACE Polymorphism and Response to Electroconvulsive Therapy in Major Depression." *Neuroscience Letters* 458.3 (2009): 122-25.

[28] Mochel, Fanny, and Ronald G Haller. "Energy deficit in Huntington disease: why it matters." *Journal of Clinical Investigation* 121.2 (2011): 493-499.

[29] Seong, IS, et al. "Huntingtin facilitates polycomb repressive complex 2." *Human Molecular Genetics* 19.4 (2010): 573-583.

[30] Voisine, Cindy, et al. "Identification of Potential Therapeutic Drugs for Huntington's Disease Using Caenorhabditis Elegans." Ed. Xiao-Jiang Li. *PLoS ONE* 2.6 (2007): E504.

[31] Woo-Kyun, Kim, et al "Negative Regulation of Hedgehog Signaling by Liver X Receptors."*Molecular Endocrinology* 23.10 (2009): 1532-543.

[32] Futter, M., et al. "Wild-type but Not Mutant Huntingtin Modulates the Transcriptional Activity of Liver X Receptors." *Journal of Medical Genetics* 46.7 (2009): 438-46.

[33] Pfaffl, Michael. "Relative Quantification." *Real Time Pcr.* Ed. T. Dorak. International University Line, 2006. 63-82.

[34] Schwab, Claudia, Andis Klegeris, and Patrick L. McGeer. "Inflammation in Transgenic Mouse Models of Neurodegenerative Disorders." *Biochimica Et Biophysica Acta (BBA) - Molecular Basis of Disease* 1802.10 (2010): 889-902.

[35] Kobal, J., et al "Autonomic Dysfunction in Presymptomatic and Early Symptomatic Huntington's Disease." *Acta Neurologica Scandinavica* 121.6 (2010): 392-99.

[36] Mochel, Fanny, et al "Early Energy Deficit in Huntington Disease: Identification of a Plasma Biomarker Traceable during Disease Progression." Ed. Ulrich Mueller. *PLoS ONE* 2.7 (2007): E647.

[37] Ciarmiello, A. "Brain White-matter Volume Loss and Glucose Hypometabolism Precede the Clinical Symptoms of Huntington's Disease." *Journal of Nuclear Medicine* 47.2 (2006): 215-22.

[38] DiProspero, Nicholas A., et al "Early Changes in Huntington's Disease Patient Brains Involve Alterations in Cytoskeletal and Synaptic Elements." *Journal of Neurocytology* 33.5 (2004): 517-33.

[39] Bjorkqvist, M., et al. "A Novel Pathogenic Pathway of Immune Activation Detectable before Clinical Onset in Huntington's Disease." *Journal of Experimental Medicine* 205.8 (2008): 1869-877.

[40] Cook, Ian H. "Prokineticin-1 (PROK1) Modulates Interleukin (IL)-11 Expression via Prokineticin Receptor 1 (PROKR1) and the Calcineurin/NFAT Signalling Pathway."*Molecular Human Reproduction* 16.3 (2010): 158-69.

[41] Galimberti, D., et al "Intrathecal Levels of IL-6, IL-11 and LIF in Alzheimer's Disease and Frontotemporal Lobar Degeneration."*Journal of Neurology* 255.4 (2008): 539-44.

[42] Schattenberg, JM, M. Schuchmann, and PR Galle. "Cell Death and Hepatocarcinogenesis: Dysregulation of Apoptosis Signaling Pathways." *Journal of Gastroenterology and Hepatology* 26 (2011): 213-19.

[43] Nelson, D. E., et al "Oscillations in NF- B Signaling Control the Dynamics of Gene Expression."*Science* 306.5696 (2004): 704-08.

[44] Khoshnan, Ali, et al "Activation of the I B Kinase Complex and Nuclear Factor- B Contributes to Mutant Huntingtin Neurotoxicity." *The Journal of Neuroscience* 24.37 (2004): 7999-8008.

Appendix

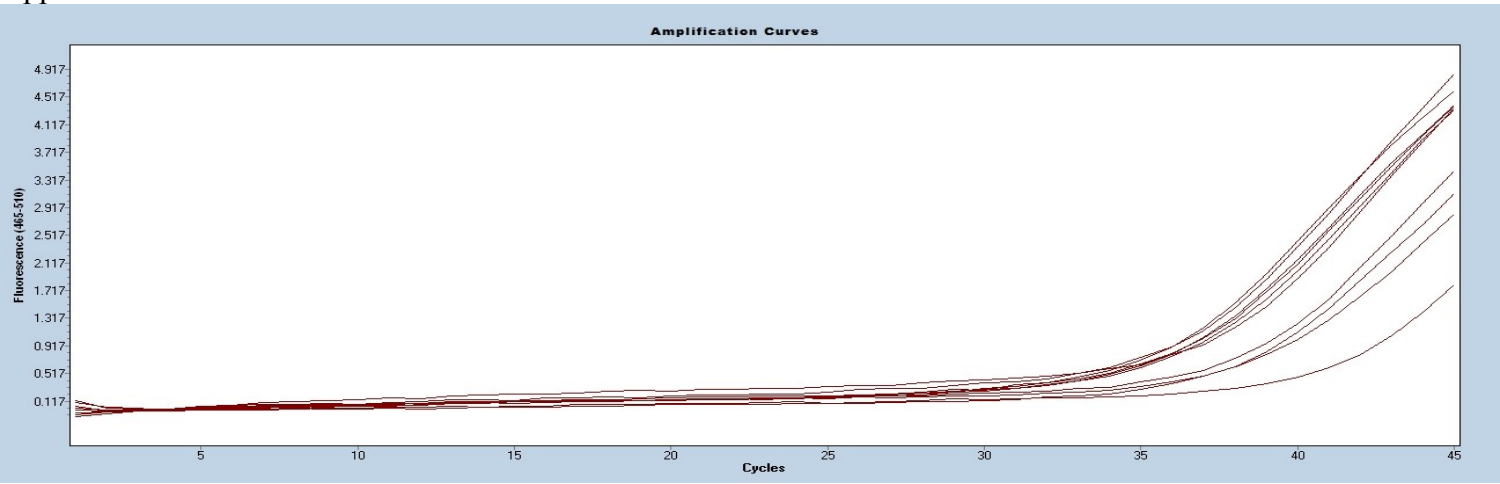

Figure 1: A standard curve that was unsuccessful. Compare with the standard curve of the internal controls from the same well

