Using Isolation and Analysis Techniques of *Drosophila melanogaster* cDNA as Models for Human Disease

Aaron M. Jacobs, Nicole Wingard and Daniel S Matasic

Abstract: *Drosophila melanogaster* genomes have been extensively examined and share many human homologs, making them great model organisms to learn more about human diseases. This experiment focuses on the extraction of a cDNA sequence from the fruit fly genome that was inserted into an *E.* coli plasmid. Through plasmid DNA isolation techniques and PCR amplification the cDNA insert was extracted and sent through gel electrophoresis to determine size. In order to determine what gene was involved, the cDNA insert was sent out for sequencing. The expected results after examining this sequence is that it will have homologous DNA sequences to humans and possible disease related ramifications. The gene, methuselah-like 8, which was found in this experiment actually didn't share homologous domains with humans, but played a key role in *Drosophila* development and potentially could show implications related to aging in humans.

Introduction: *Drosophila melanogaster* has been a model organism over a span of 100 years in many science related fields. Commonly known as the fruit fly, *D. melanogaster* has had its entire genome sequenced and it has been found that around 75% of disease-related genes in humans are homologous to those of flies. This large similarity is due to the fact that a high number of domains are conserved across the species. The fly can reproduce very quickly and are known to have life spans of only two weeks. Such short life spans allow for a total analysis from the beginning of the flies life cycle to the flies death, this produces the unique opportunity to see how a disease or disorder works from start to finish. Also, an important aspect of *D.*

melanogaster is that they share many of the same responses to drugs as mammalians would. This relationship allows for scientists screen potential drugs for their effectiveness against a wide range of diseases and disorders such as nervous system related disorders, cancers, diabetes and even cardiovascular diseases. By being able to see how these diseases work in a laboratory setting scientists can isolate specific proteins related to these diseases and can develop improved treatments for humans with these same disorders and diseases[1].

Diabetes is a disease that results in high blood sugar levels because of a deficiency of some kind in insulin production or the lack of a cellular response to insulin[2]. Around 4% of the population is stricken with diabetes making this a good area to be further studied using *Drosophila* drug discovery. Due to recent discoveries that show a relatively strong relationship between metabolic pathways and the endocrinology of humans and flies, *Drosophila* could make a good model organism for future diabetic related research. Though the pathways relating to molecular mechanisms are relatively conserved, the actual structures that house these pathways vary greatly. When *Drosophila* insulin producing cells are destroyed, they show increased glucose levels and resistance to starvation. Glucose related pathways and the proteins between them are homologous, but insulin related receptors are less known. In *Drosophila* a decrease in development and body size is shown when there are insulin deficiencies, creating an easy phenotype to look for when running drug related tests to find drugs that stop insulin deficient mutants[1]. Diabetes, along with other diseases, can be examined further using *Drosophila* as a model organism, eventually leading to effective drugs for treatment purposes.

Methods: The *E. coli* colonies were prepared by Daniel Matasic, the bacteria's plasmids contained ampicillin resistant genes as well a cDNA sequence from the *Drosophila melanogaster*. We extracted two colonies from the agar plates, using inoculating techniques and

placed the colonies in Luria broth and penicillin. The reason penicillin was used was because the bacteria that weren't carrying the cDNA from the flies would be killed off only leaving the cDNA carrying, ampicillin resistant bacteria. After these were allowed to incubate further techniques were used to extract and isolate the cDNA in question, specifically this experiment used the QuickLyse Miniprep kit[3].

In order to amplify the small amount of cDNA extracted, PCR amplification had to be used. PCR is a common tool used in amplifying small quantities of DNA in a small amount of time. PCR amplification uses a master mix of a buffer, dNTPs, two primers and taq polymerase. The primers focus on SP6 and T7 phage RNA polymerase promoters, which are on opposite sides of the inserted region of the plasmid. 24ul of master mix was combined with 1ul of both plasmids extracted from the two bacteria colonies as well as with 1ul of sterile water in order to determine if any contamination occurred, this is referred to as the negative. These were run through the PCR machine, amplifying any *D. melanogaster* cDNA present. The amplified cDNA is then subjected to agarose gel electrophoresis[3].

Agarose gel electrophoresis is a method used to determine the size of multiple DNA fragments. Since DNA has a negative charge, due to the phosphate groups, the DNA fragments migrate away from the negative pole and move toward the positive pole because of an electrical current that is sent through the gel. The size of these fragments determines the length in which the fragments move, the larger they are the shorter they will go. The gel contains wells where the samples are loaded into; we inserted into the first well 5ul of DNA ladder. The DNA ladder acts as a reference point in determining the size of the various samples that were loaded. The second and the fourth well contained the plasmids A and B, respectfully, without PCR amplification; these were a mixture of 2ul of cDNA, 8ul of water and 2ul of 6x loading dye. The

third and fifth well contained the plasmids A and B, respectfully, which underwent PCR amplification, these were a mixture of 4ul of 6x loading buffer added to the PCR products. 12ul of all the cDNA related samples were added to their respected wells as well as 12ul of negative control to the sixth well. The plasmid cDNA that showed that amplification occurred, which could be seen by the gel electrophoresis, were sent off for sequencing[3].

