LARVAL ZEBRAFISH IN VIVO MODEL TO STUDY GUT MOTILITY DISORDERS

A Thesis
by
VICTOR NASR

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE
Cellular and Molecular Biology

May 2022
Department of Biology

LARVAL ZEBRAFISH IN VIVO MODEL TO STUDY GUT MOTILITY DISORDERS

A Thesis
by
VICTOR NASR
May 2022

APPROVED BY:

Mary D. Kinkel, Ph.D.
Chairperson, Thesis Committee

Cortney Bouldin, Ph.D.
Member, Thesis Committee

Theodore Zerucha, Ph.D.
Member, Thesis Committee

Zack E. Murrell, Ph.D.
Chairperson, Department of Biology

Marie Hoepfl, Ph.D.
Interim Dean, Cratis D. Williams School of Graduate Studies

Abstract

LARVAL ZEBRAFISH IN VIVO MODEL
TO STUDY GUT MOTILITY DISORDERS

Victor Nasr
B.A., University of Central Florida
M.S., Appalachian State University

Chairperson: Mary D. Kinkel, Ph.D.

The goal of this research is to establish zebrafish as a model organism to study intestinal motility diseases in humans to aid in the understanding of the pathophysiology of such diseases. The study of exactly how and why the human intestine develops certain gut motility disorders needs to be more carefully investigated to develop medicinal or surgical cures for gut motility disorders such as Hirschsprung's disease, a congenital disorder of the intestine affecting 1 in 5000 children. First, I was able to identify several zebrafish candidate genes from the literature that are of interest, and determined how conserved they are using multi-sequence alignments to compare DNA sequences with humans, mice, and zebrafish. Second, I performed a feeding behavior study to determine the age at which zebrafish larvae are mature enough to capture brine shrimp. The results show day 9 post-fertilization to be the best and most efficient day a zebrafish larvae captures brine shrimp as a meal. Therefore I concluded that a gut transit assay is best conducted at this age. The results of this study aided in developing a gut transit assay which is a

way to quantify the passage time of food and waste from feeding to clearing in a zebrafish larvae. In future studies, the Kinkel lab will build on this work to map the expression of the identified genes along the zebrafish digestive tract using RT-PCR. Using this information, diseases like Hirschsprung's disease and irritable bowel syndrome could be further investigated in greater details.

Acknowledgments

I would like to give special thanks to my PI and advisor Dr. Mary Kinkel for allowing me to be a part of the Kinkel's lab, and for assisting me throughout my research. I would like to also thank my committee members Dr. Ted Zerucha and Dr. Cortney Bouldin for their time and effort. I would like to also thank my lab mates, Kitt Franse, Savanna Sheridan, Alena Norton, Tierney Daw, and Latanya Gordon for their contributions to this research. All procedures and research methods were approved by the Institutional Animal Care and Use Committee at Appalachian State University.

Table of Contents

List of Tables

List of Figures

Introduction

Gastrointestinal (GI) diseases are manifested in many different ways and affect millions of people throughout the world. Researchers have been searching for answers of how and why certain humans develop gut motility diseases such as Hirschsprung's disease, and Irritable bowel syndrome (IBS). Hirschsprung's disease is a congenital disorder of the intestine affecting 1 in 5000 children. It varies in severity depending if all or part of the GI tract is affected, however it affects patients mostly in the distal portion of the gastrointestinal tract, where there is slow or cessation of motility movements due to the absence of neuronal ganglion cells in the myenteric and submucosal plexus. The direct effect of this leads to the accumulation of feces. Patients' symptoms vary but most complain of prolonged constipation, and in severe cases, patients present with acute bowel obstruction (Haikal et al., 2022).

When it comes to the general health of the human body, all organs work together as one machine to maintain homeostasis and optimum health. One of the most important organs for nourishment and maintenance of health is the intestine. The most important functions of the intestine are food and water absorption. These functions are carried out by two separate yet continuous organs: the small intestine and the large intestine, also known as the colon. (Figure1).

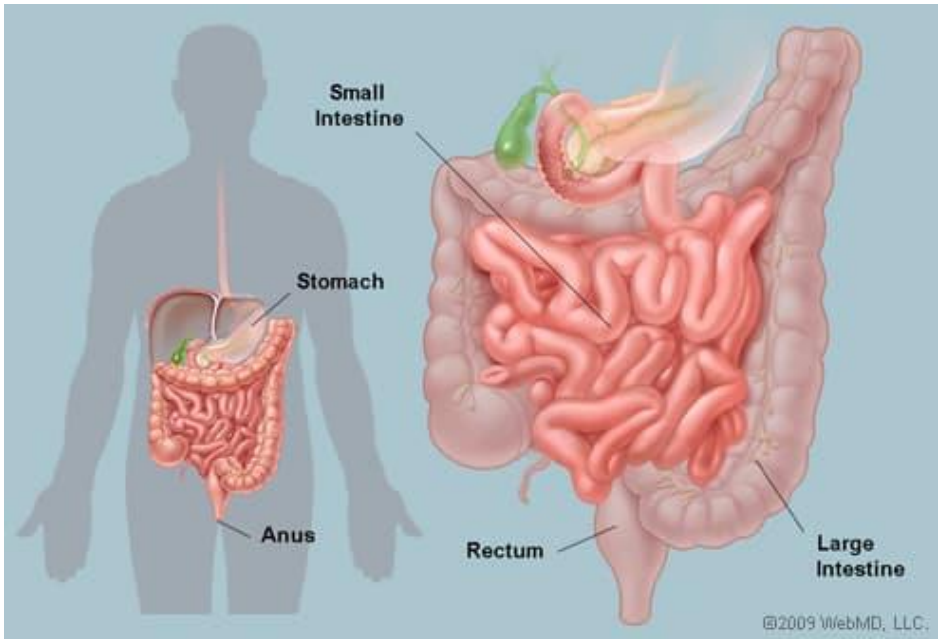

Figure 1. The human digestive system with its different organs. From https://www.webmd.com/digestive-disorders/picture-of-the-intestines

The small intestine is involved in finishing up digestion begun in the stomach and has a major role in absorption of nutrients from consumed foods. The large intestine absorbs water from waste, creating feces to be excreted. In a certain number of the population, due to one or more of several known and unknown causes, the large intestine may start functioning abnormally with chronic excessive activity leading to chronic abdominal discomfort and other gastrointestinal problems. This is known as irritable bowel syndrome (IBS).

Irritable Bowel Syndrome

IBS is a chronic disease affecting the gastrointestinal system of around 10% of the population globally with the fulfillment of the Rome III criteria for diagnosis (Ford et al., 2020). Women are three to four times more likely to develop IBS possibly due to hormonal factors. Women with IBS are more likely to be diagnosed with IBS with constipation (IBS-

C) than IBS with diarrhea (IBS-D). Most people who develop IBS are in their childhood to young adults under the age of 45 years of age (Radovanovic-Dinic et al., 2018). It affects the western world such as the United States and Australia primarily due to diets mainly dependent on refined carbohydrates and fast foods. In fact, it is estimated that over 3.6 million physician visits annually in the US are solely for IBS which consists of care and expenses of over $20 billion (Raskov et al., 2016). Gathering statistical information regarding the prevalence of IBS subtypes is complex because many IBS subtypes overlap in terms of symptoms and many patients switch between subtypes during the course of the disease (Enck et al., 2016).

In irritable bowel syndrome, the large intestine's different layers and segments do not have smooth coordination in movement, and thus the function of the colon is disturbed yet its anatomical parts remain normal without obvious defects. This disorder is characterized by recurrent or chronic abdominal pain that is relieved or exacerbated by defecation, changes in bowel habits, and bloating that lack a known structural or anatomic explanation making it difficult to diagnose and treat (Holtmann et al., 2016).

Signs and symptoms of IBS

Many people experience gastrointestinal symptoms whether they have been diagnosed with IBS or not. For instance, many people are allergic to certain foods and may find other foods are indigestible or hard to digest as in lactose intolerance. These cases can produce many different abdominal symptoms including pain, discomfort, and gas. IBS patients usually complain of abdominal pain and discomfort that gets better with defecation. The bowel habits of these patients are usually altered ranging from constipation to diarrhea to a mix of both depending on the type of IBS a patient has: IBS with diarrhea (IBS-D), IBS

with constipation (IBS-C), or a mix IBS (IBS-M)(Enck et al., 2016). IBS-C patients might experience incomplete bowel evacuation from a few days to a few months due to constipation, or heavy stool movements of more than four times per day which sometimes consists of urgency (having to rush to the toilet) as in the case of IBS-D. In each case, intestinal motility is affected differently and produces different symptoms. Bloating and abdominal distention which include a sensation of gas-filled bowel are also characteristic symptoms of IBS. Most patients have episodes of symptom exacerbation, followed by periods of remission depending on many factors and etiologies, directly and indirectly, affecting their condition (Radovanovic-Dinic et al., 2018). Other symptoms are also associated with IBS. These include dysphagia (difficulty swallowing), nausea, vomiting, and certain chronic pain disorders in the muscles and joints such as chronic fatigue syndrome and fibromyalgia. Psychiatric disorders such as anxiety and depression are often associated with IBS as well (Raskov et al., 2016).

Etiologies of IBS: Brain-gut axis and psychological disorders

A variety of possible causes of IBS have been proposed. For many years, patients complained of abdominal symptoms that were triggered by anxiety, depression, and several other mental disorders without knowing if there is any relationship between the two different complaints. Recent studies have shown that anxiety, mood changes, and other brain-related disorders could directly or indirectly cause IBS-like symptoms in the abdomen (Figure 2) (Holtmann et al., 2016). Furthermore, studies have also shown that the opposite is also true: IBS symptoms could lead to anxiety, mood, and personality changes (Radovanovic-Dinic et al., 2018). Patients with IBS may have suicidal ideation as a result of their bowel symptoms. Additionally, several studies have shown higher levels of sexual and physical abuse among

IBS patients. Psychological disorders such as anxiety and depression can affect the brain-gut axis, promoting the release of corticotropin-releasing hormone, which can influence mood, inflammatory pathways, and digestive motility via neuroendocrine and autonomic pathways. Stress in IBS patients increases the levels of pro-inflammatory interleukins, activating both the hypothalamic-pituitary-adrenal (HPA) and the hypothalamic-autonomic nervous system axes and consequently increases cortisol levels, known as the stress hormone which has positive feedback on the brain-gut axis (Figure 2) (Radovanovic-Dinic et al., 2018). As

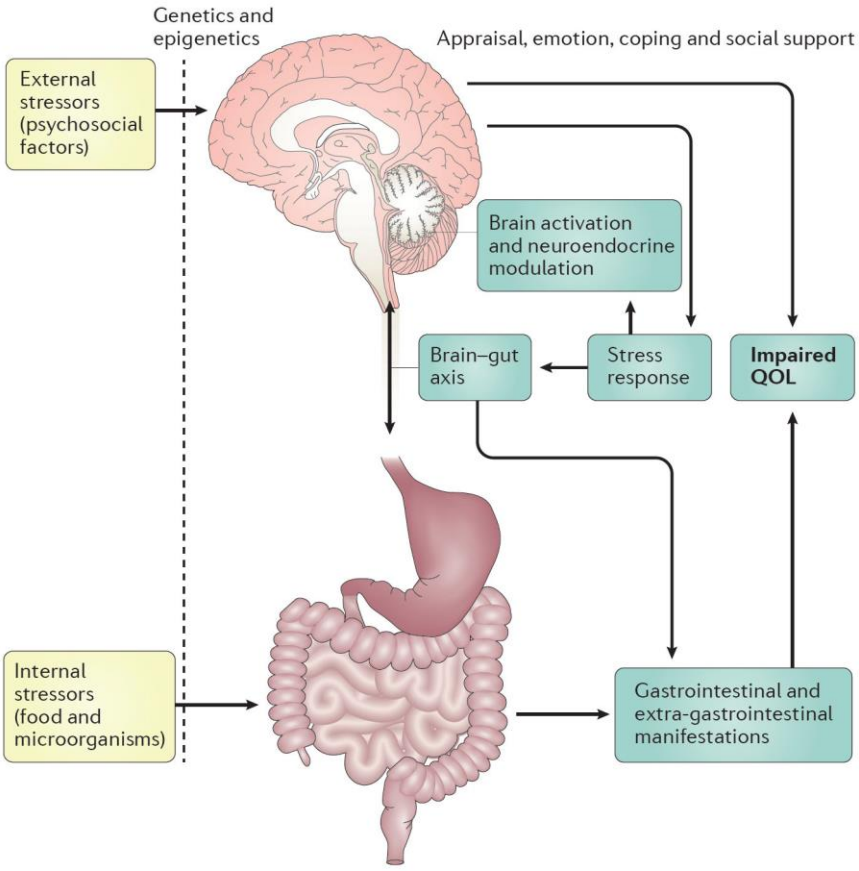

Figure 2. Concept of multiple etiologies of IBS and its effect on the brain-gut axis. Etiologies of IBS include external stressors such as stress and internal ones such as food and microorganisms. Their direct influence and effect is shown on the both the brain and the gastrointestinal system. The result is manifestation of GI and extra-gastrointestinal symptoms which ultimately impairs the quality of life (QOL). From Enck et al. (2016).

a result, people with IBS have increased or decreased motility manifested as diarrhea or constipation respectively due to the intestines being highly sensitive to stress levels.

Serotonin, or 5-hydroxytryptamine (5-HT), is produced in the brain and also in the gut by intestinal enterochromaffin cells and is an important signaling molecule in the central and peripheral nervous systems of humans as well as the gut (Qin et al., 2014). It helps regulate the normal functioning of the brain, cardiovascular system, and gastrointestinal tract by acting on several types of receptors. Dysfunction in serotonin signaling has been identified in several diseases, including anxiety disorders, depression, hypertension, and irritable bowel syndrome (Szeitz and Bandiera, 2018). This hormone binds on the 5-HT3 and 5-HT4 receptors to directly affect gastrointestinal motility, secretion, and sensation. It has been observed that when serotonin concentration is low, patients tend to have IBS with constipation. In the case of females, serotonin concentration levels are normal although as mentioned earlier females tend to have higher incidents of IBS with constipation, it is not well understood the reasoning behind this. However, when this hormone's concentration is high, patients tend to have IBS with diarrhea, demonstrating its role in the motor dysfunction associated with this condition (Saha, 2014).