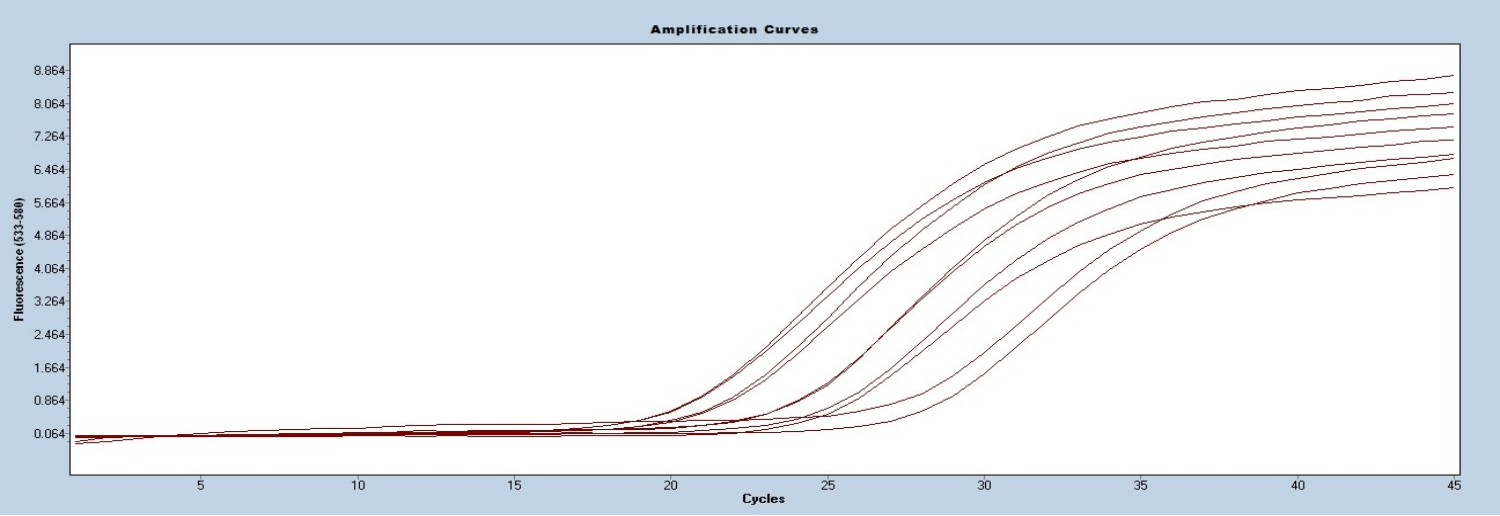

Figure 2: The standard curve from the internal controls of Figure 1. You can see how late the target begins to amplify as compared to the control

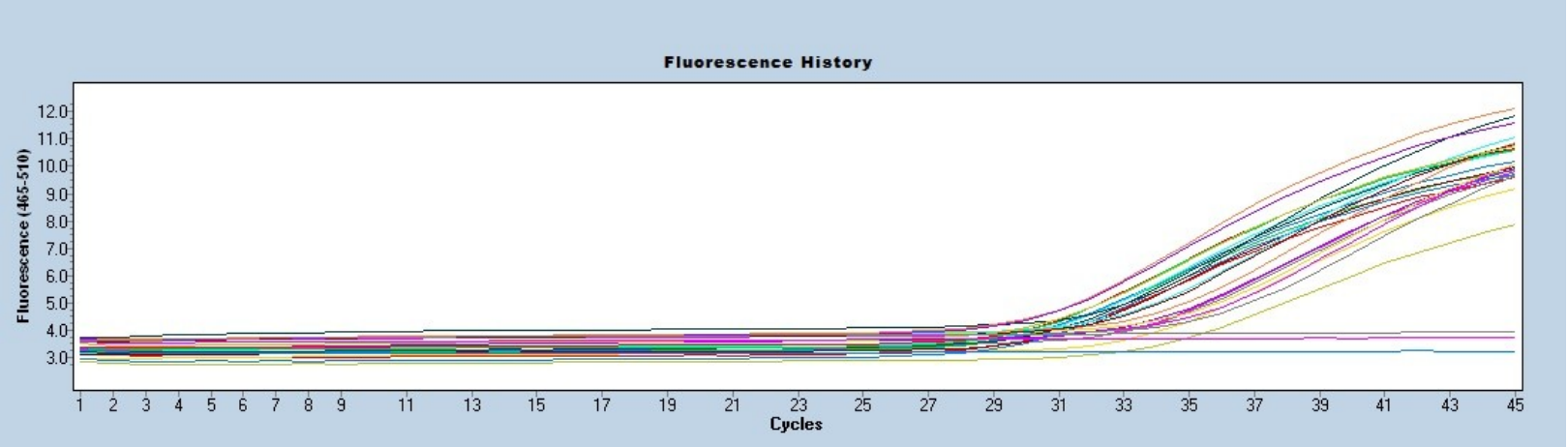

Figure 3: The –RT control for the *Pomc* target. 2 different –RT controls were run. One was DNase treated but did not undergo reverse transcription. The other was not DNase treated and did not undergo reverse transcription. The three flat lines are the water blank and the two –RT controls.