The sequencing technique used for this lab was the dideoxy method, which was carried out by the Nucleic Acid Facility on campus. Once the results were received, bioinformatics were then used to find further information about what particular *D. melanogaster* cDNA was extracted. The sequence was extracted from MEGA 5, refer to figure 4, but it showed complications. Instead an alternative sequence, refer to figure 5, was entered into the sequencing public database created by the NCBI in order to determine the gene it encoded and in turn the protein that gene encoded for. The information that this database provided allowed us to determine if the *D. melanogaster* protein was homologous to proteins that cause diseases in humans[3].

Results: The Agarose Gel Electrophoresis, refer in figure 1, represents the various sizes of plasmid cDNA. Well two shows cDNA from plasmid A and well three shows the PCR product of that cDNA, both show equal size between 100-300bp when compared to the DNA ladder as shown in figure 2. Well three doesn't show any cDNA migration, which should have shown plasmid B, but well four, containing the PCR product of the cDNA from plasmid B, showed migration with a size between 100-300bp. The negative control showed no contamination, since there was no migration of DNA material of any kind.

The alternative sequence that was received from the Nucleic Acid Facility on campus showed, once sent through the mega software, the sequence shown in figure 5. The plasmid A

sequence, which shows multiple waves from multiple DNA fragments, is shown in figure 4. The protein ID assigned to the alternative sequence after being run through the online database was AAF47322.2. The database showed this to be the Methuselah-like 8 protein, found in *Drosophila melanogaster*. There was only one conserved domain, which was the Mth_Ecto domain. The database showed that this domain was not conserved in humans.

Discussion: The alternative sample of plasmid cDNA, figure 5, was run though the online databases showing that the Methuselah-like 8 protein, shown in figure 3, was what this particular sequence coded for. This specific protein is not homologous to humans, it is only found in organisms like *Drosophila, Arthropoda, Hexapoda and Insecta*[4]. The Methuselah-like proteins are receptors, specifically GPCR, and play a major role in the life span and stress response in *Drosophila.* When a mutation occurs in the Methuselah gene, the average life span of these flies can increase and resistances to stresses are enhanced[5]. Mutants actually receive a 35% increase in average lifespan and resistance to stresses including things like starvation, oxidative damage and heat[6]. The Methuselah receptor has also been found to be crucial for the development of the flies, because in larval and pre-adult stages mutations in this gene resulted in death[5]. Even though there are no human homologs of this gene, would a drug affecting such a receptor in humans be worth the life span increase[7]?

The investigation into modifying signaling pathways in humans, like those in *Drosophila*, to cause a life span increase could come at a price. Drugs that would delay aging by affecting these types of pathways have shown pleiotropic effects in model organisms like *Drosophila*. In mutated Methuselah flies some of these resulting effects include a reduction in reproductive output as well as insulin production creating diabetic phenotypes. Researchers also found that increasing life expectancy didn't necessarily mean a healthier life beyond the expected age of

death. The delay in aging didn't seem to be related to old age problems, like motor activity[7]. Even though the Methuselah gene was not homologous to humans, it still showed future application to human related problems like aging and possibly diabetes. As described in the introduction, *Drosophila* is being considered as a possible route for diabetes related drugs. The Methuselah-like proteins could be the key to unlocking the cure for diabetes, as mentioned earlier when there are insulin deficiencies the size of the flies' decreases and there are negative side effects to development as well as a resistance to stresses like starvation[1]. This same relationship was found in mutated Methuselah flies, meaning that a key player in insulin related problems could be the Methuselah-like proteins[7]. The relationship between insulin and this protein could give insight into possible closely related proteins and signaling pathways in humans. Using *Drosophila* as model organisms for research into humans and human related diseases, as shown in this lab, can be a very effective and easy way to collect important data in a short amount of time. Using flies allows for multiple trials to be run on multiple generations in a short amount of time and can produce results that have vast implications, even to humans.