The impact of microbiota and the role of infection in IBS

In daily life, a person encounters many different microbes from simply drinking water to eating different foods, and from touching contaminated surfaces. While some microbes might not adversely affect us, others might find their way into our bodies and specifically into the digestive system. As a result of this microbial invasion, the immune system reacts by recruiting white blood cells such as T cells, B cells, and mast cells which have different specialties in attacking and eliminating the threat. Due to this activation, inflammation

increases in the intestinal tract and throughout the body causing a variety of different symptoms like abdominal pain, diarrhea, and many others as shown in Figure 3. Enteric inflammation has been shown to be present in some patients with IBS after prolonged infectious enteritis (post-infectious IBS) such as in the case of *Giardia lamblia*, a bacteria which patients acquire from drinking water from streams while

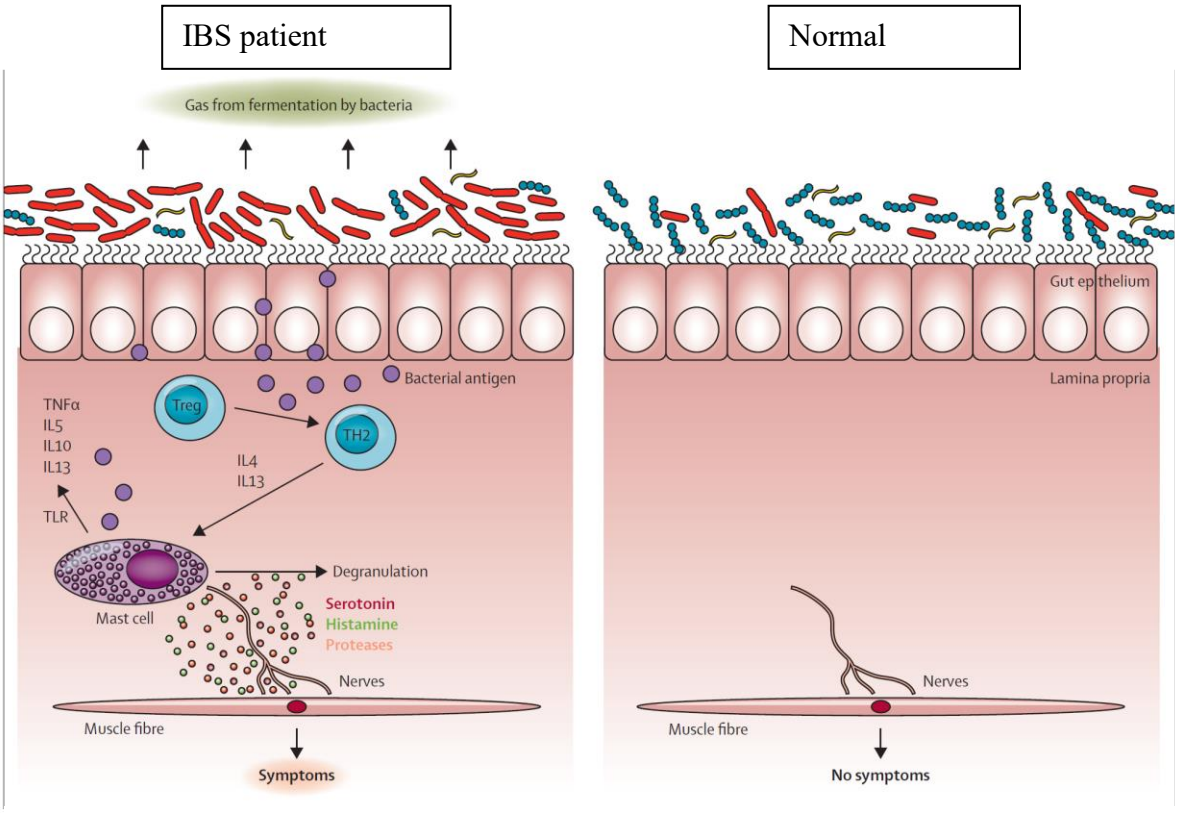

Figure 3. Infection/antigen effect on intestinal permeability. A depiction of a comparison between a normal intestine and an intestine of an IBS patient. The IBS intestine shows intestinal permeability by a bacterial antigen which triggers an immune response activation by recruiting immune cells T-helper-2-cell (TH2), and mast cells releasing different types of interleukins, histamine, proteases and serotonin. This cascade of signaling results in neural excitation and smooth muscle contraction, leading to disturbed intestinal transit with many different IBS symptoms, as well as extra-intestinal symptoms when these cytokines are released into the circulation (Holtmann et al., 2016).

camping or hiking (Radovanovic-Dinic et al., 2018). Acute enteric infections have been shown to frequently precede the onset of IBS, particularly IBS with diarrheal symptoms, or other functional gastrointestinal disorders (Holtmann et al., 2016). Previous studies have used quantitative immunohistochemistry to show increased infiltration of mast cells and T cells in the mucosal layer of the small and large intestine of some patients with IBS, especially post-infection (Figure 3) (Enck et al., 2016). This could suggest a role in hypersensitivity or overreaction of the immune cells when the infection was present. Thus, mucosal tissues and smooth muscle as well as the embedded enteric nervous system might directly be impacted by this immune response leading to the development of IBS symptoms, and directly influencing intestinal motility (Ye et al., 2021). A study using mice in which mucosal inflammation was induced chemically, suggested that the severity of the immune response and inflammation is linked to the severity of subsequent visceral hypersensitivity in the mucosal layers of the intestinal tissues, which might be one of the underlying mechanisms for the development of IBS-like symptoms (Enck et al., 2016).

Patients with previous anxiety symptoms and other mental disorders have been shown to be increasingly affected by hypersensitivity post-infection with a microbe (Luo et al., 2018). A previous study showed that there was an inverse correlation between CD4-positive T lymphocytes expressing interleukin 2 and anxiety scores pre-exposure to a microbe (Holtmann et al., 2016). Additionally, it was also noticed that there was an association between a T-helper-2 cell cytokine phenotype, an immune cell type, at the time of exposure and the subsequent development of post-infectious IBS. This suggests that susceptibility to IBS after an acute enteric infection results from the recruitment of a T-helper-2 immune-cell response as part of eliminating of the microbial threat (Holtmann et al., 2016).

Diet and IBS

For a long time, it has been known that consumed food can directly affect the human body and alter health through its role in maintaining normal body functions. In recent years, most diets have relied heavily on carbohydrates such as refined sugars and other harmful substances that were minimally or never consumed in past diets. The human body's digestive system, having not evolved for processing high-sugar diets, and as a result, it has had numerous digestive problems manifested in the form of IBS symptoms as well as other serious health problems such as diabetes. Chronic inflammation has been exacerbated due to the high-sugar diets as well.

Bread and other carbohydrates have a protein called gluten. Humans have a digestive enzyme called dipeptidyl peptidase-IV which helps break down gluten. However, this enzyme does not break down gluten proteins completely and undigested gluten travels to the small intestine where some people react negatively to it and fall under either gluten sensitivity group or celiac disease group. It has been observed in one multicenter double-blind trial of 140 patients that the patients experienced gluten sensitivity, which is characterized by IBS-like symptoms following gluten ingestion but in the absence of celiac disease (Elli et al., 2016) . In another randomized trial, 33 in 45 patients with IBS with diarrhea symptoms who were given either a gluten-containing diet or, gluten-free diet, lactulose and mannitol excretion which are markers for intestinal permeability were increased in gluten-containing diet serving as direct evidence for small-bowel mucosal permeability (Vazquez-Roque et al., 2013). Based on studies like these, many researchers think that gluten alters the intestinal permeability, also known as leaky gut, and allows

undigested foods into the bloodstream. This activates the enteric and autonomic nervous system which typically produces the IBS symptoms (Radovanovic-Dinic et al., 2018).

Certain types of carbohydrates such as fermentable oligosaccharides, monosaccharides, and disaccharides and polyols (FODMAPs), which are present in legumes, stone fruits, lactose-containing foods, and artificial sweeteners, might exacerbate symptoms with bloating, pain, and others in a large group of IBS patients due to their fermentation and osmotic effects by causing water retention in the small bowel according to MRI studies (Rodiño-Janeiro et al., 2018). The anaerobic fermentation by the microbiota in the large intestine of indigestible polysaccharides such as dietary fiber and resistant starch produces the metabolites short-chain fatty acids (SCFAs). The SCFAs have many known benefits for the gut including improvement in gut barrier function, energy supply to gastrointestinal epithelial cells, and a decrease in inflammation.

It has been shown that SCFA-producing bacteria such as *Roseburia*, *Blautia*, and *Veillonella*, are increased in healthy patients thus increasing the levels of SCFA which lowers inflammation. In IBS patients, the opposite has been seen in which there is an altered level in SCFA-producing bacteria (Raskov et al., 2016). Researchers have also observed an increase in gas-producing bacteria in IBS patients which would explain the flatulence and abdominal pain. Patients that have IBS with diarrhea are most affected because the excessive production of gas cause faster colonic transit because the colon of these patients is more sensitive to increased intestinal volume. Methanogenic archaea which are responsible for intestinal gas removal is depleted in patients with IBS and is negatively correlated with the presence of loose stools. However, a significant increase in this microbial group is characteristic of patients with slow transit, and constipation specifically, IBS-C patients (Enck et al., 2016).

Recent studies have shown that when IBS patients consumed a low FODMAP diet, they experienced a significant reduction in IBS-related symptoms due to the lower amount of abundant gas, and volume in the intestinal lumen (Halmos and Gibson, 2019).

Current treatments of IBS

For many years, researchers sought to quickly find treatments for IBS due to its bothersome symptoms which affected the quality of life. However due to the different types of severity of this disease, IBS treatments range from non-pharmacological to pharmacological approaches that depend on how tolerant the person is, and the specific type of the diagnosed IBS, being IBS-C or IBS-D or both.

Non-pharmacologic treatments are many and involve no drugs. Physical activity is one of the most important current treatments for IBS-C. When the person stays active, this increases blood flow in the whole body while increasing metabolism. Additionally, exercise helps the flow of waste products in the intestine due to gravity, and push and pull forces on the body. This will help anything that is blocked or constipated in the colon to move to be excreted. Diet monitoring is also crucial to improving patient outcomes. Patients are educated to keep a record of the foods they eat and the items that produce their IBS symptoms such as the known gluten effect and are advised by their physicians to avoid these foods. Additionally, patients are encouraged to incorporate foods that are known to be anti-inflammatory such as legumes, vegetables, and some spices (Radovanovic-Dinic et al., 2018). Lastly, patients are encouraged to reduce stress and increase stress-relieving habits like meditation and relaxation. Stress has a direct effect on the intestine due to the interconnections between the central nervous system including the brain and the spinal cord, and the enteric nervous system that comprises of the network of nerves embedded in the

digestive system. Stress also increases cortisol level, the stress hormone, which in itself also exacerbates symptoms by directly influencing smooth muscle activity in the gut controlling motility (Saha, 2014).

Pharmacological treatments are encouraged and prescribed when non-pharmacological treatments are insufficient in controlling patients' symptoms. For IBS-D, patients are prescribed antidiarrheals such as loperamide, a synthetic opioid, that acts on intestinal smooth muscles to prolong transit time and inhibit peristalsis by decreasing fecal volume and frequency, as well as improving stool consistency. Antidepressants such as tricyclic antidepressants (TCA), selective serotonin reuptake inhibitors (SSRI), and serotonin-norepinephrine reuptake inhibitors are another class of drugs for IBS treatment. Based on several studies, these drugs are believed to act via centrally-mediated antinociceptive pathways in the brain decreasing abdominal pain associated with IBS. Additionally, these agents may affect the gastrointestinal tract transit times due to their effect on the nervous system (Wall et al., 2014). Physicians usually prescribe antispasmodics such as alverine or dicyclomine for IBS-C patients. These are drugs that relax the smooth muscles of the gut via anticholinergic mechanisms or calcium channel antagonism. Osmotic laxatives like polyethylene glycol (PEG) are another class of drugs that treat the IBS-C subtype. These drugs work by increasing water volume in the intestinal lumen in order to decrease intestinal transit time which in turn relieves constipation. Increasing the overall water consumption could potentially help patients with the IBS-C subtype.

Bulking agents such as fiber as well as probiotics containing beneficial bacteria are natural products that can benefit all types of IBS patients. Fiber products such as psyllium husk are often used for their ability to increase stool frequency which benefits IBS-C

patients, and transit time which benefits IBS-D patients. Probiotics contain live beneficial bacteria that enhance the gut flora of the intestine which in turn benefits all subtypes of IBS disease (Saha, 2014).

Current research related to IBS

Due to the ever-expanding research into IBS, its pathophysiology, and how it affects the body, physicians have been able to treat IBS patients effectively with and without pharmacological treatments. Despite numerous studies, the etiopathogenesis of irritable bowel syndrome is still not completely understood because of its complexity in which several body systems and organs are interacting to produce or alleviate the symptoms. Current research is focused on understanding the primary causes of IBS whether be it an infection, autoimmunity, or other etiology. Taking that into consideration, the feedback between the different body systems especially the central and enteric nervous systems in addition to the entire digestive system needs to be understood in greater detail in order to find improved treatments and potential cures.

Recently, researchers have experimented with genetic engineering of bacteria, specifically transgenic *Lactococcus lactis* expressing mature human interleukin-10 instead of the normal thymidylate synthase in 10 patients with IBS (Rodiño-Janeiro et al., 2018). Results showed improvement in the symptoms of these patients. Another recent trial was an experiment in the use of fecal microbiota transplantation in which it was hypothesized that a healthy microbiota in the gut could be restored by fecal microbiota transplantation and improve the overall IBS symptoms. However, these recent experiments consisted of only a few uncontrolled small studies. Therefore, larger and more in-depth experimentation is

needed to confirm if these would be potential treatments for IBS (Rodiño-Janeiro et al., 2018).

Gut Motility Diseases and Zebrafish

Gut motility diseases such as IBS and Hirschsprung's disease have affected humans for decades even when most diets have dramatically changed over time. Researchers and scientists have been searching for answers of how and why certain numbers of humans develop gut motility diseases. This research involved the identification and testing of candidate genes and their direct potential involvement in motility dysfunction.

Zebrafish could serve as an effective model for understanding human intestinal disorders. These disorders result in failure to properly move food through the digestive tract. Common symptoms in most of these diseases include constipation, diarrhea, and malnutrition (Keller et al., 2018). These symptoms are usually pharmacologically treated, but the actual cause is rarely treated. Several zebrafish lines already exist for modeling specific genetic mutations that disrupt gut motility. Motility is a key factor in digestion, energy balance, and homeostasis (Murphy and Bloom, 2006). The smooth muscles that control motility are innervated by enteric nerves, causing them to contract and relax, thus producing motility (Cummings and Overduin, 2007). One way to measure motility is using a gut transit assay. This assay quantifies the time it takes for food to pass through the entire gastrointestinal tract. However, to clearly establish normal physiology, we need to understand the anatomy and physiology of the intestine.

Due to many advantages, zebrafish can be used as models for understanding human intestinal diseases. Firstly, previous studies have established that zebrafish gene expression is conserved with humans: similar expression patterns are observed during embryonic

development. They have at least one orthologous gene for 70% of all human genes and they have orthologous genes for 82% of disease-associated genes in humans (Howe et al., 2013). The Kinkel lab has previously identified several candidate genes from the literature that are known to be expressed in the gut of zebrafish as well as some being conserved with humans or other organisms such as mice, shown in Table 1. My current research identified additional candidate genes that are of interest from the literature.