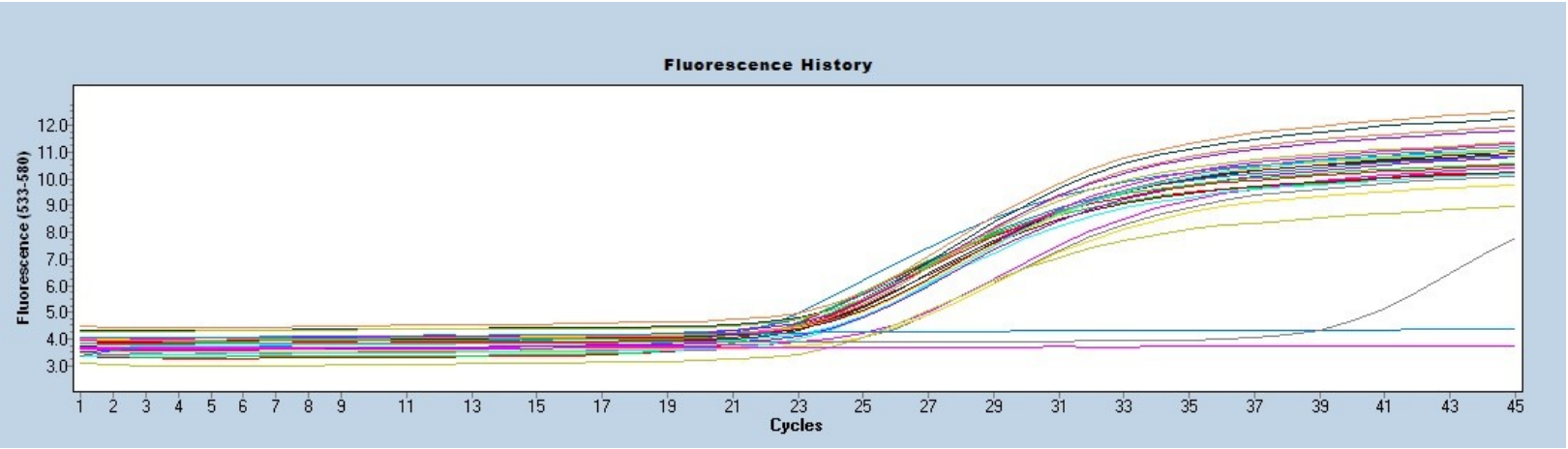

Figure 4: The –RT control for the actin internal control of the same *Pomc* run. The two flat lines are the water blank and the –RT control that was DNase treated. The curve that begins to amplify late is the sample that was not DNase treated.

Mouse ID	Sex	Genotype & Repeat Count	Age at Death	Date of Death	Used for Brain Gene Assays	Used for Liver Gene Assays
L1765	F	Wildtype	17.5 weeks	10/20/2010	Yes*	No
L1773	F	Wildtype	17.5 weeks	10/20/2010	Yes*	No
L1775	M	Knock-In; 123	17.5 weeks	10/20/2010	Yes	Yes
L1776	M	Wildtype	17.5 weeks	10/20/2010	Yes	Yes
L1777	M	Wildtype	17.5 weeks	10/20/2010	Yes	Yes
L1778	M	Knock-In; 126	17.5 weeks	10/20/2010	Yes	Yes
L1784	M	Knock-In; 113	17.5 weeks	10/20/2010	Yes	Yes
L1785	M	Knock-In; 130	17.5 weeks	10/20/2010	Yes	Yes
M1285	M	Wildtype	18 weeks	12/13/2010	Yes*	Yes
M1106	M	Wildtype	20 weeks	12/13/2010	Yes*	Yes

Figure 5: A list of the mice used in different stages of the experiment. M1285 and M1106 had RNA that was degraded, but this was not discovered until after experiments had been run. * Indicates data not shown.

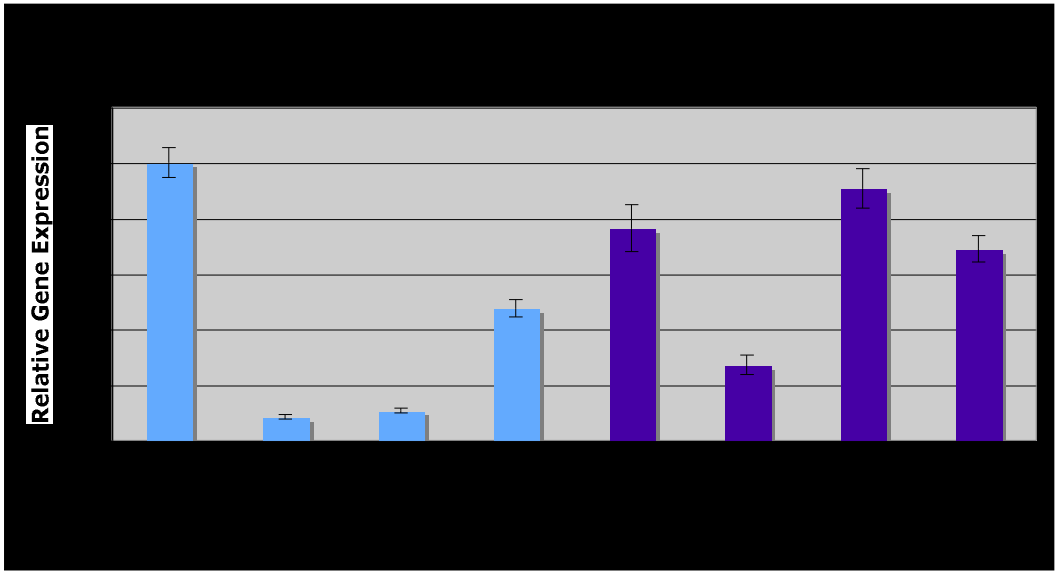

Figure 6: Relative expression of *Dscaml1* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. The low concentration actin primer was used. The error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample. This graph includes the two samples that were found to have RNA degradation (M1285 and M1106)