The alternative sequence shown in figure 5 was used in bioinformatics because the original plasmid cDNA of plasmid A and B, A being shown in figure 4, showed contamination of some kind. The multiple layers of waves shown in figure 4 represent various different samples of DNA, this means that at some point in the preparation of the sample an outside source of DNA contaminated the plasmid's cDNA, the same result was shown for plasmid B. Even though the negative in the electrophoresis, figure 1, didn't show any contamination that was only the master mix and not the sample. There were also some problems that occurred during the gel electrophoresis of our samples. Our group had poor well formation, which resulted in the poor migration seen in well 5 and the absence of any genetic material in well 4. Some other errors

that occurred included the misinterpretation of the process in which the plasmid cDNA was supposed to be prepared, leaving well 2 inapplicable. This improper preparation may have resulted in the reason behind why well 2 and 3, in figure 1, were the same size.

References:

1. Pandey, U.B., and Nichols, C.D., 2011. Human Disease Models in Drosophila melanogaster and the Role of the Fly in Therapeutic Drug Discovery. *Pharmacological Reviews*, 63, 411-436.
2. "Diabetes Mellitus." *Wikipedia*. Wikimedia Foundation, 11 July 2012. Web. 07 Nov. 2012. <http://en.wikipedia.org/wiki/Diabetes_mellitus>.
3. Burpee, D. Bio 230 Laboratory Activity: cDNA Isolation and Analysis. Pennsylvania State University, 2012
4. "Conserved Domains on [gi|23092651|gb|AAF47322|]." *NCBI*. U.S. National Library of Medicine, n.d. Web. 07 Nov. 2012. <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?RID=9NK06DMH01R>.
5. Brody, T. & Cravchik, A. *Drosophila melanogaster* G Protein–Coupled Receptors. *The Journal of Cell Biology* 150, F83–F88 (2000).
6. West AP, Jr, Llamas LL, Snow PM, Benzer S, Bjorkman PJ. Crystal structure of the ectodomain of Methuselah, a *Drosophila* G protein-coupled receptor associated with extended lifespan. Proc Natl Acad Sci USA. 2001;98:3744–3749.
7. Alic, Nazif, and Linda Partridge. "Antagonizing Methuselah to Extend Life Span." *National Center for Biotechnology Information*. U.S. National Library of Medicine, 31 Aug. 2007. Web. 07 Nov. 2012. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2374980/>.

Table and Figures:

Figure 1.

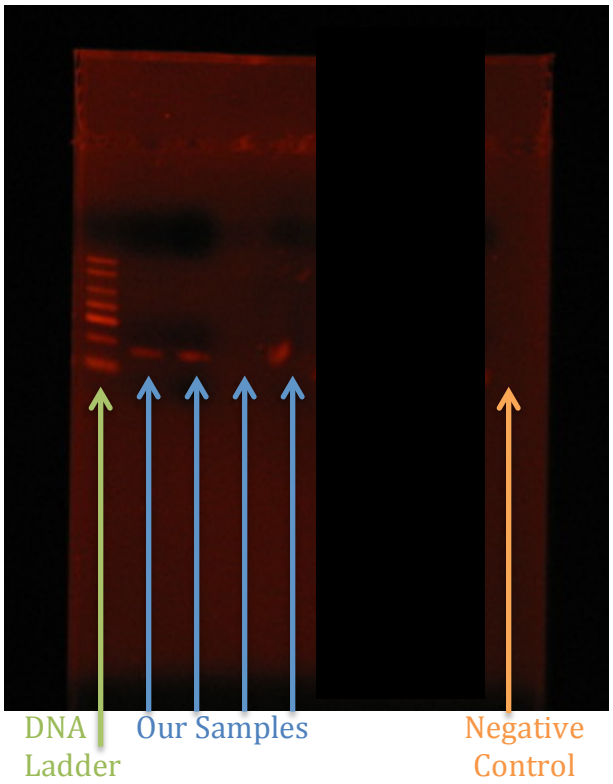

Figure 1: Agarose Gel Electrophoresis. This picture shows our groups plasmid cDNA after it was run through electrophoresis.

Figure 2.

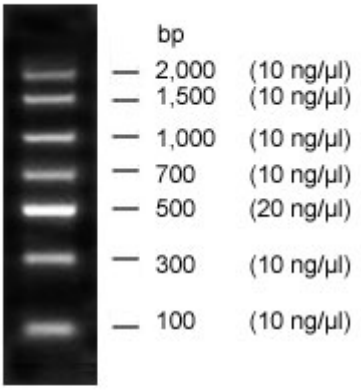

Figure 2: DNA ladder scale. This is used in determining the size of the various sample DNA[labmanual].

Figure 3.