Table 1. Zebrafish gut motility candidate genes.

Gene	**Location**	**Zebrafish intestine expression pattern**			
		embryonic 0-72 hrs	early larval 3-9 dpf	late larval 10-21 dpf	adult >21 dpf
Anoctamin 1 (*ano1*)	ICC	?	✓ IHC	?	✓ IHC
Kit receptor a (*kita*)	ICC	?	✓ IHC	?	✓ IHC
Motilin receptor (*mlnr*)	intestinal, not localized	?	?	?	✓ qPCR
Ghrelin receptor (*ghsra*)	intestinal, not localized	?	?	?	✓ RT-PCR
Transient receptor potential cation channel, subfamily A, member 1a (*trpa1a*)	?	?	?	?	?
Transient receptor potential cation channel, subfamily A, member 1b (*trpa1b*)	EEC	?	✓ qPCR	?	?

Table compiled by Savanna Sheridan, Hollyn Franklin, Cosmin Serban and Mary Kinkel. Sources: (Ball et al., 2012) (*kita*), (Eom et al., 2014), 2014 (*ghsra*), keller. (*mlnr*), Uyttebroek et al., 2013 (*ano1*), Ye et al., 2021 (*trpa1b*), ZFIN searched March 2021 for *trpa1a* expression.

Another advantage of larval zebrafish includes a transparent body wall that allows imaging of the gut contents in live animals. Part of the overall research was to examine the specific life stages in wild type zebrafish. Examination requires the identification of the cell types that lines the intestine during all stages of zebrafish development. This information will help us to interpret the effects of defects that arise with respect to morphology and function

of the gut. A gut transit assay that takes into account the specific morphology and epithelial composition of the gut at specific life stages would be much more powerful and informative than simply relying on the young larval stage and also allows the ability to differentiate between normal and defective guts.

Intestinal tract in zebrafish

There are key apparent differences between regions of the zebrafish intestine early in development. Larval zebrafish have a straight simplified gut from anterior to posterior, unlike the highly looped mammalian gut (Figure 4).

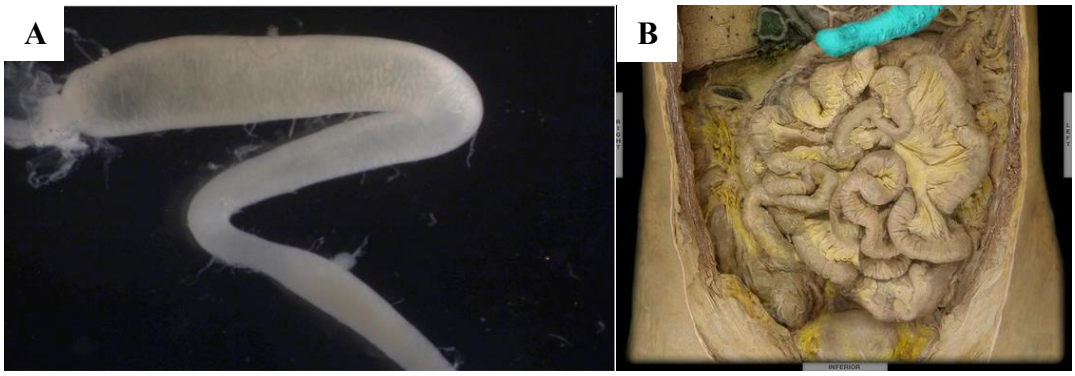

Figure 4. A comparison between the zebrafish and human intestine A) The zebrafish digestive tract is a simple tube that lacks a stomach. B) The human intestine is highly looped. Panel B from quizlet.com.

The region that corresponds with the area of the gastrointestinal tract from the stomach to the anus in humans is subdivided into three main sections in the zebrafish (Wallace et al., 2005). There is an anterior section, known as the intestinal bulb, similar functionally to a stomach; a small intestine that contains the loop in adulthood; and finally, the posterior intestine, where waste is excreted. A key difference in the zebrafish intestine is a lack of sphincters clearly subdividing each of these sections (Wang et al., 2010). The cell type distribution may be the most useful characteristic for understanding the regions of the intestine, especially when it is developing from the larval stage to the adult stage (Lickwar et

al., 2017). Zebrafish provides a high throughput model that allows for the examination of the intestine throughout the life of an organism, and to understand the morphological changes that occur between these two life stages aiming to bridge the gap between larval and adult morphology for a more robust ontogenetic study of the intestine.

Candidate genes from the literature

Part of my research was the identification of candidate genes that are of interest and that relates directly to the zebrafish intestine, and potentially have an involvement in the motility and propulsion of food and waste products. The following information involves the identified genes and descriptions of what is currently known from the literature regarding their function in the GI tract.

Apolipoprotein 1, apoa1

Cholesterol from the diet or that the body produces travels through the blood on proteins called "lipoproteins." Two types of lipoproteins carry cholesterol throughout the whole body: on the one hand, there is LDL (low-density lipoprotein) also known as the "bad" cholesterol, which constitutes most of a human body's cholesterol. On the other hand, there is HDL (high-density lipoprotein), also known as the "good" cholesterol, because it absorbs cholesterol and transports it back to the liver. Apolipoprotein 1 or *apoa1* is the main protein component of HDL particles. *Apoa1* regulates cholesterol flowing out from cells to HDL particles, the first step in transporting peripheral cholesterol to the liver to be excreted in the form of bile (Otis et al., 2015). *Apoa1* is produced by the intestine and secreted as a component of chylomicrons following lipid-rich meals. The *apoa1* gene can be considered a molecular marker for the small intestine because, in 36 human tissues and 45 mouse tissues

examined in previous literature, expression of mammalian *apoa1* was highly restricted to the small intestine and liver (GSE2361 and GDS182, GEO database, NCBI).

In zebrafish, the mRNA expression of *apoa1* has been characterized from embryogenesis of the eight-cell stage and throughout early larval development until 6 days post-fertilization (dpf) (Otis et al., 2015) (Figure 5). Specifically, it was expressed in the yolk syncytial layer (YSL), as shown in Figure 5. YSL processes yolk lipids and secretes very-low-density lipoprotein (VLDL) cholesterol also known as the "bad cholesterol" to support embryonic and larval development (Otis et al., 2015) highlighting VLDLs' importance in transporting lipids from the yolk cell to the developing embryo, and subsequently in the digestive organs including intestine and liver.

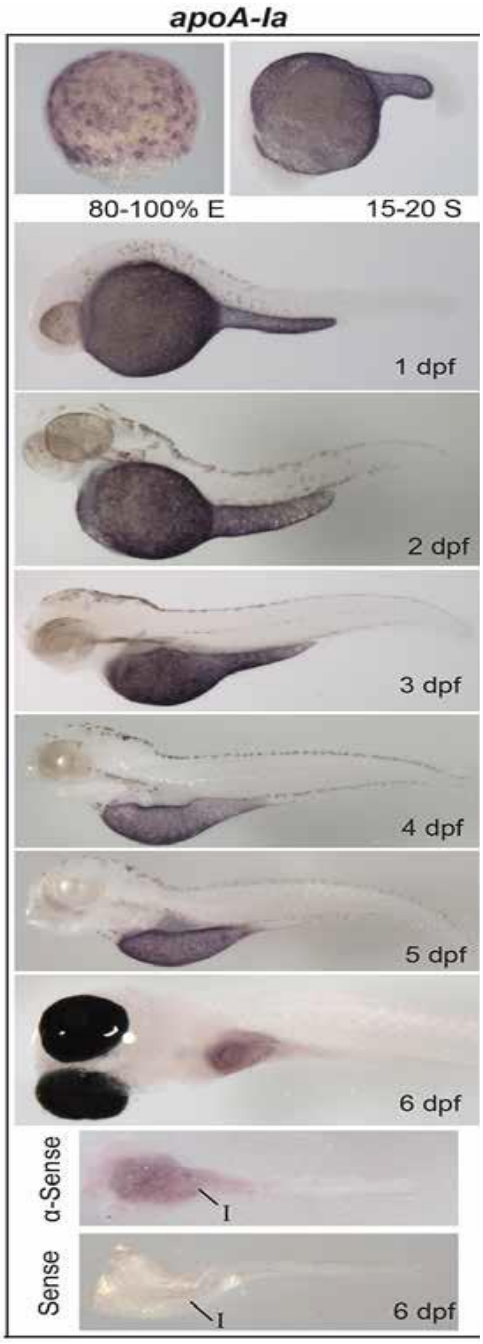

Figure 5. Developmental mRNA expression patterns of *apoa1*. In situ hybridization of *apoa1* during gastrulation (80-100% epiboly, E), somitogenesis (15-20 somites, S), and daily until 6 dpf. (Otis et al., 2015)

As shown in Figure 5, *in situ* hybridization (ISH) was done with antisense and control sense riboprobes specific to apolipoprotein. During very early development, gastrulation (80-100% epiboly) is associated with the earliest expression of *apoa1* mRNA in the YSL. At 80-100% epiboly, *apoaI* mRNA was found to be localized to perinuclear YSL regions; and from 15-20 somites through 6 dpf. As the digestive organs develop further (4-6 dpf), *apoa1* begins to be expressed mainly in the intestine as shown by ISH (Babin et al., 1997). However, *apoa1* expression was not localized to a specific region of the intestine or to a specific cell type but was expressed in the whole intestine according to the whole-mount ISH.

In another study by (Wang et al., 2010) the adult zebrafish gut was dissected and cut into seven equal-length sections labeled S1 to S7 (Figure 6). The cross-section from S1 to S6 contains three tissue layers: mucosa, muscularis externa, and serosa, however, segment S7 has simple epithelium. It was noticed that segments S1-S4 possessed features of a small intestine because *apoa1* was more highly expressed in these sections than the rest of the anterior-posterior axis of zebrafish intestine as shown in Figure 7 (Wang et al., 2010). The *apoa1* gene can be considered a molecular marker for the small intestine as previous literature has shown, and in this case, it was highly expressed in the anterior portion of the zebrafish intestine (GSE2361 and GDS182, GEO database, NCBI).

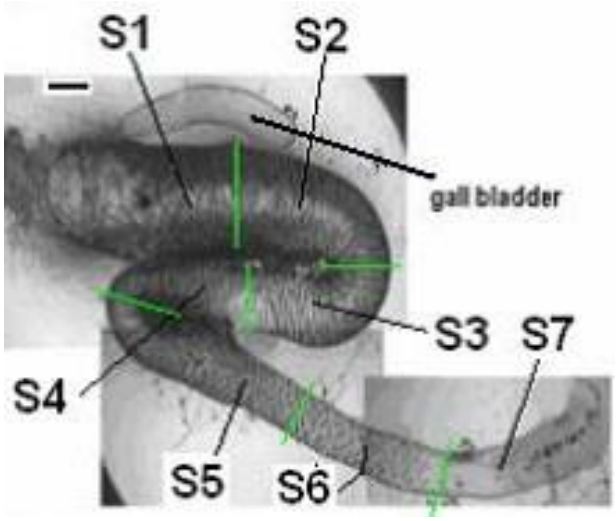

Figure 6. An isolated zebrafish intestine after removal of the surrounding mesentery. The isolated intestine was divided into seven roughly equal-length segments as indicated by green lines denoting S1 to S7. The gall bladder is indicated (Wang et al., 2010).

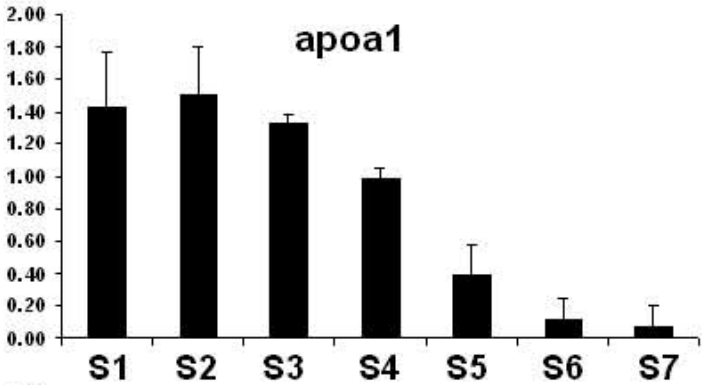

Figure 7. *Apoa1* expression pattern from S1 to S7 intestinal segments. The gene was highly expressed from S1 to S4. (Wang et al., 2010).

Aquaporin 3, aqp3

Aquaporin 3 is an osmoregulatory channel protein on the membrane of epithelial cells especially in the large intestine of mammals. Aquaporin 3 is well known to be a key component of fecal dehydration in the mammalian colon (Silberstein et al., 1999). These

water channel proteins increase water absorption by facilitating water movement and permeability across a cell membrane (Lance et al., 2004). Since it is known that water is absorbed in the large intestine mostly, one can predict that *aqp3* would be highly expressed in that region of the intestine more than other regions. Wang and colleagues examined this gene's expression in six-month-old zebrafish (Wang et al., 2010). Their microarray data and RT-PCR agrees with the mammalian expression of *aqp3*. Their data shows that this gene is highly expressed in segments S6 and S7, similar to the colon in mammals as shown in Figure 8.

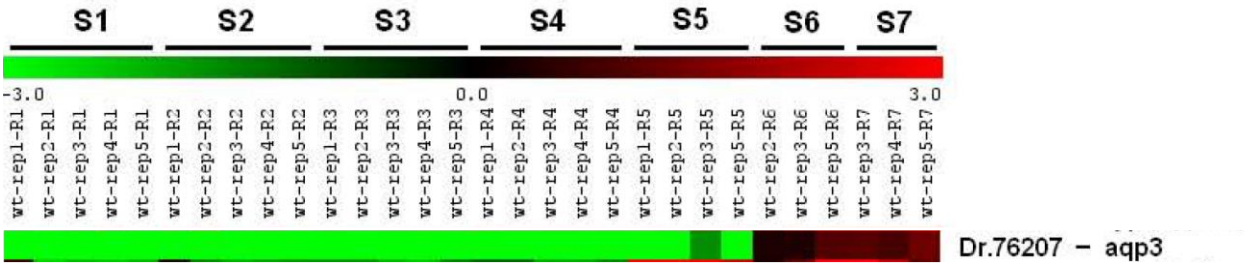

Figure 8. Expression pattern of selected intestinal gene (*aqp3*). Expression pattern of *aqp3* gene based on microarray data. The red area indicates positive high expression. The green area indicates low expression. (Wang et al., 2010).