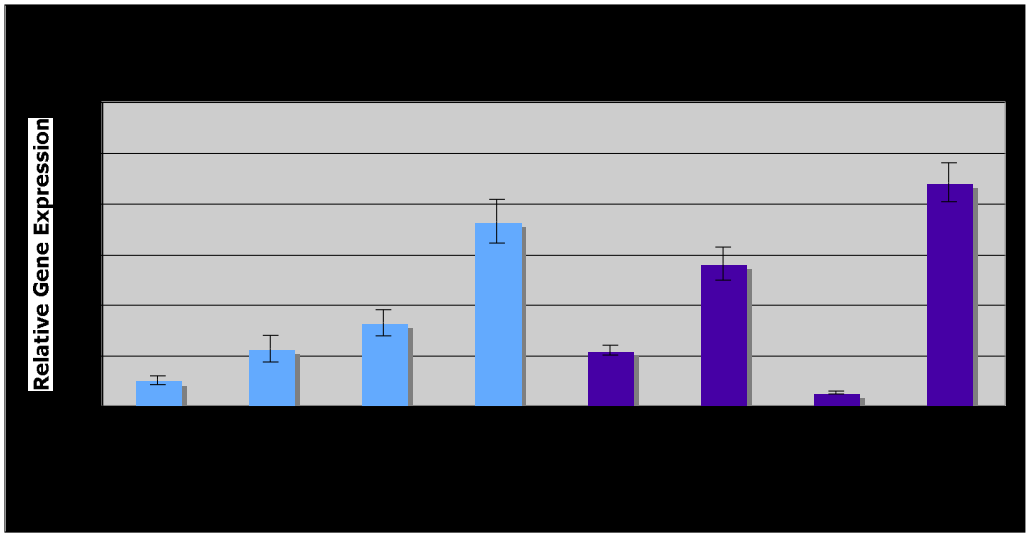

Figure 7: Relative expression of *Pomc* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. The low concentration actin primer was used. The error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample. This graph includes the two samples that were found to have RNA degradation (M1285 and M1106)

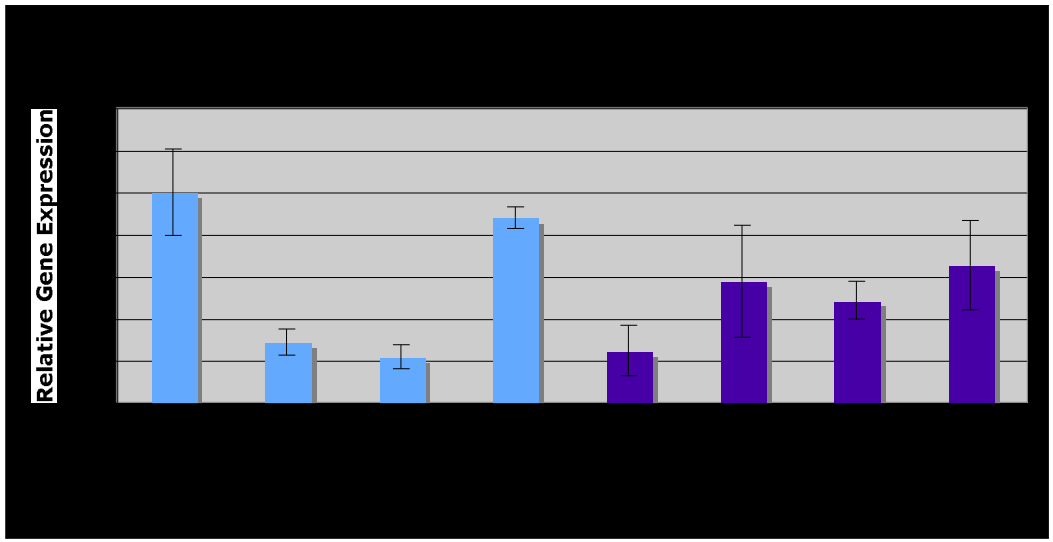

Figure 8: Relative expression of *Prok1* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. The low concentration actin primer was used. The error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample. This graph includes the two samples that were found to have RNA degradation (M1285 and M1106)

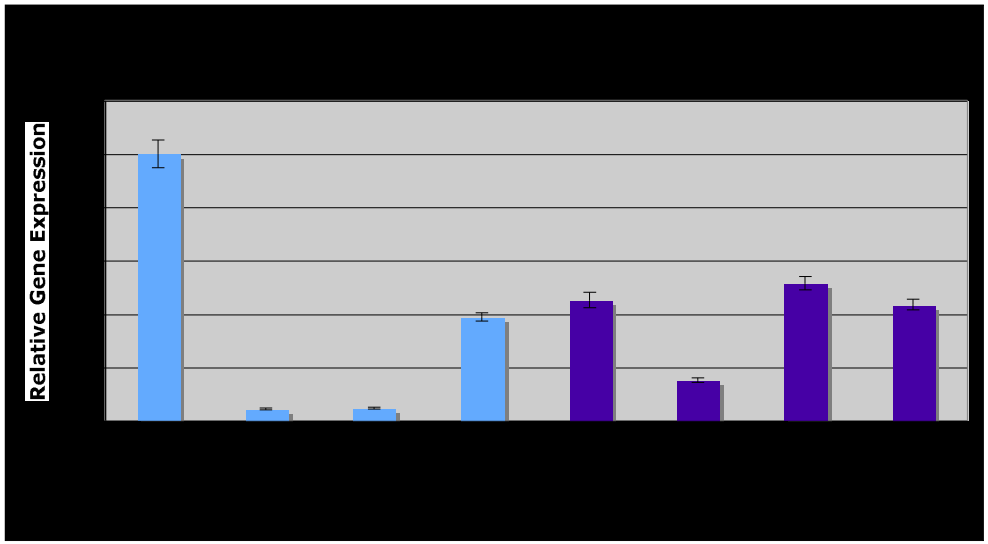

Figure 9: Relative expression of *Prdm16* by mouse normalized to one wild-type sample (L1776) after being normalized to the internal actin control level. The low concentration actin primer was used. The error bars are the normalized standard error between the call times ratio of target to reference of the triplicate wells of each sample. This graph includes the two samples that were found to have RNA degradation (M1285 and M1106)