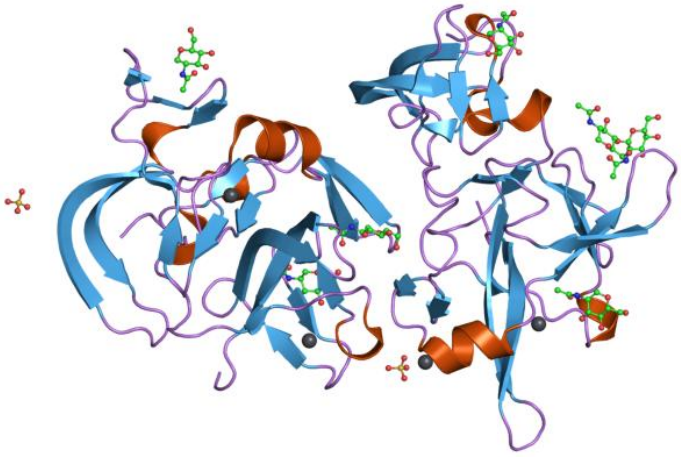

Figure 3: Methuselah-Like Protein. This figure shows the structure of the protein that was found through sequencing. (Wikipedia, 2012 Methuselah-like proteins)

Figure 4.

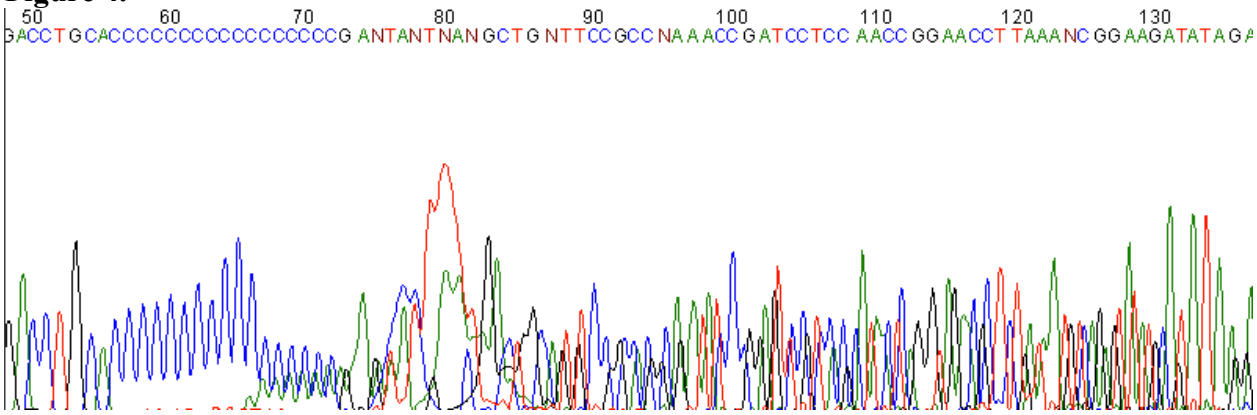

Figure 4: Sequence of cDNA for Plasmid A. This figure shows only a small portion of the whole genetic code of cDNA sent for sequencing.

Figure 5.

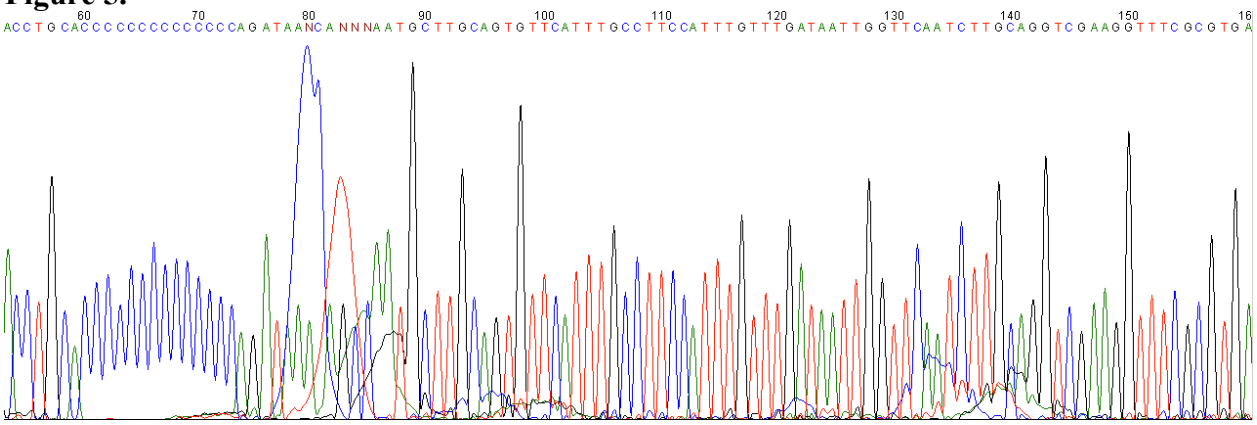

Figure 5: Alternative Sequence of cDNA. This sequence was given to me, this figure shows only a small portion of the whole genetic code of the cDNA sent for sequencing. This was used then for bioinformatics.