Cathepsin, ctsl1

Cathepsins are lysosomal cysteine protease enzymes that can induce apoptosis, catabolize intracellular proteins, degrade the extracellular matrix, and process prohormones into active forms. Song and colleagues examined the temporal and spatial RNA expression of *Ctsl1* in the placentae and endometria of pigs (Song et al., 2010). In situ hybridization analysis of *Ctsl1* mRNA in the small intestine from a piglet revealed expression in the absorptive epithelium (enterocytes) of villi and, to a lesser extent, the cryptal epithelium on Day 1 postpartum. In fact, it may be involved in the transport of IgG antibodies from

colostrum across the intestine epithelium, primarily by enterocytes lining the intestine, into the blood of piglets during the first 24 to 36 h postpartum to establish passive immunity. This could suggest a direct role for *ctsl1* in mechanisms for fluid-phase pinocytosis and transport of proteins during transport across placental and gut epithelia (Song et al., 2010). Wang and colleagues examined the *ctsl1* gene's expression in a six-month-old zebrafish with its dissected intestine into seven segments (S1-S7) as shown in (Figure 6). *Ctsl1* was highly expressed in caudal segments of the intestine from S5 to S7 (Wang et al., 2010) (Figure 9).

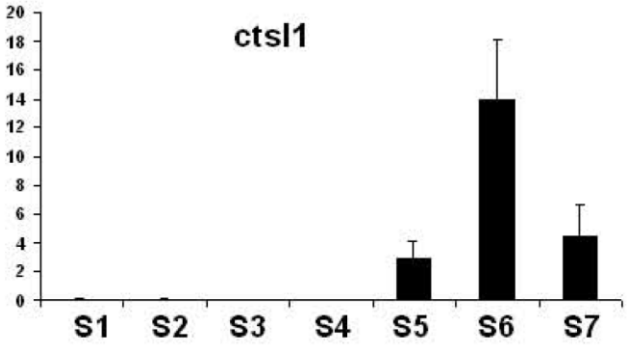

Figure 9. *ctsl1* expression pattern from S1 to S7 intestinal segments. The gene was highly expressed from S5 to S7. (Wang et al., 2010).

Cofilin 1, cfl1

Cfl1 belongs to a family of actin-binding proteins and mediates dynamic stabilization of actin filaments. According to a previous analysis of rat EST database, it suggests that *cfl1* is expressed in the large intestine but not in the small intestine (Unigene EST profile viewer, Unigene Rn.11675, NCBI). This gene was investigated also by Wang and his colleagues and their methods of dividing the gut of a six-month-old (adult) zebrafish into seven segments as mentioned previously. Their real-time RT-PCR data (Figure 10) indicates that *cfl1* is primarily expressed in segments S5-S7, but down-regulated in the first four

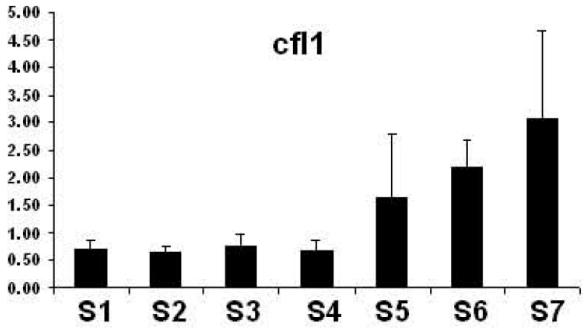

Figure 10. *cfl1* expression pattern from S1 to S7 intestinal segments. The gene was highly expressed from S5 to S7. (Wang et al., 2010).

segments of the intestine. Therefore mimicking its expression patterns in the rat where it is expressed in the large intestine, and thus a known large intestine gene is expressed in S5-S7 according to Wang and colleagues.

In a separate study by Lin and colleagues, the *cfl1* gene was characterized further by investigating the spatial and temporal expression of it during embryogenesis as shown in figure 11. To test the importance and the role of cofilin-1 during gastrulation, the researchers knocked this gene down. The loss of cofilin-1 disrupted the cadherin-dependent adhesion between different epithelial cell layers during gastrulation in zebrafish, as shown in Figure 12 (Lin et al., 2010).

Fatty acid binding protein 2, fabp2

The intestinal *fabp2* gene encodes a fatty acid-binding protein that is specifically involved in reversibly binding the hydrophobic ligands and transporting them throughout cellular compartments, including the peroxisomes, endoplasmic reticulum, mitochondria, and nucleus in the small intestine (Smathers and Petersen, 2011). Hydrophobic ligands include fatty acids (FAs) and their acyl-CoA derivatives (FA-CoA), serve many biological functions

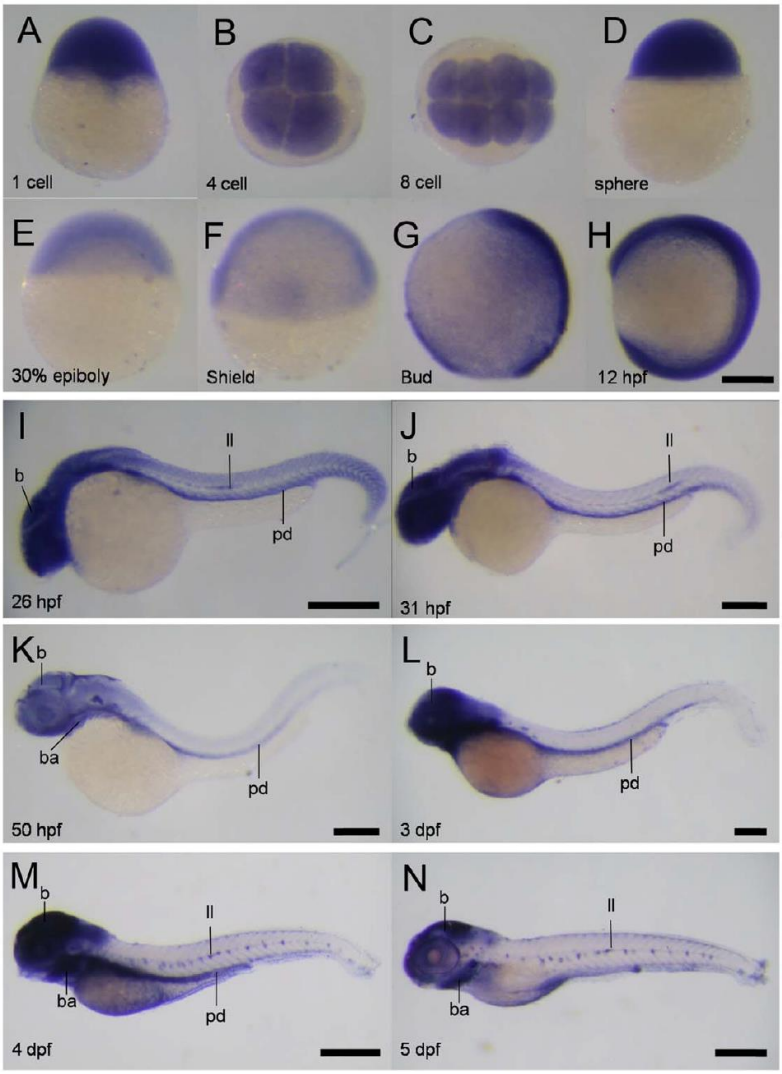

Figure 11. Spatial and temporal expression of *cfl1* during development. Representative in situ hybridization results reveal the expression patterns of *cfl1* from 1-cell to 5 dpf (Lin et al., 2010).

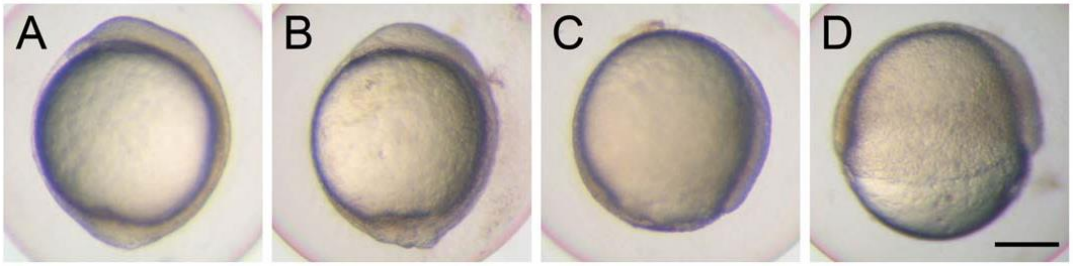

Figure 12. Knockdown of *cfl1* causes epiboly defects. Embryos imaged at 10 hours post-fertilization (hpf). The *cfl1* MO caused epiboly defects of different severities. A) Sham-injected embryo reached 100% epiboly. B) *cfl1* morphant with malformed tail bud. C) *cfl1* morphant that reached 90% epiboly. D) A *cfl1* morphant that reached 50% epiboly. Scale bar, 200 mm. (Lin et al., 2010).

such as metabolic energy sources, signaling molecules for internal and external cell signaling as well as serving as substrates within the cell (Martinez-Lopez et al., 2015).

Based on previous literature, the *fabp2* gene is highly conserved across many species including mammals, teleosts, amphibians, and avians. In earlier literature, a Red Fluorescent Protein (*RFP)* transgenic zebrafish line under the intestinal *fabp2* promoter, *Tg(fabp2:RFP)*, was generated, and the RFP reporter gene is specifically expressed in the intestine (Wang et al., 2010). In previous work done by (Wang et al., 2010), the *fabp2* expression was further studied by isolating an intestine from a 3-month-old *Tg(fabp2:RFP)* transgenic zebrafish. It was discovered that RFP fluorescence, a marker for *fabp2* gene expression, was high in segments S1-S4, but quickly diminished around the second turn of the intestine at the level of segment S5. This expression pattern was also confirmed by direct detection of endogenous *fabp2* mRNA expression by *in situ* hybridization (Figure 13).

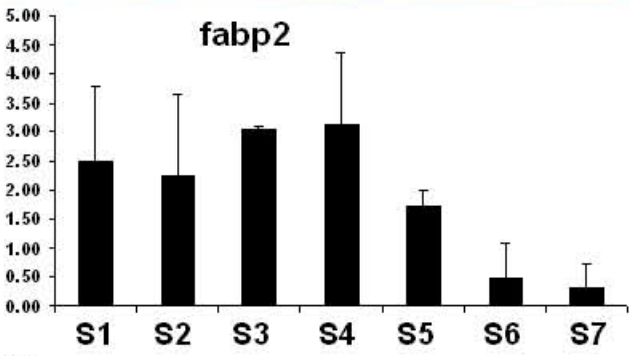

Figure 13. *Fabp2* expression pattern from S1 to S7 intestinal segments. The gene was highly expressed from S1 to S4. (Wang et al., 2010).

In a separate study (Venkatachalam et al., 2013), the gene expression of *fabp2* was studied to verify its response to several different diets containing different types of fatty

acids. Five-month-old zebrafish were fed the fatty acids then the intestine and the liver were dissected out and the expression of *fabp2* was studied. The expression of *fabp2* mRNA did not change in the liver (Figure 14A) of zebrafish fed any of the fatty acid diets. The expression level of *fabp2* mRNA was noticeably higher in the intestine due to the high levels of fatty acids (Figure 14B).

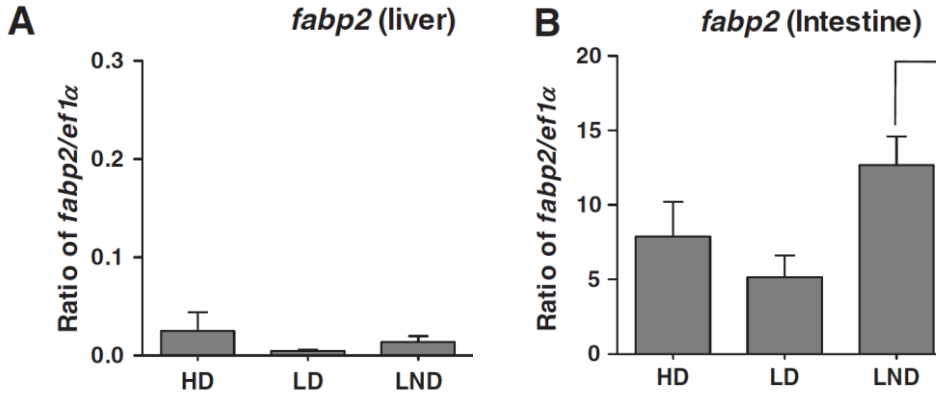

Figure 14. The expression levels of *fabp2* mRNA in liver and intestine in response to diet. Zebrafish were fed diets differing in fatty acid content including a highly unsaturated FA-rich diet (HD), linoleic acid-rich diet (LD), and linolenic acid-rich diet (LND). The expression levels of *fabp2* mRNAs in the liver were minimal when compared to the intestine. (Venkatachalam et al., 2013).

Triacylglycerol lipase, lipf

Lipf encodes gastric lipase, an enzyme in humans involved in the digestion of dietary triglycerides in the GI tract, and is responsible for 30% of fat digestion processes. It is known to be secreted by mammalian gastric chief cells (Armand, 2007). It is only expressed in the esophagus, stomach, and several other tissues in humans, but not in the intestine (Unigene EST profile viewer, UniGene Hs.523130, NCBI database). However, according to Wang and colleagues, the gene was identified as being expressed in the whole gut of a six-month-old zebrafish (adult) intestine from segments S1 to S7 (Figure 15), using the intestinal segmentation the researchers employed in their study (Wang et al., 2010).

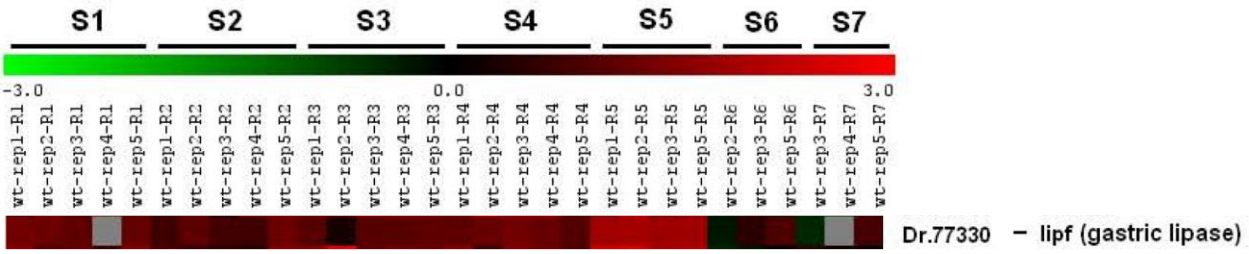

Figure 15. Expression pattern of selected intestinal gene (*lipf*). Expression pattern of *lipf* gene based on microarray data and following the seven segments S1-S7 as shown in Figure 6. The red area indicates positive high expression. The green area indicates low expression. (Wang et al., 2010).