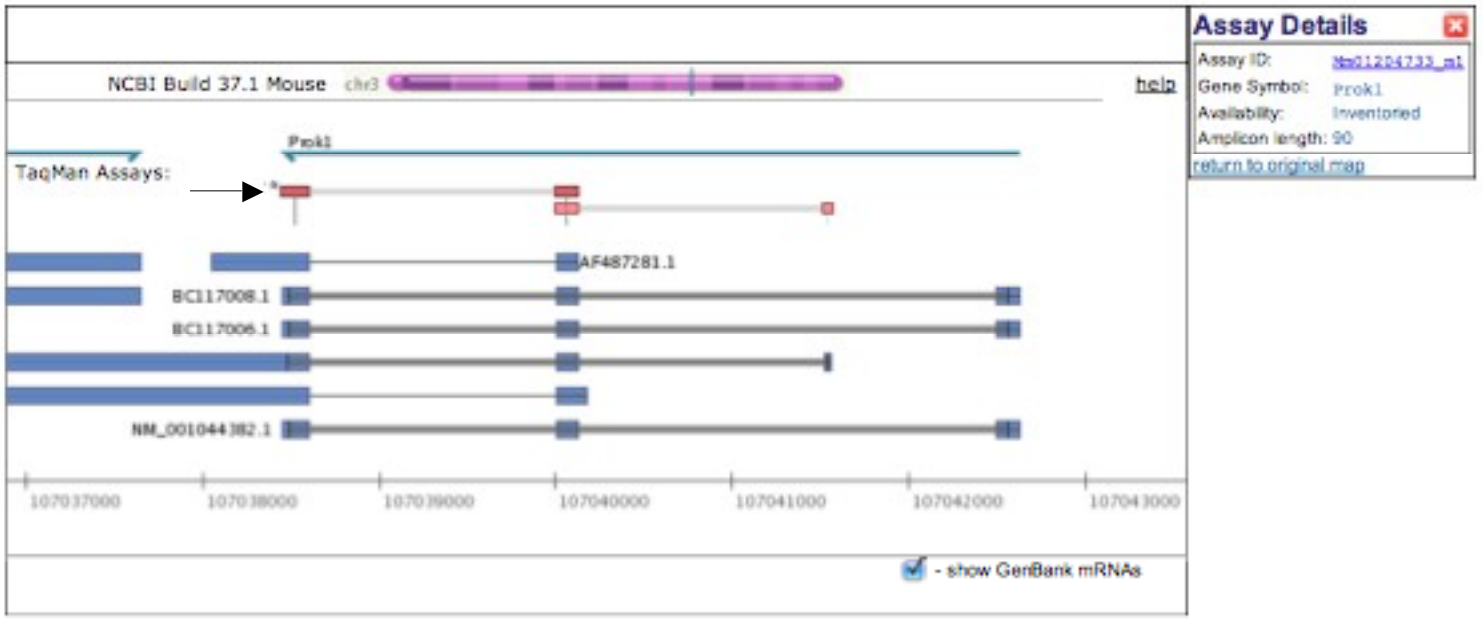

Figure 10: Alignment map of *Prok1*. Arrow represents the primer probe set we used; the blue bars are different splice variants, where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website

Figure 11: Alignment map of *Prl*. Arrow represents the primer probe set we used; the blue bars are different splice variants, where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website