Villin1, vil1

Villin is a protein that regulates actin that is associated with the formation of microvillar actin filaments and is expressed in significant amounts in the gastrointestinal epithelial cells of mammals (Bretscher and Weber, 1979). Villin regulates epithelial cell motility, and cell morphology, and actin reorganization. In the research by Wang and colleagues in (Wang et al., 2010), the expression of *vil1* resembled that of *fabp2* expression being highly expressed in segments S1 to S5 while being less expressed in segments S6 and S7 according to the RT-PCR (Figure 16).

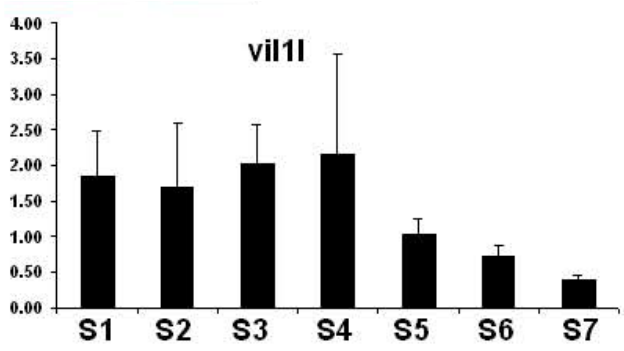

Figure 16. *vil1* expression pattern from S1 to S7 intestinal segments. The gene was highly expressed from S1 to S5. (Wang et al., 2010).

In a separate study done by (Zhao et al., 2018), PIK3C3 was being studied as a cause for lethality with inflammatory bowel disease-like features in zebrafish. PIK3C3 is a class III phosphatidylinositol-3-kinase that catalyzes the production of phosphatidylinositol 3-phosphate (PI3P) from PI, which is a critical component of the vesicular membrane. It was reported in the study that PIK3C3 deficiency in zebrafish suppresses PI3P synthesis in intestinal epithelial cells (IEC), induces a proinflammatory response in the intestine, and causes epithelial damage. Total RNA samples were prepared from whole, wild-type and transgenic zebrafish fish line (Tg(mpx:EGFP) and Tg2(mpeg1:EGFP)) at 6, 7, and 8 dpf, and the samples were sequenced using RNA-Seq. The RNA-Seq data revealed that villin 1 *(vil1)*, an IEC-specific gene, was down-regulated in the pik3c3 transgenic fish line. Additionally, *in situ* hybridization analysis was performed which revealed that the expression of intestinal v*il1* was reduced in the mutant larvae, as shown in Figure 17.

There are many genes that may have a direct or indirect role in the normal functionality of the zebrafish intestine. In this review of the literature, I have explored candidate genes that were chosen based on their previously known expression in zebrafish or other organisms. I have learned about the spatial and temporal expression of these genes in the zebrafish intestine which could lead to further information about the normal function of the gut and could lead to investigations of intestinal motility diseases with specifics on where and how they may arise.

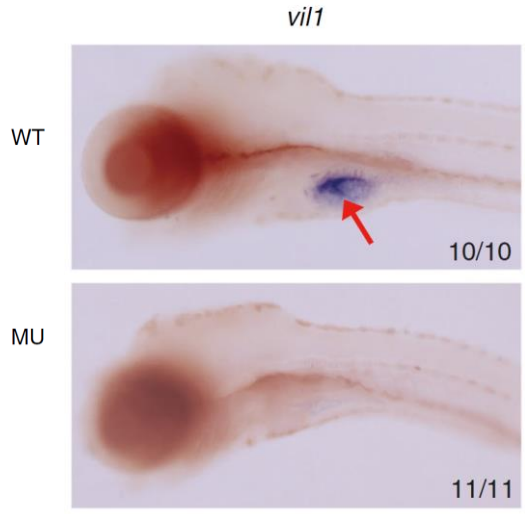

Figure 17. In situ hybridization of 8 dpf larvae with a *vil1* probe. pik3c3 mutants compared with wild-type zebrafish. Red arrow points to *vil1* expression in the intestine. Numbers indicate larvae with the indicated phenotype/total analyzed. (Zhao et al., 2018).

Objectives of This Study

The primary purpose of this research is to aid in the establishment of zebrafish as a model organism for intestinal motility disorders in humans. Using the literature, I identified candidate genes that are of interest based on their previously known expression in the zebrafish GI tract. I studied the spatial and temporal expression of these genes to learn more about their functional and anatomical aspects in the GI tract. To better understand the genes' conservation with other organisms such as humans and mice, I aligned the DNA sequences of all three organisms using multi-sequence alignment tools. In order to study the motility of zebrafish GI tract, and to assist in the development of a gut transit assay, I performed a feeding behavior study to determine the age at which zebrafish larvae are mature enough after their fertilization to capture brine shrimp. A gut transit assay utilized this information to determine the best age to perform this assay at, which is a way to quantify the time it takes

for food and waste products to pass through the whole GI tract of a zebrafish from feeding to clearing. The results of this study and its research will contribute to the overall knowledge of gastrointestinal diseases, and may contribute to the development of either new or enhanced medical and surgical cures related to patients diagnosed with gastrointestinal diseases.

Materials and Methods

Zebrafish Husbandry

Zebrafish (*Danio rerio*) were maintained according to standard procedures (Westerfield, 2000), and were housed and maintained in the Appalachian State University vivarium. To conduct experiments on larval zebrafish including the feeding behavior study, adult zebrafish were bred by having several male and female in breeding tanks, and fertilized eggs were collected. Eggs were transferred into a glass dish filled with 120 mL of E3 medium (containing 0.17mM KCl, 0.33 mM $MgSO_4$, 0.33 mM $CaCl_2$, 5 mM NaCl) with 0.01% methylene blue. The eggs were bleached to prohibit fungal growth at around 24 hours after the eggs' fertilization. Eggs were bleached with a solution of bleach and facility water, and rinsed with facility water three times. The bleached eggs were then transferred and separated into 50 eggs per bowl containing fresh E3 medium and maintained at 28.5ºC in an incubator

Beginning at 5 dpf, larval zebrafish were transferred to fish tanks and raised in a tabletop nursery, following the protocols described in Norton et al., 2019. Briefly, wild-type zebrafish larvae were raised in tabletop nurseries using larval tanks containing 280 mL of E3 medium and 50 zebrafish per tank kept at approximately 28ºC. The larval tanks were cleaned using a siphoning method which involves partially replacing the E3 medium to remove waste. Siphoning occurred daily approximately 15 minutes before 9 a.m. and 3 p.m. Feeding included both dry food Golden Pearl (GP) and Larval Diet (Brine Shrimp Direct), and newly-hatched brine shrimp *(Artemia franciscana).* Dry food was given to each tank at 9 a.m. and 3 p.m. while brine shrimp were given at 12:00 p.m. daily. At 21 dpf, the larvae were moved to the fish room with recirculating water at the Appalachian State University animal facility.

Until the zebrafish reached adulthood, they were fed dry food at 9 a.m. and 3 p.m., and 48-hour brine shrimp cultures at 12:00 p.m. Once they became adults, they were fed dry food at 9 a.m. and 48-hour brine shrimp at 3 p.m. All procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at Appalachian State University.

Candidate Gene Analysis

Candidate genes were identified by searching for previously identified genes in the literature using the Unigene database from the National Center for Biotechnology Information (NCBI). Default settings were used to search each gene of interest. EST profiles were analyzed to obtain the desired gene expression of the cDNA. After the identification of the candidate genes, a database search was done on each of the accession numbers to further learn more about them.

Multiple Sequence Alignment

Candidate genes were analyzed using Clustal Omega to compare and differentiate between zebrafish, *Homo sapiens*, and *Mus musculus*. The DNA sequence for each candidate gene was obtained from the nucleotide NCBI reference sequence database. The sequences were then uploaded and aligned using the Clustal Omega MSA tool.

Primer Design

Primers to be used for RT-PCR of candidate genes were designed using NCBI Primer3 and BLAST. Using the default settings in Primer3, introns were excluded, and the primer search was performed. Table 2 shows a list of primers for RT-PCR.

Table 2. Primers for the genes of interest. All primers are oriented 5'-3'.

Gene	Primer sequence	Size (bp)
*apoa1*F	CCA ATT TGT TCC AGG CTG AT	233
*apoa1*R	CAA CTG GGT GGA GAT GGT CT	
*aqp3*F	TCC CTG ATT GTG TGC ATC CT	156
*aqp3*R	TCC AAA ATC TCG AGC TGG GT	
*ctsl1*F	TGA GTG AGC AGA ACC TGG TG	152
*ctsl1*R	GCA GGA TTC ATC CTT TCC AA	
*cfl1*F	TGT TGA CGG TCT TCA ACG AG	199
*cfl1*R	CAA TCA TTG GGA GGA AGC AT	
*fabp2*F	CAA CGT GAA GGA AGT CAG CA	163
*fabp2*R	TAC CTT TCC GTT GTC CTT GC	
*lipf*F	GTC GAC CCC TCC ACT TTA CA	175
*lipf*R	CCC CCA GAT AAA ATC CAG GT	
*vil1*F	AGG AGA AGC AGA GGT GTG GA	238
*vil1*R	CTG CAC CGG TTC TCC ATT AT	

Larval Feeding and Anesthesia

After overnight fasting to clear gut contents, larvae were fed in the morning dry food mix that is free of colored supplements such as Spirulina. This ensured that the gut contents could be distinguished from brine shrimp. At noon, larvae were fed brine shrimp. After 15 minutes, larvae were anesthetized by gradually lowering water temperature to approximately 13-17℃ using ice made of E3 medium (containing 0.17mM KCl, 0.33 mM $MgSO_4$, 0.33 mM $CaCl_2$, 5 mM NaCl). This allowed the larvae to be immobile for short periods, allowing imaging to be completed.

Live Imaging of Intestinal Contents

Anesthetized larvae were transferred to an agarose mold made from 3% agarose in E3 medium. A wide bore fire-polished Pasteur pipet was used for the transfer. Larvae were submerged in warm E3 medium. Larvae were imaged using an Olympus SZX12 stereoscope and a Canon 70D camera. Imaging involved capturing the whole intestine of each larva whether it had captured a shrimp or not. Immediately after imaging, larvae were returned to a tank of fresh, 28.5ºC E3 medium and monitored for swimming behavior. The study was conducted from 6 dpf to 15 dpf. . If a larva captured shrimp, this was noted and counted for. Finally, the percentage was determined for each age group based on whether a larva captured a shrimp or not.

Results

Candidate gene expression patterns

Understanding the anatomy and physiology of the gastrointestinal tract is crucial for studies of digestive diseases. Learning about the genes that are expressed in the GI tract and how their expression provides functionality to the GI tract with regard to motility, digestion, and absorption is of utmost important.

Part of my research is identifying candidate genes that are relevant for understanding normal gut motility. This is done by using literature searches to identify intestinal genes of interest that were previously identified to be expressed in the zebrafish GI tract. I have identified several candidate genes from the literature. These genes are listed in Table 3 below.

Table 3. Genes of interest expression patterns. The location data and function refer to what is known for zebrafish intestinal tissues including spatial and temporal expression patterns.

Gene Name	Gene Symbol	Definition/Function	Intestinal Expression	Stages
Apo-lipoprotein 1	*apoa1*	Main protein component of HDL: Cholesterol metabolism	Adult intestinal cells, anterior segments (by qPCR, RT-PCR)	Embryonic to 6 dpf; adult
Aquaporin 3	*aqp3*	Water channel protein in large intestine: Fecal dehydration	Intestinal cells (by qPCR, RT-PCR)	Adult
Cathepsin L1	*ctsl1*	Lysosomal cysteine protease: Roles in fluid phase pinocytosis and transport of protein across gut epithelia	Intestinal cells, posterior segments (by qPCR, RT-PCR)	Adult
Cofilin 1	*cfl1*	Actin-binding protein: Mediates dynamic stabilization of actin filaments by inhibiting depolymerisation	Adult intestinal cells, posterior segments (by qPCR, RT-PCR) and microarray; different intestinal epithelial layers (by in-situ hybridization)	Embryonic to early larval (0-72 hpf); adult
Fatty Acid Binding Protein 2	*fabp2*	Intracellular protein: Transport of fatty acids in the small intestine	Intestinal cells, anterior segments (by qPCR, RT-PCR) and microarray	Adult
Triacyl-glycerol lipase	*lipf*	Gastric acidophilic lipase: Digests lipids	Intestinal cells, whole segments (by qPCR, RT-PCR) and microarray	Adult
Villin 1	*vil1*	Crossfilament protein: Links microfilament core to the inner side of the microvillus membrane	Larval epithelial cells (by RNA-Seq); Adult intestinal cells, anterior segments (by qPCR, RT- PCR) and microarray	Early larval (3-9 dpf), adult

Sources: *apoa1*,(Otis et al., 2015); (Wang et al., 2010); *aqp3, ctsl1, lipf,* (Wang et al., 2010)*; cfl1,*(Lin et al., 2010) ; (Wang et al., 2010); *fabp2,* (Venkatachalam et al., 2013); (Wang et al., 2010); *vil1*, (Wang et al., 2010)*;*(Zhao et al., 2018).

The results show that while some of the candidate genes were found to be expressed at the embryonic to larval stage, other genes are expressed later at the adult stage. This implies that some gut functions are absent in the early larvae but becomes apparent when the larvae becomes an adult. Thus, motility could be indirectly affected if a gene is not expressed early on at the larval stage. In order to explore if the same genes have similar sequences as in humans and mice, a multiple sequence alignment was done next. This could lead into exploring if these genes have possibly the same functions, and are expressed at similar stages in different organisms.

Multiple sequence alignment (MSA) shows degree of similarity

To determine the conservation of the candidate genes across multiple organisms, I compared the sequences for humans (*Homo sapiens),* mice (*Mus musculus),* and zebrafish (*Danio rerio)*. This study involved the research and identification of the nucleotide sequences of each gene for each of the organisms to compare DNA sequences to better understand how the gene sequences are conserved across the three organisms, and thus better understand their relationships. To do this step, Clustal Omega was used to compare the sequences using multiple sequence alignments. I was able to obtain the multiple sequence alignments for *apoa1, aqp3, ctsl1, cfl1, fabp2, lipf, and vil1* candidate genes, shown in Appendix A. The results in Table 4 below indicate that the percent identities of the seven genes of interest across the three organisms were between 62% and 86% for each gene when their sequences were compared.

Table 4. Multiple sequence alignment percent identity. The nucleotide sequences of each gene were aligned across three organisms: humans (*Homo sapiens),* mice (*Mus musculus),* and zebrafish (*Danio rerio).* The percent identity was identified for each alignment. Results obtained from NCBI and Clustal omega.