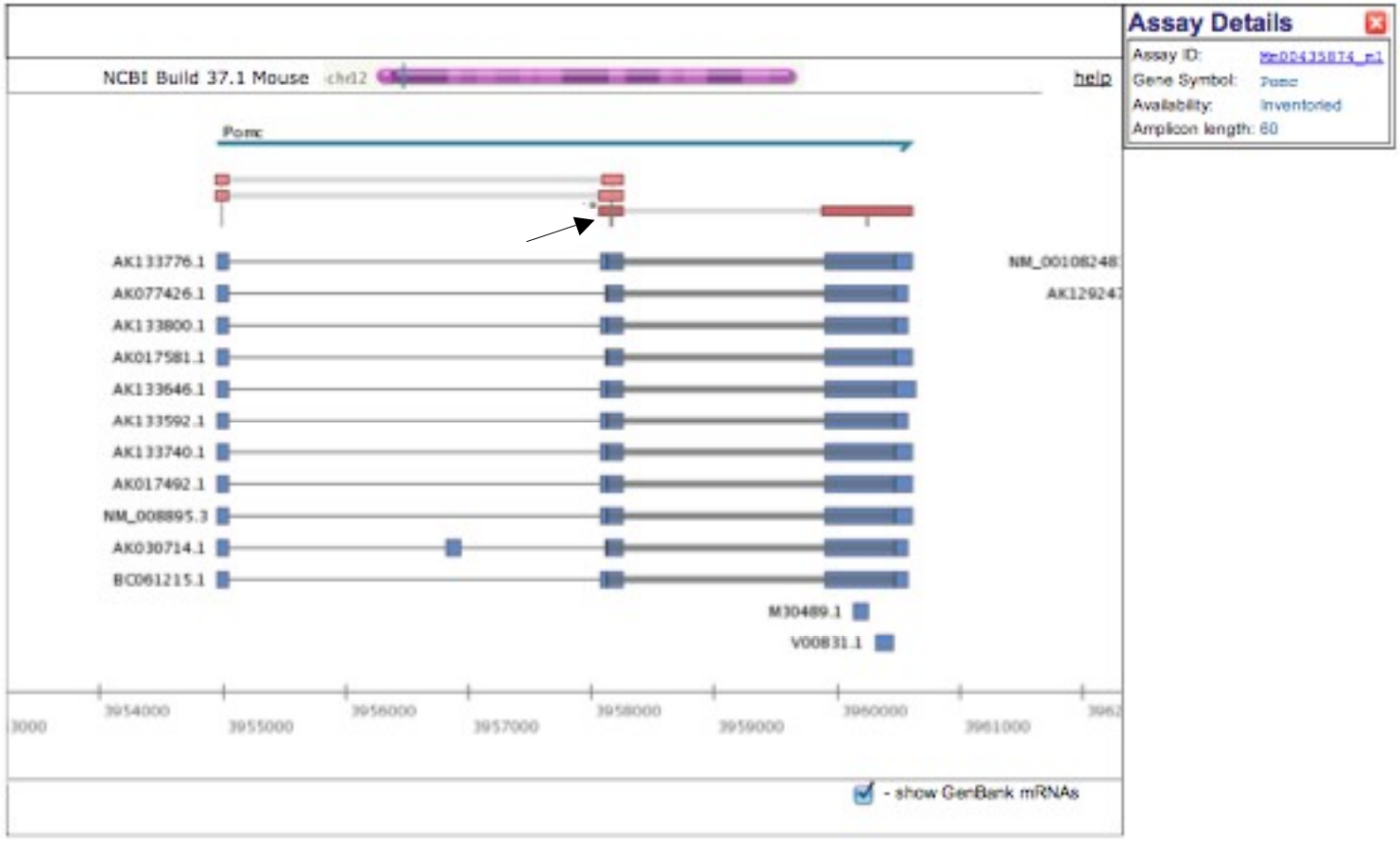

Figure 12: Alignment map of *Pomc*. Arrow represents the primer probe set we used; the blue bars are different splice variants, where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website

Figure 13: Alignment map of *Nfkbia*. Arrow represents the primer probe set we used; the blue bars are different splice variants, where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website

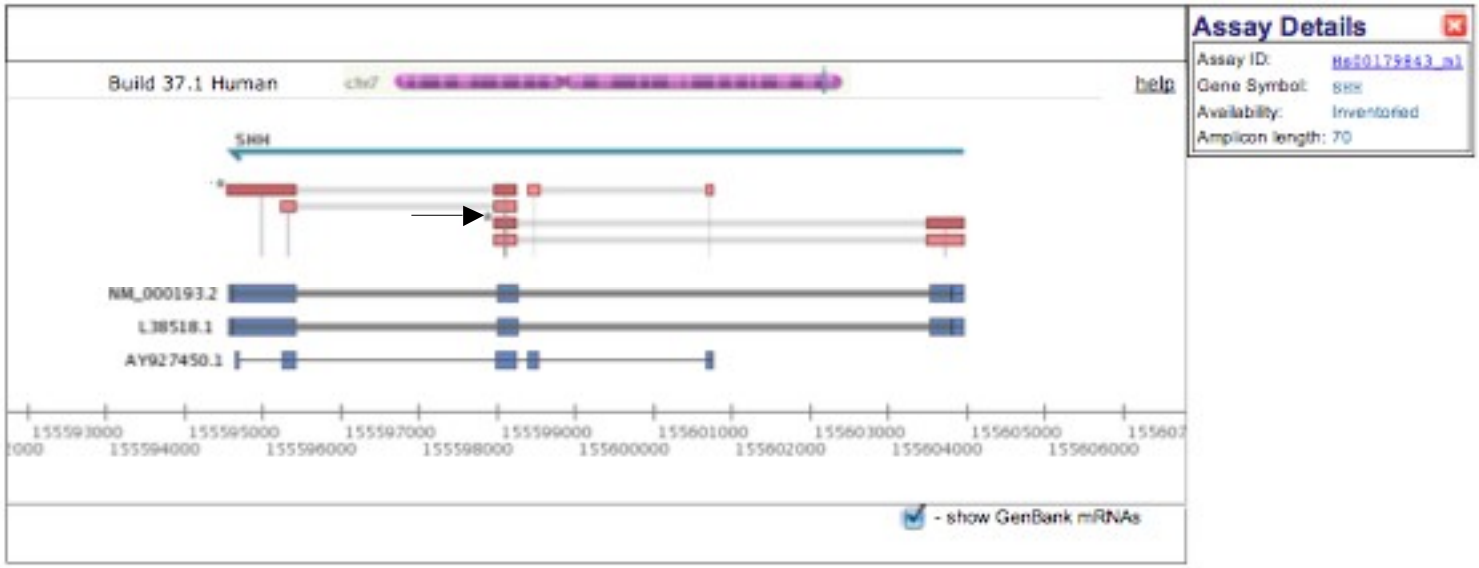

Figure 14: Alignment map of *Shh*. Arrow represents the primer probe set we used; the blue bars are different splice variants, where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website

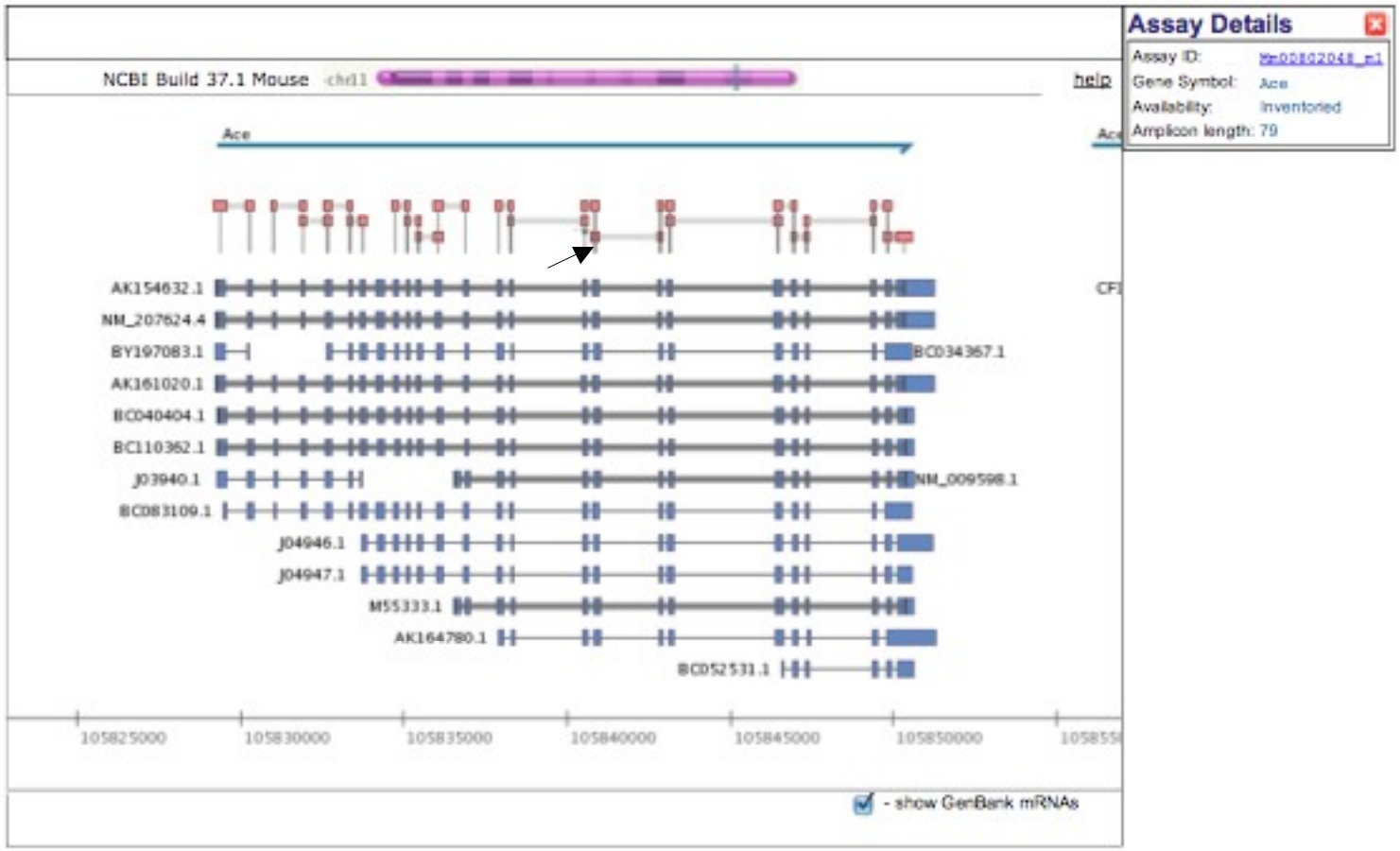

Figure 15: Alignment map of *Ace.* Arrow represents the primer probe set we used; the blue bars are different splice variants, where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website