Gene Name	Gene Symbol	Percent Identify
Apo-lipoprotein 1	*apoa1*	75.07%
Aquaporin 3	*aqp3*	62.48%
Cathepsin L1	*ctsl1*	76.76%
Cofilin 1	*cfl1*	86.29%
Fatty acid binding protein 2	*fabp2*	72.06%
Triacyl-glycerol lipase	*lipf*	77.86%
Villin 1	*vil1*	79.39%

From the results of the multiple sequence alignments, I was able to conclude that all seven candidate genes are conserved across the three organisms. In future studies, it would be important to identify these genes expression's location using the identified sequence-specific primers in the GI tract of zebrafish using reverse transcription-polymerase chain reaction (RT-PCR). This would allow researchers to compare the genes' expression locations with their known functions to better understand their role in the gastrointestinal motility in zebrafish and humans.

Given the results from the multiple sequence alignments, it can be concluded that studying these candidate genes in zebrafish would be important to understand their possible role in intestinal motility. Therefore, a gut transit assay would be a foundation to be able to test a normal versus an abnormal intestinal motility in zebrafish. To be able to conduct a gut transit assay, the optimal zebrafish age needs to be determined. The best approach to accomplish this is a feeding behavior study. Some of the candidate genes are not expressed in the early larval stages as mentioned previously as part of this research, therefore it would be crucial knowing if their absence would alter the GI tract's motility.

Feeding behavior study results

For the purposes of better understanding the function of the GI tract, zebrafish were used as a model to investigate the time it takes from feeding to clearing the intestine for normal wild type zebrafish. The results of the passage time can then be compared to a mutant zebrafish with a mutation in a specific gene of interest. This experimentation is known as a gut transit assay. This assay helps in the understanding of normal and abnormal motility of the GI tract.

Brine shrimp were used as a meal because it is a normal and digestible meal for zebrafish that can be easily detected under the microscope when it is consumed due to the transparency of zebrafish larvae and the ingested brine shrimp's orange color. In order to conduct a gut transit assay, the earliest and most efficient age at which zebrafish larvae are mature enough to capture brine shrimp needs to be determined. Several feeding behavior experiments were conducted for comparative purposes of the different feeding behavior results. The experiments were focused between 6 and 15 dpf. Earlier than 6 dpf, the larvae are too immature to be able to capture brine shrimp (Clift et al., 2014). At noon each day, larvae were fed brine shrimp. After 15 minutes, fish were anesthetized by gradually lowering water temperature using ice, then imaged under the microscope. Larvae were scored as to if they captured or did not capture brine shrimp as indicated by the appearance of orange gut.

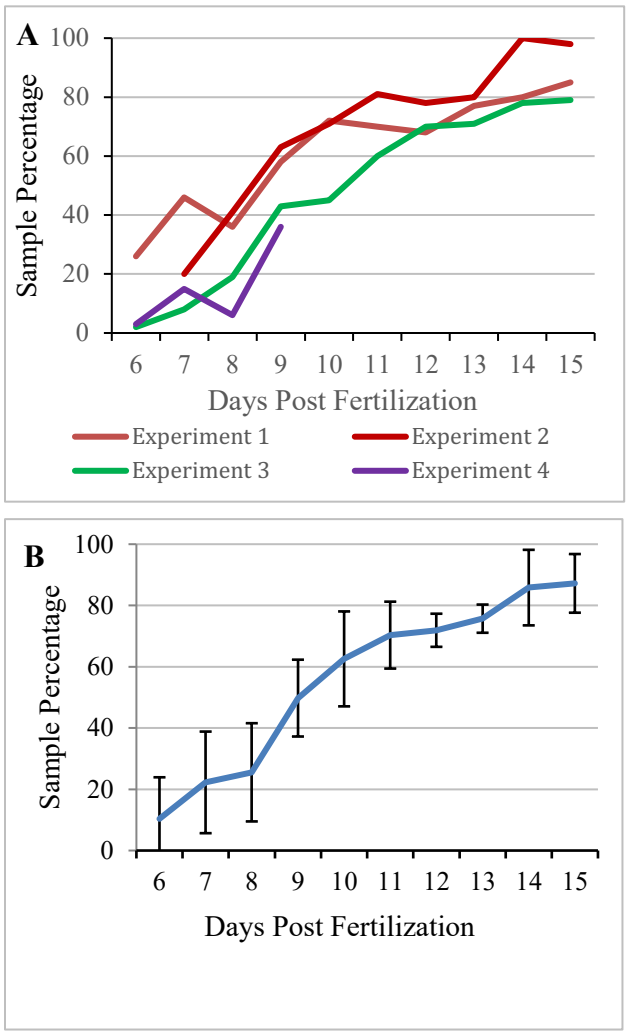

Figure 18. Larval zebrafish shrimp capture ability. A) The averages of the four experiments are shown individually. Sample sizes ranged from 30 to 50 larvae (see Table 6). B) The average percentage of larval zebrafish that captured brine shrimp across four separate feeding behavior studies combined in the lab. Error bars indicate standard deviation.

To investigate the shrimp capture results more closely, four independent experiments including my experiment and three additional similar experiments by three other persons in the lab were conducted and were plotted separately, as shown in Figure 18A. Sample sizes ranged each dpf from 30 to 50 larvae as shown in Table 5. Experiment 1 shows that at 9 dpf, more than half of the larvae captured shrimp. Experiment 2 had similar results. However, experiments 3 and 4 were both under fifty percent at 9 dpf. For instance, experiment 3

showed only a 43% shrimp capture rate and experiment 4 showed 36%. Additionally, the overall shrimp capture success was lower in experiments 3 and 4 in the earlier dpf range before 9 dpf.

Figure 18B demonstrates the average percentage of larval zebrafish that captured brine shrimp across the four separate feeding behavior studies. At 9 dpf, all four experiments converged at an average of fifty percent of the percentage of larvae that captured brine shrimp. Additionally, at 9 dpf the precision of the results noted by the standard deviation was smaller than at 6, 7, 8, and 10 dpf. This smaller deviation suggests that the results for the four experiments agree that 9 dpf is the earliest day at which a gut transit assay could be conducted with a reasonable shrimp capture success rate.

Table 5. The sample sizes used in the four feeding behavior experiments.

DPF	Experiment 1	Experiment 2	Experiment 3	Experiment 4
6	50	--	48	30
7	50	35	50	48
8	50	37	47	51
9	50	32	47	42
10	50	38	47	--
11	50	43	47	--
12	31	41	40	--
13	31	39	34	--
14	25	40	40	--
15	26	44	38	--

Discussion

The ultimate goal of this research project was to aid in the understanding of gastrointestinal motility diseases and to help establish zebrafish as a model organism for such diseases, and their etiologies. This study identified candidate genes that may be directly or indirectly involved in GI tract motility, and a defect in one or more of these genes might make the GI tract dysfunctional as seen in IBS and Hirschsprung's disease. The candidate genes identified from the literature were found to be conserved with humans using the multiple sequence alignment studies, thus verifying them as possibly involved directly in GI tract function or dysfunction. This study also established a suitable age of larval zebrafish for conducting a gut motility assay using live, intact larvae. This assay can be used in the future to study gut motility by monitoring GI tract movements from feeding to emptying.

Candidate gene expression patterns

The gastrointestinal tract involves many different cell types that vary in anatomy and physiology. Differential gene expression is observed in specific intestinal cell types in zebrafish (Otis et al., 2015). The spatial location of where a gene is expressed is crucial to the proper function of each region of the GI tract. Consistent with this, the identified genes from the literature showed specific segmental expression in the adult zebrafish GI tract. Some of the candidate genes were expressed mainly anteriorly in the intestine, others posteriorly, while others were expressed throughout the whole intestine (Wang et al., 2010). This different expression might correspond to specific cell types that have different roles in gut function including motility. For example, genes that are expressed mainly anteriorly, in the

intestinal bulb might be more involved in functions related to digesting a meal. Genes that are expressed posteriorly, in the colon, likely have roles in nutrient absorption.

Temporal information is important for giving rise to specific cell types with specific functions at a specific point in time during the zebrafish larval development. Therefore, it can be learned that if a gene is not expressed in zebrafish until the adult stage, certain cell types with specific functions will not be detected in the gastrointestine, and thus cannot contribute to specific function such as smooth muscle movements, also known as peristalsis allowing GI tract motility. Therefore, a motility dysfunction in the GI tract will depend on specific mutations in specific genes.

Feeding behavior study results

Gut transit assays allow us to study the peristaltic movements of the digestive system. Ideally, we can study motility from the time food is consumed until the time it is cleared as waste products. This would allow us to observe the effects of specific mutations on the peristaltic movements of the gastrointestinal tract. In order to develop a gut transit assay using larval zebrafish and a controlled meal size, the earliest age at which zebrafish larvae are mature enough to capture a brine shrimp needed to be determined. Therefore, I conducted a feeding behavior study to determine at what age most larvae could capture a brine shrimp.

The feeding behavior study results shown previously in figure 20 involved four independent trials of the same experiment, conducted by four different people. It is worth noting that the overall shrimp capture success was lower in experiments 3 and 4 for larvae younger than 9 dpf, compared to the results for experiments 1 and 2. This difference in results could be because different people used different strategies for delivering shrimp to the tanks. Later tests of this (not shown here) determined that this was the case. For experiments

1 and 2, the shrimp were delivered to the tanks in a zig-zag fashion from front to back. By contrast, in experiments 3 and 4, the shrimp were delivered to the center of each tank. It was noticed that experiments 1 and 2 had a better shrimp capture success rate than experiments 3 and 4 during the early days of the experiments from 6 dpf to 12 dpf. From this, we propose that immature larvae tend to swim less towards the brine shrimps. There are several possible reasons for this such as their inability to swim fast at this age to compete with other larvae, or their vision along other senses are not fully developed, or the possibility that they are not used to brine shrimp at this age. Thus, the capture rate is lower. However, as the larvae's age increase, so too does their brine shrimp capture rate because they are swimming rapidly towards the meal at this age with possibly more mature visual and olfactory senses. In support of this claim, at 12 dpf and beyond, the experiments coincide with each other regardless of the feeding method used. The conclusion of the feeding behavior study showed that at 9 dpf, all four experiments converged at an average of more than fifty percent of the percentage of larvae that captured brine shrimp. Therefore, 9 dpf would be the earliest day for efficiently conducting a gut transit assay using brine shrimp as a meal.

Multiple sequence alignment (MSA) shows degree of similarity

The ultimate goal of this research is to study gastrointestinal (GI) motility disease such as Hirschsprung's disease, and Irritable bowel syndrome (IBS) using live organisms with the possibility of applying findings and concepts to the human GI tract. As a result, surgical and pharmaceutical treatments could be found and applied. In the case of this research, zebrafish was used as a model organism to study such diseases. However, in order for scientists to use information found in this research, such as the identified candidate genes from the literature, the link that connects both organisms, humans and zebrafish, needed to be

identified. In the case of genes identified as expressed in the GI tract of zebrafish, the same genes would ideally show conservation with human genes. My multiple sequence alignments showed the genes to be between 62 to 86% conserved across humans, mice, and zebrafish. This similarity suggests that further study of these genes that are known to be expressed in the zebrafish GI tract will provide useful insights for understanding gut function. Future research on these genes may reveal connections with gut motility, digestion, and absorption, and can therefore be directly beneficial to aid the research of GI tract motility diseases in humans. For instance, if a gene is found to be mutated and results in loss of function in the zebrafish intestine, it can be said that this could be the case with the conserved gene found in humans. This assumption could be further explored with additional research.

Future Directions

Intestinal Transit Assay

To study gastrointestinal motility diseases, zebrafish is used as a model organism. The transparency of a zebrafish gut allows the ability to carefully monitor the passage of food and waste products. Mainly, motility comprised by several peristaltic movements is studied. The lab has developed the idea to use an intestinal transit assay which allows us to monitor the duration it takes for a bolus to travel from feeding to clearing in the zebrafish gut. The feeding behavior studies allowed the identification of the best day post-fertilization of zebrafish to conduct such an assay. The objective of such an assay is study the duration it takes for a bolus to travel the gut in wild type normal zebrafish and compare it to a mutant zebrafish with knocked down specific gene. This will aid in the overall research of gastrointestinal diseases.

Previous gut transit assays from published studies showed that it took 1 hour to feed the zebrafish, and up to 24 hours or more to completely excrete the waste products (Field et al., 2009). These studies used plastic florescent beads mixed into the meal and the beads could have caused constipation in the fish (Field et al., 2009). Therefore, it is suggested that this assay be repeated using digestible meals such as brine shrimp that the fish consume in their habitat.

Dissection

In order to more thoroughly understand the seven candidate genes and their role in the gastrointestinal tissue of zebrafish, the GI tract can be dissected out, the various anterior-posterior regions of the gut can be separated, and nucleic acids can be isolated (Eames Nalle et al., 2017). The dissection should follow the different anatomical regions of the GI tract to isolate four segments: intestinal bulb, intestinal loop, small intestine, and the colon. This will allow isolation of RNA samples from each of the major regions of the intestine. RT-PCR could then be performed using gene-specific primers.

RT-PCR

In order to perform RT-PCR, cDNA will be reverse transcribed from each RNA sample (Bustin, 2004). A set of PCR primers has already been designed for each of my candidate genes. The purpose of this future direction is to determine which anatomical segment of the gut expresses the genes of interest. Previous gene expression work from (Wang et al., 2010) had dissected the zebrafish intestinal tract into seven equal-length segments. While this approach resulted in a useful gene expression map, it ignored anatomical distinctions between the major regions of the gut. Therefore, potential functional differences cannot be clearly determined from the previous work. Dissecting the zebrafish

intestine based on anatomical landmarks will allow future researchers to better understand gene function.

References

Armand, M. 2007. Lipases and lipolysis in the human digestive tract: where do we stand? *Curr. Opin. Clin. Nutr Metab. Care*. 10:156-164.

Babin, P.J., C. Thisse, M. Durliat, M. Andre, M.A. Akimenko, and B. Thisse. 1997. Both apolipoprotein E and A-I genes are present in a nonmammalian vertebrate and are highly expressed during embryonic development. *Proc. Natl. Acad. Sci. U. S. A*. 94:8622-8627.

Ball, E.R., M.M. Matsuda, L. Dye, V. Hoffmann, P.M. Zerfas, E. Szarek, A. Rich, A.B. Chitnis, and C.A. Stratakis. 2012. Ultra-structural identification of interstitial cells of Cajal in the zebrafish Danio rerio. *Cell Tissue Res*. 349:483-491.

Bretscher, A., and K. Weber. 1979. Villin: the major microfilament-associated protein of the intestinal microvillus. *Proc. Natl. Acad. Sci. U. S. A*. 76:2321-2325.

Bustin, S.A. 2004. A-Z of quantitative PCR. International University Line, La Jolla, CA.

Clift, D., H. Richendrfer, R.J. Thorn, R.M. Colwill, and R. Creton. 2014. High-throughput analysis of behavior in zebrafish larvae: effects of feeding. *Zebrafish*. 11:455-461.

Cummings, D.E., and J. Overduin. 2007. Gastrointestinal regulation of food intake. *J. Clin. Invest*. 117:13-23.

Eames Nalle, S.C., K.F. Franse, and M.D. Kinkel. 2017. Analysis of pancreatic disease in zebrafish. *Methods Cell Biol*. 138:271-295.

Elli, L., C. Tomba, F. Branchi, L. Roncoroni, V. Lombardo, M.T. Bardella, F. Ferretti, D. Conte, F. Valiante, L. Fini, E. Forti, R. Cannizzaro, S. Maiero, C. Londoni, A. Lauri, G. Fornaciari, N. Lenoci, R. Spagnuolo, G. Basilisco, F. Somalvico, B. Borgatta, G. Leandro, S. Segato, D. Barisani, G. Morreale, and E. Buscarini. 2016. Evidence for the presence of non-celiac gluten sensitivity in patients with functional gastrointestinal symptoms: Results from a multicenter randomized double-blind placebo-controlled gluten challenge. *Nutrients*. 8:84.

Enck, P., Q. Aziz, G. Barbara, A.D. Farmer, S. Fukudo, E.A. Mayer, B. Niesler, E.M. Quigley, M. Rajilić-Stojanović, M. Schemann, J. Schwille-Kiuntke, M. Simren, S. Zipfel, and R.C. Spiller. 2016. Irritable bowel syndrome. *Nat. Rev. Dis. Primers*. 2:16014.

Eom, J., A. Hong, Y.H. Kang, H.J. Yoo, E.J. Chang, S.W. Kang, S.Y. Yoon, S.Y. Kim, and Y. Song. 2014. Molecular cloning, regulation, and functional analysis of two GHS-R genes in zebrafish. *Exp. Cell Res*. 326:10-21.

Field, H.A., K.A. Kelley, L. Martell, A.M. Goldstein, and F.C. Serluca. 2009. Analysis of gastrointestinal physiology using a novel intestinal transit assay in zebrafish. *Neurogastroenterol. Motil*. 21:304-312.

Ford, A.C., A.D. Sperber, M. Corsetti, and M. Camilleri. 2020. Irritable bowel syndrome. *Lancet*. 396:1675-1688.

Haikal, Z., Q.H. Setiadi, S. Sunanto, and I. Hasanah. 2022. Modified Swenson-like pull-through with temporary stump for redo pull-through after failed transanal endorectal pull-through. *J. Taibah Univ. Med. Sci*. 17:150-154.

Halmos, E.P., and P.R. Gibson. 2019. Controversies and reality of the FODMAP diet for patients with irritable bowel syndrome. *J. Gastroenterol. Hepatol*. 34:1134-1142.

Holtmann, G.J., A.C. Ford, and N.J. Talley. 2016. Pathophysiology of irritable bowel syndrome. *Lancet Gastroenterol. Hepatol*. 1:133-146.

Howe, K., M.D. Clark, C.F. Torroja, J. Torrance, C. Berthelot, M. Muffato, J.E. Collins, S. Humphray, K. McLaren, L. Matthews, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott, J.C. Barrett, R. Koch, G.J. Rauch, S. White, W. Chow, B. Kilian, L.T. Quintais, J.A. Guerra-Assunção, Y. Zhou, Y. Gu, J. Yen, J.H. Vogel, T. Eyre, S. Redmond, R. Banerjee, J. Chi, B. Fu, E. Langley, S.F. Maguire, G.K. Laird, D. Lloyd, E. Kenyon, S. Donaldson, H. Sehra, J. Almeida-King, J. Loveland, S. Trevanion, M. Jones, M. Quail, D. Willey, A. Hunt, J. Burton, S. Sims, K. McLay, B. Plumb, J. Davis, C. Clee, K. Oliver, R. Clark, C. Riddle, D. Elliot, G. Threadgold, G. Harden, D. Ware, S. Begum, B. Mortimore, G. Kerry, P. Heath, B. Phillimore, A. Tracey, N. Corby, M. Dunn, C. Johnson, J. Wood, S. Clark, S. Pelan, G. Griffiths, M. Smith, R. Glithero, P. Howden, N. Barker, C. Lloyd, C. Stevens, J. Harley, K. Holt, G. Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Gordon, K. Auger, D. Wright, J. Collins, C. Raisen, L. Dyer, K. Leung, L. Robertson, K. Ambridge, D. Leongamornlert, S. McGuire, R. Gilderthorp, C. Griffiths, D. Manthravadi, S. Nichol, G. Barker, et al. 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*. 496:498-503.

Keller, J., G. Bassotti, J. Clarke, P. Dinning, M. Fox, M. Grover, P.M. Hellström, M. Ke, P. Layer, C. Malagelada, H.P. Parkman, S.M. Scott, J. Tack, M. Simren, H. Törnblom, and M. Camilleri. 2018. Expert consensus document: Advances in the diagnosis and classification of gastric and intestinal motility disorders. *Nat. Rev. Gastroenterol. Hepatol*. 15:291-308.

Lance, S.L., A.S. Peterson, and M. Hagedorn. 2004. Developmental expression of aquaporin-3 in zebrafish embryos (Danio rerio). *Comp. Biochem. Physiol. C Toxicol. Pharmacol*. 138:251-258.

Lickwar, C.R., J.G. Camp, M. Weiser, J.L. Cocchiaro, D.M. Kingsley, T.S. Furey, S.Z. Sheikh, and J.F. Rawls. 2017. Genomic dissection of conserved transcriptional regulation in intestinal epithelial cells. *PLoS Biol*. 15:e2002054.

Lin, C.W., S.T. Yen, H.T. Chang, S.J. Chen, S.L. Lai, Y.C. Liu, T.H. Chan, W.L. Liao, and S.J. Lee. 2010. Loss of cofilin 1 disturbs actin dynamics, adhesion between enveloping and deep cell layers and cell movements during gastrulation in zebrafish. *PLoS One*. 5:e15331.

Luo, J., L. Han, L. Liu, L. Gao, B. Xue, Y. Wang, S. Ou, M. Miller, and X. Peng. 2018. Catechin supplemented in a FOS diet induces weight loss by altering cecal microbiota and gene expression of colonic epithelial cells. *Food Funct*. 9:2962-2969.

Martinez-Lopez, E., F. Curiel-Lopez, A. Hernandez-Nazara, L.E. Moreno-Luna, M.E. Ramos-Marquez, S. Roman, and A. Panduro. 2015. Influence of ApoE and FABP2 polymorphisms and environmental factors in the susceptibility to gallstone disease. *Ann. Hepatol*. 14:515-523.

Murphy, K.G., and S.R. Bloom. 2006. Gut hormones and the regulation of energy homeostasis. *Nature*. 444:854-859.

Otis, J.P., E.M. Zeituni, J.H. Thierer, J.L. Anderson, A.C. Brown, E.D. Boehm, D.M. Cerchione, A.M. Ceasrine, I. Avraham-Davidi, H. Tempelhof, K. Yaniv, and S.A. Farber. 2015. Zebrafish as a model for apolipoprotein biology: comprehensive expression analysis and a role for ApoA-IV in regulating food intake. *Dis. Model. Mech*. 8:295-309.

Qin, H.Y., C.W. Cheng, X.D. Tang, and Z.X. Bian. 2014. Impact of psychological stress on irritable bowel syndrome. *World J. Gastroenterol*. 20:14126-14131.

Radovanovic-Dinic, B., S. Tesic-Rajkovic, S. Grgov, G. Petrovic, and V. Zivkovic. 2018. Irritable bowel syndrome - from etiopathogenesis to therapy. *Biomed. Pap. Med. Fac. Univ. Palacky. Olomouc. Czech Repub*. 162:1-9.

Raskov, H., J. Burcharth, H.C. Pommergaard, and J. Rosenberg. 2016. Irritable bowel syndrome, the microbiota and the gut-brain axis. *Gut Microbes*. 7:365-383.

Rodiño-Janeiro, B.K., M. Vicario, C. Alonso-Cotoner, R. Pascua-García, and J. Santos. 2018. A review of microbiota and irritable bowel syndrome: Future in therapies. *Adv. Ther*. 35:289-310.

Saha, L. 2014. Irritable bowel syndrome: pathogenesis, diagnosis, treatment, and evidence-based medicine. *World J. Gastroenterol*. 20:6759-6773.

Silberstein, C., A. Kierbel, G. Amodeo, E. Zotta, F. Bigi, D. Berkowski, and C. Ibarra. 1999. Functional characterization and localization of AQP3 in the human colon. *Braz. J. Med. Biol. Res*. 32:1303-1313.

Smathers, R.L., and D.R. Petersen. 2011. The human fatty acid-binding protein family: evolutionary divergences and functions. *Hum. Genomics*. 5:170-191.

Song, G., D.W. Bailey, K.A. Dunlap, R.C. Burghardt, T.E. Spencer, F.W. Bazer, and G.A. Johnson. 2010. Cathepsin B, cathepsin L, and cystatin C in the porcine uterus and placenta: potential roles in endometrial/placental remodeling and in fluid-phase transport of proteins secreted by uterine epithelia across placental areolae. *Biol. Reprod*. 82:854-864.

Szeitz, A., and S.M. Bandiera. 2018. Analysis and measurement of serotonin. *Biomed. Chromatogr*. 32:10.1002/bmc.4135.

Vazquez-Roque, M.I., M. Camilleri, T. Smyrk, J.A. Murray, E. Marietta, J. O'Neill, P. Carlson, J. Lamsam, D. Janzow, D. Eckert, D. Burton, and A.R. Zinsmeister. 2013. A controlled trial of gluten-free diet in patients with irritable bowel syndrome-diarrhea: effects on bowel frequency and intestinal function. *Gastroenterology*. 144:903-911.e3.

Venkatachalam, A.B., D.L. Sawler, and J.M. Wright. 2013. Tissue-specific transcriptional modulation of fatty acid-binding protein genes, fabp2, fabp3 and fabp6, by fatty acids and the peroxisome proliferator, clofibrate, in zebrafish (Danio rerio). *Gene*. 520:14-21.

Wall, G.C., G.A. Bryant, M.M. Bottenberg, E.D. Maki, and A.R. Miesner. 2014. Irritable bowel syndrome: a concise review of current treatment concepts. *World J. Gastroenterol*. 20:8796-8806.

Wallace, K.N., S. Akhter, E.M. Smith, K. Lorent, and M. Pack. 2005. Intestinal growth and differentiation in zebrafish. *Mech. Dev*. 122:157-173.

Wang, Z., J. Du, S.H. Lam, S. Mathavan, P. Matsudaira, and Z. Gong. 2010. Morphological and molecular evidence for functional organization along the rostrocaudal axis of the adult zebrafish intestine. *BMC Genomics*. 11:392.

Westerfield, M. 2000. The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Danio rerio). Univ. of Oregon Press, Eugene.

Ye, L., M. Bae, C.D. Cassilly, S.V. Jabba, D.W. Thorpe, A.M. Martin, H.Y. Lu, J. Wang, J.D. Thompson, C.R. Lickwar, K.D. Poss, D.J. Keating, S.E. Jordt, J. Clardy, R.A. Liddle, and J.F. Rawls. 2021. Enteroendocrine cells sense bacterial tryptophan catabolites to activate enteric and vagal neuronal pathways. *Cell Host Microbe*. 29:179-196.e179.

Zhao, S., J. Xia, X. Wu, L. Zhang, P. Wang, H. Wang, H. Li, X. Wang, Y. Chen, J. Agnetti, Y. Li, D. Pei, and X. Shu. 2018. Deficiency in class III PI3-kinase confers postnatal lethality with IBD-like features in zebrafish. *Nat. Commun*. 9:2639.

Appendix A

Multiple sequence alignment, *apoa1*. The first line in each bundle of sequences is for humans (*Homo sapiens, HS*). The second line is for Mice (*Mus musculus, MM*). The third line is for Zebrafish (*Danio rerio, DR*). The asterisks under each bundle of sequences signify DNA bases are identical. Program: Clustal Omega.