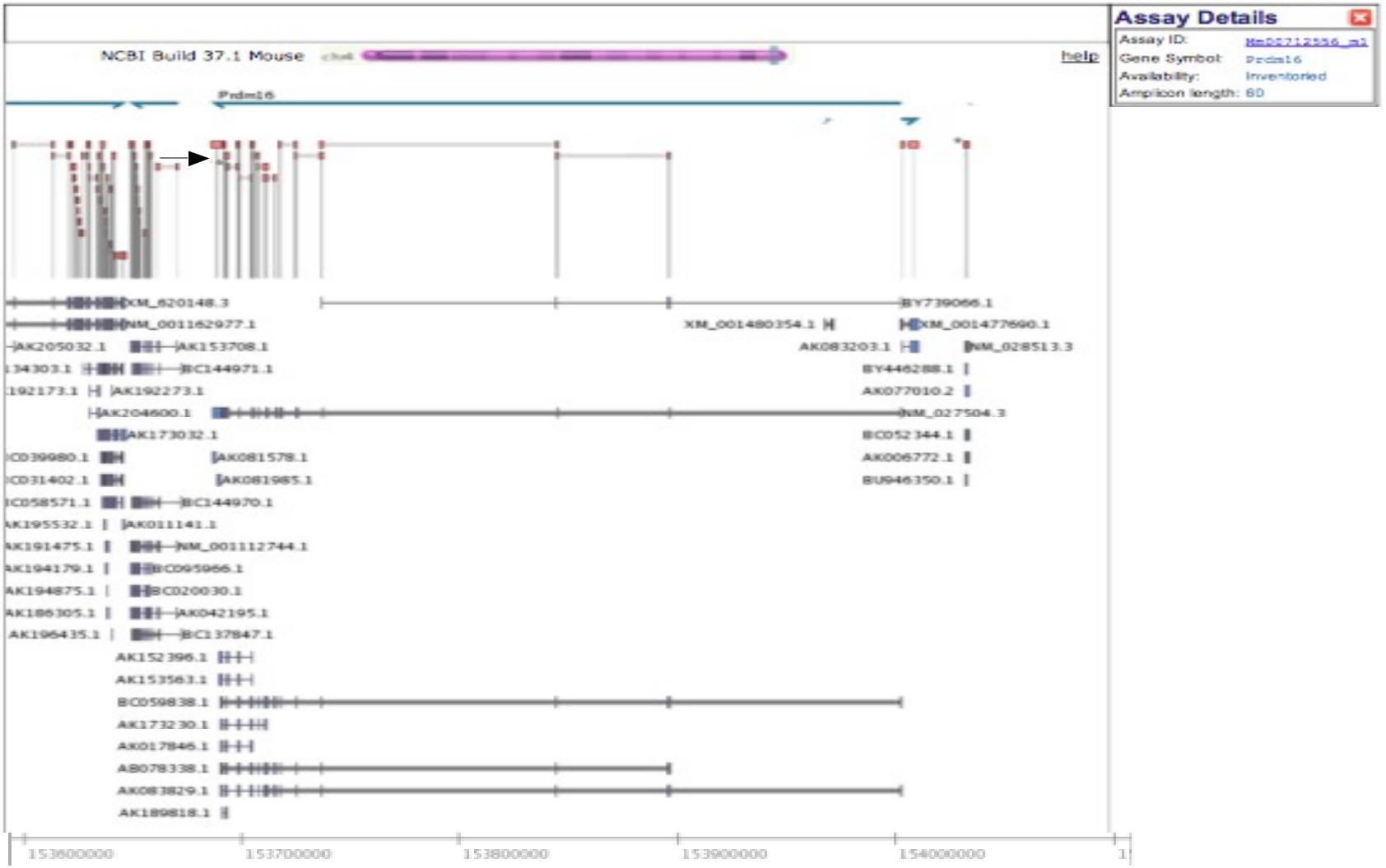

Figure 16: *Prdm16* alignment map. Arrow represents the primer probe set we used; the blue bars are different splice variants, where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website

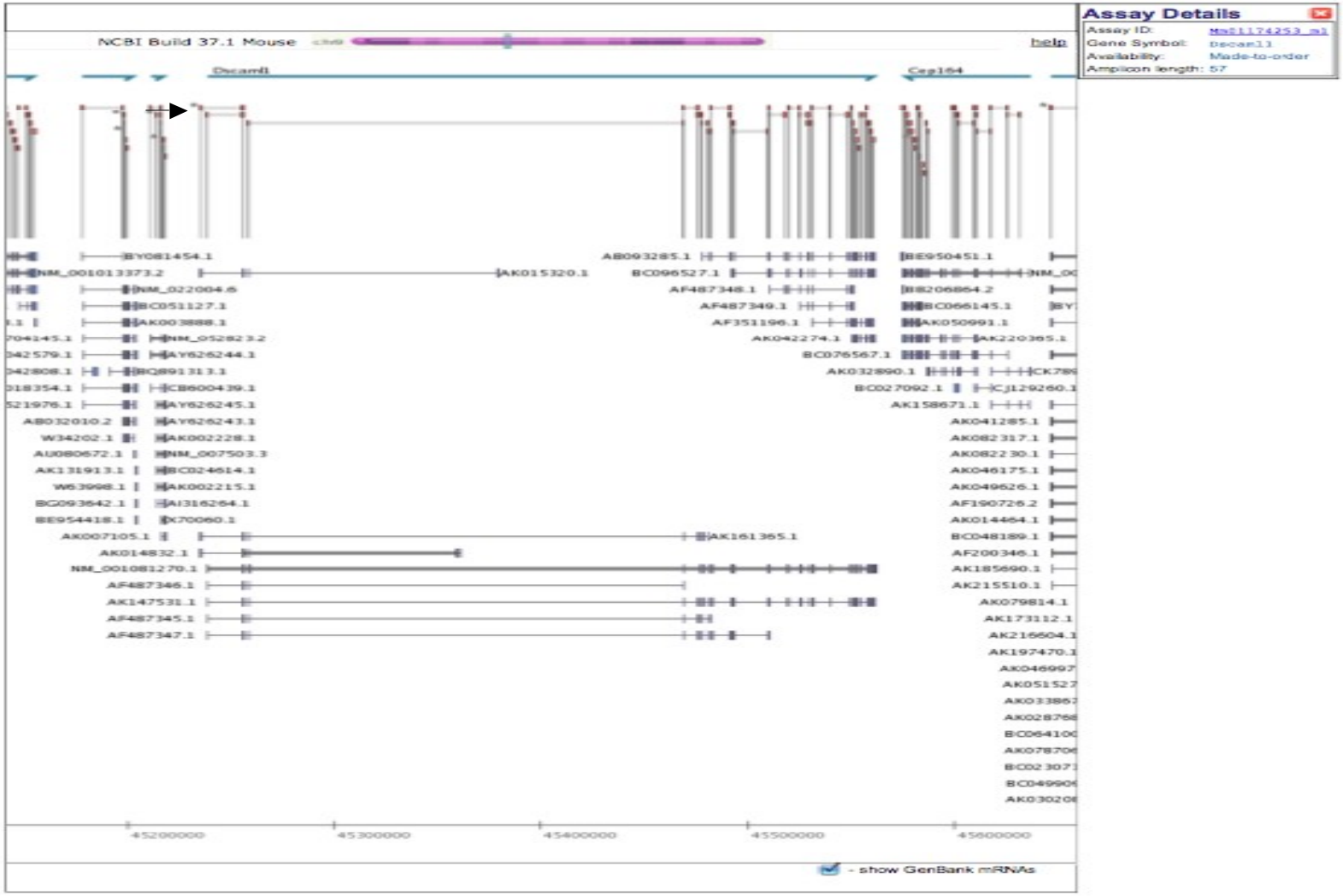

Figure 17: Alignment map of *Dscaml1*. Arrow represents the primer probe set we used; the blue bars are different splice variants,

where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website

Figure 18: Alignment map of Actin-β . Arrow represents the primer probe set we used; the blue bars are different splice variants, where the blue boxes represent exons and the grey lines indicate introns. Data from TaqMan (Applied Biosystems) Website