Apolipoprotein 1 (apoa1)

```
NM_213468.1         --AGCA--CACGTCTCTGAGGAACATACAGCTTGTTTCATCCAGACAAAGGAAAACATCT         56
NM_001318144.2      ----------GCTCGCTGCGCCACCGCCTCCCGCCACCCCTGCCCGCCCGA--CAGCGCC         48
NM_016689.2         ACCGCTCTCGGTGCCTTGCGCTAGCTACTTT-GCACTCGTACGCCGCCGGACCTCGCCGC         59
                                 *  ** *  *    *         *           *

NM_213468.1         GGACAAGTCTGAAAGATGGGTTGGCAGAAAAGCGTTCTGGATAAGCTTGCGCAGACTTTC         116
NM_001318144.2      GCCGCCTGCCCCGCCATGGGTCGACAGAAGGAGCTGGTGTCCCGCTGCGGGGAGATGCTC         108
NM_016689.2         TGCCTGCCTCGCGCCATGGGTCGACAGAAGGAGTTGATGAATCGTTGTGGGGAGATGCTT         119
                                  ****** * *****     *  **         * * ***   *

NM_213468.1         CAAATCCGCAACAAGTTACTGCGCCAGGGATTGGCTGAATGCTTAGGAACTCTCATCCTG         176
NM_001318144.2      CACATCCGCTACCGGCTGCTCCGACAGGCGCTGGCCGAGTGCCTGGGGACCCTCATCCTG         168
NM_016689.2         CACATCCGCTACCGGCTGCTTCGCCAGGCGCTGGCGGAGTGCCTGGGGACCCTCATCCTT         179
                    ** ****** **  * * ** ** ****   **** ** *** * ** ** ********

NM_213468.1         GTGATGTTTGGCTGTGGTTCATTGGCCCAGTTGAAACTAAGCGAAGGTTCTCATGGTCTC         236
NM_001318144.2      GTGATGTTTGGCTGTGGCTCCGTGGCCCAGGTTGTGCTCAGCCGGGGCACCCACGGTGGT         228
NM_016689.2         GTGATGTTTGGCTGTGGCTCCGTGGCTCAGGTGGTGCTCAGCCGTGGCACCCATGGTGGC         239
                    ***************** **  **** *** *    ** ***   **  * ** ***

NM_213468.1         TTTCTCACTGCAAACCTTGCTTTTGGGTTTGGTGCTACTCTTGGAATCTTGGTTTGCGGC         296
NM_001318144.2      TTCCTCACCATCAACCTGGCCTTTGGCTTTGCTGTCACTCTGGGCATCCTCATCGCTGGC         288
NM_016689.2         TTCCTCACCATCAACTTGGCTTTTGGCTTCGCTGTCACCCTTGGCATCTTGGTGGCTGGC         299
                    ** *****    *** * ** ***** ** * **  ** ** ** *** *  *    ***

NM_213468.1         CAGGTGTCAGGCGGACATTTAAATCCTGCTGTTACATTTGCTCTCTGCCTCTTGGGAAGA         356
NM_001318144.2      CAGGTCTCTGGGGCCCACCTGAACCCTGCCGTGACCTTTGCCATGTGCTTCCTGGCTCGT         348
NM_016689.2         CAGGTGTCTGGTGCCCACTTGAACCCCGCTGTGACCTTCGCAATGTGCTTCCTGGCACGA         359
                    ***** ** ** *  **  * ** ** ** ** ** ** **  * *** ** ***   *

NM_213468.1         GAAAAATGGAGAAAGTTTCCTGTGTACTTTCTGTTCCAAACACTCGGATCCTTCTTGGGT         416
NM_001318144.2      GAGCCCTGGATCAAGCTGCCCATCTACACCCTGGCACAGACGCTGGGAGCCTTCTTGGGT         408
NM_016689.2         GAGCCCTGGATCAAGCTGCCCATCTATGCACTGGCACAGACACTGGGGGCCTTCTTGGGC         419
                    **    ****  *** * **  * **    ***   ** ** ** **  **********

NM_213468.1         GCTGCTATTATCTTTGCCGAATACCATGACGCAATTTATGATTATGCTGGAGAATCAAAT         476
NM_001318144.2      GCTGGAATAGTTTTTGGGCTGTATTATGATGCAATCTGGCACTTCGCCGACAA------C         462
NM_016689.2         GCTGGGATTGTTTTTGGGCTGTACTACGATGCAATCTGGGCCTTTGCCAACAA------T         473
                    ****  **  * ****     **  * ** ***** *     *  **     *

NM_213468.1         GAATTGCTTGTACTGGGTGAAAAAGAAACAGCTGGGATTTTTGCTACATACCCAAGCAAA         536
NM_001318144.2      CAGCTTTTTGTTTCGGGCCCCAATGGCACAGCCGGCATCTTTGCTACCTACCCCTCTGGA         522
NM_016689.2         GAGCTTTTCGTCTCTGGCCCCAACGGCACAGCTGGAATCTTTGCCACCTATCCCTCTGGA         533
                     *  *  * **    **    ** *  ***** ** ** ***** ** ** **      *
```

```
NM_213468.1         TATCTCACCCCCCTAAATGGATTTTTTGACCAGGTGATAGGCACAGCATCCCTGATTGTG        596
NM_001318144.2      CACTTGGATATGATCAATGGCTTCTTTGACCAGGACCGGCCAGCATTGGTGGTGGGTGCC        582
NM_016689.2         CACTTGGACATGGTCAATGGCTTCTTTGATCAGTTCATAGGCACAGCCGCCCTTATTGTG        593
                     *  *        * ***** ** ***** ***          **       *   **

NM_213468.1         TGCATCCTGGCCATTGTGGACCCCTACAACAACCCGATCCCTCAAGGTCTTGAGGCCTTC        656
NM_001318144.2      CATCGTGTCCCCACTCCTGGGCTCCATTG----------------------CGGGTGTC        619
NM_016689.2         TGTGTACTGGCCATCGTTGACCCTTATAACAACCCTGTGCCCCGTGGCCTGGAGGCTTTC        653
                           *  ***     *  *   *                              **   **

NM_213468.1         ACAGTGGGATTCAGCGTCCTTATCATTGGTCTCTCCATGGGCTTCAATTCTGGCTATGCA        716
NM_001318144.2      TTCGTGTACCAGCTGATG------ATCGGCTGCCACCTGGAGCA-------GCCCCCACC        666
NM_016689.2         ACTGTGGGCCTGGTGGTCCTGGTCATTGGAACCTCCATGGGCTTCAATTCTGGCTATGCC        713
                       ***          *       ** **   *  * ***          * *    *

NM_213468.1         GTAAACCCAGCTCGAGATTTTGGACCTCGTCTTTTTACTGCCATGGCTGGTT-------G        769
NM_001318144.2      CTCCAACGAGGAAGAGAATGTGAAGCTGGCCCATGTGAAGCACAAGGAGCAGATCTGAGT        726
NM_016689.2         GTCAACCCTGCCCGTGACTTTGGACCTCGCCTCTTCACCGCCCTGG--C-----TGGCTG        766
                     *  * *  *   * ** * ** * ** * *  *     **    *

NM_213468.1         GGGTAGTGAAGTCTTCACAGCCAGGGATTATTGGTTTTTGGTGCCCATCTTTGCTCCGTT        829
NM_001318144.2      GGGCAGGGGCCATCTCCCCACTCCGCTGCCCTGGCCTTGAGCATCCACTGACTGTCCAAG        786
NM_016689.2         GGGCTCAGAAGTCTTCACGACTGGCCGGCACTGGTGGTGGGTACCCATTGTCTCCCCACT        826
                    ***    *      ** *  *          ***   *  *   ***        **

NM_213468.1         CATTGGAGCCGTTATTGGTGTGATTGTGTACCAGCTGATGGTGGGATGGCATGTGGAAGG        889
NM_001318144.2      GGCCACTCC--------------------------------CAAGAAGCCCCCTTCACGA        814
NM_016689.2         CCTGGGTTCCATCGCTGGTGTCTTCGTGTACCAGCTCATGATTGGTTGCCACCTGGAGCA        886
                            *                                      *  * *   *  *

NM_213468.1         AGAGGCACGAGATAAGAAAGCTAAAGCTAGAGAGGAGGTGATGAACCTCAATGACGTCG-        948
NM_001318144.2      TCCACCCTTTCAGGCTAAGGAGCTCCCT------------ATCTACCCTCACCCCACGAG        862
NM_016689.2         GCCCCCACCCTCCACCGAGGAAGAGAAT------------GTGAAGCTGGCCCACATGAA        934
                         *          * *       *            *  * *       *

NM_213468.1         CCAGCAAGGAATGATTTTCAGGATCTCCATCAGCCCTCAAC-ACCACAAATCCACAT---        1004
NM_001318144.2      ACAGCCCCTTCAGGATTT--------CCACTGGACCTTGCCCAAATAGCACCTT------        908
NM_016689.2         ACACAAGGAGCAGATCTGAGTGGGCAGCAGCCCCCCTCCCCCACTGTGCACTCTCCTGAG        994
                     **        *   *          **     ***   * *     *

NM_213468.1         AAGCTTCTGCTTAAGGATCTCTGGGTTCAATAAAAACCACTTGCACCACACAAAACTCCC        1064
NM_001318144.2      AGGCCACTGC-----------------CCCTAAGCTGG-GGT--GGAACCGGAATTTGGG        948
NM_016689.2         TGTCCACTGACTGTGTGGGGACCAGTCCCCGAAAGCCC-TTTGTGATGCCTCTCTCGGGC        1053
                       *  ***                  *   **        *     *

NM_213468.1         TCATCTGTTTCAGAGGTGATTCTCACAGTTTTGCAGATCTTCCACTGGACTTTGAATAAA        1124
NM_001318144.2      TC----------------AATACATCCTTTTG-----TCTCCCAAGGGAAGAGAATGGGC        987
NM_016689.2         TA----------------AACCGCTCCCTGTG-----TCCACCCCTG--CTGGATGGGCC        1090
                    *                *     *  * *     **  **   *
```

```
NM_213468.1         AAGCGTTATACATTTCTATACTTTACAAATTCTGGGTTGT---TTTAACCAAATTTGGGT        1181
NM_001318144.2      AGCAGGTAT--GTGTGTGTG---TGCATGTGTGTGCATG--------------TGTGTG        1027
NM_016689.2         CTCCAGAAT--TTCTATGAACTCTGCCCATTAGGGCATTAGGTTCCCACCCACCTTTAAG        1148
                           **   * * *     * *   *    *  *               *

NM_213468.1         ACAATTTTGACTAACCCAATGGTTGGCTTAAAA--------------------------        1214
NM_001318144.2      CA----------------------------------------------------------        1029
NM_016689.2         CCAAGGTAGGATAGCAAATAAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG        1208

NM_213468.1         ATTTTATTTAAAAATTTAACCAACTGCTGGGTTATTGCTTTCCAACCCAGCATTATTTTA        1274
NM_001318144.2      TGTGTGTGCAGGGGTGTGTGTGTGTGGGGGGGG--TTCCCAGATATTCAGGGCAAGGGAC        1087
NM_016689.2         AGAGAGAGAATGAATGTGTACATGTGTGCTGTT--TTCTAAGCTGAATGATGCAAAGGCA        1266
                          *    * *       **    *    * *                *

NM_213468.1         GATTTCCAACTTTATATCAGACATTTGATTGTTTTTCAAGCATGACTCTTATTTCTTTCT        1334
NM_001318144.2      CAGTCGGAAGGGATTCTGG-CTATTGGG-------GGAGCCCAGAGACAGGGGAAGGCAG        1139
NM_016689.2         AGGGACCAAGTTTTCAAAACAAACTGTA-------GCAG------CTCAGGGGAAGGGAG        1313
                           **            * *           *        *

NM_213468.1         TCCACAT-----GCTTTATTCTTCCATTAAGGGAACAAGGTGTTACTTCAACAGGACGAA        1389
NM_001318144.2      CC--TGTCCATCTGTGCATAAGGAGAGGAAAGTTCCAGGGTGTGTATGTTTCAGG-----        1192
NM_016689.2         CCCAGGGGAAGGGAGAAAGTGAGTCAGGAATGTGCCAGAGTGTGCATGCTTCAGG-----        1368
                     *              *       *  ** *   **  ****   *    ****

NM_213468.1         TTATATTAATTCACTCTTTTGAAAAGTTAGTAAAAGTTTAGATCTTTGAAATGAAACTGT        1449
NM_001318144.2      ------GGCTTCACATGGAGGAGC---T-----------------GCAGATAGATAT-G        1224
NM_016689.2         ------GACTCCTCCATGTGGAGG---T-----------------GGACCCAGAAGT-G        1400
                           * * *      **     *                    *     *  *

NM_213468.1         TGTTGTTACACCTATCAATATATACAGTATTTCATGCTTTTGTGCTACTA---ATATCTA        1506
NM_001318144.2      TGTTTCTGTGTATGTGTATGTCTGCCTTTTTTTCTAAGTGGGGGCTTCTACAGGCTTTTG        1284
NM_016689.2         AGTTTCTAAGTATGCGTGTGCCTACTGTTTTTTTTTTTTTTTTTGAAATG--GACTTCTA        1458
                     ***  *     *     *   * *  * ***  *   *         *      * *

NM_213468.1         A-------ATATGTTGCCATTCACGCTGAGAAAGTTGAGAAGTTACCCAAAAAAGCTACA        1559
NM_001318144.2      GGAAGTAGGGTGGATGTGGGTAGGGCTGGGAGGAGGGGGCCACAGCTTAGGTTTGGAGCT        1344
NM_016689.2         G----------GCTTGGGGAGGGGGAAGGGATAAGAAGGGTGTAGCTCACATCTGGAGCT        1508
                                **        *  * **       *      *  *     *   *

NM_213468.1         GTACTTTTTAAAAAAATTGTAAATATTTTGTATTGTCATTTTCATTTAAAGCATTTTTTT        1619
NM_001318144.2      CTGGATGTACATACATAAGTAGGAGCAGTGG--------------------GACGTGTTT        1384
NM_016689.2         A-------TGACCCTTGACTGGGGGCTGTGT--------------------AATATGTTT        1541
                             *        *        **                      *  * ***

NM_213468.1         TTGTTGTATTACATGTATGGATTACAACTTAACATGTACACTTACGTCTGATAGTGG---        1676
NM_001318144.2      CTGTCATAATGCAGGCATGAAGGGTGGAGTGAAGTCAGGTCATAAGTTTCATGTTTGCTT        1444
NM_016689.2         CTGTTATAAGATAGACATTGGGAGGGGCTGAAGTCCAGGTCGTAAGTTTCATAATTTGTT        1601
                     ***  **    *   **             *        * ** ** * **  *
```

```
NM_213468.1         --AGAGATCCTGACATTTATTTGACATTGCAGTTTTTAGACTGTTAGATCTGTTCATTTC        1734
NM_001318144.2      TTGTTTTGTTTTGTTTTTAATGTATGTAGCAGATGTTACAGTCTTAGGGATCCGGGA--T        1502
NM_016689.2         TTTTAAATAT--ATAAATATATACATACATATATGTTACAGCCCTAGGAATAGGGGT--G        1657
                              **          *  * *** *    ***   *

NM_213468.1         ATAACATGTTAATTTTGAAAGTGAA----TGTTTTAAATCAACTGAT---TGTTATCCTA        1787
NM_001318144.2      GGGAGACCCCACTTTAGAAAGGGTCGTCACTCCTTTAATCCTCTACTCAACAATGTACTC        1562
NM_016689.2         GGAAACTCCACTTTTTAAAAGGGGTTTCCTTTCTTTAATCCTCCAATCAACAATGTACTG        1717
                       *        ***  **** *          ** ****  *   *      * * **

NM_213468.1         ATAAATTTTTTTAAAGAAGAAAAAAAAAAAAAAAAAA---------      1824
NM_001318144.2      TTTTACTTTTATATTAAAAAAAAT-AAAATAAATATGTGCCTAAAA      1607
NM_016689.2         TTGCCTTTTATATATAAAAAAGAATAAAACGTATACATGCTACAGG      1763
                     *    ***       ** ** *  ****   * *
```

Appendix B

Multiple sequence alignment, *aqp3*. The first line in each bundle of sequences is for humans (*Homo sapiens, HS*). The second line is for Mice (*Mus musculus, MM*). The third line is for Zebrafish (*Danio rerio, DR*). The asterisks under each bundle of sequences signify DNA bases are identical. Program: Clustal Omega.

Aquaporin 3 (aqp3)

```
NM_213468.1         --AGCA--CACGTCTCTGAGGAACATACAGCTTGTTTCATCCAGACAAAGGAAAACATCT        56
NM_001318144.2      ----------GCTCGCTGCGCCACCGCCTCCCGCCACCCCTGCCCGCCCGA--CAGCGCC        48
NM_016689.2         ACCGCTCTCGGTGCCTTGCGCTAGCTACTTT-GCACTCGTACGCCGCCGGACCTCGCCGC        59
                                 *  ** *  *    *         *           *

NM_213468.1         GGACAAGTCTGAAAGATGGGTTGGCAGAAAAGCGTTCTGGATAAGCTTGCGCAGACTTTC        116
NM_001318144.2      GCCGCCTGCCCCGCCATGGGTCGACAGAAGGAGCTGGTGTCCCGCTGCGGGGAGATGCTC        108
NM_016689.2         TGCCTGCCTCGCGCCATGGGTCGACAGAAGGAGTTGATGAATCGTTGTGGGGAGATGCTT        119
                                  ****** * *****     *  **         * * ***   *

NM_213468.1         CAAATCCGCAACAAGTTACTGCGCCAGGGATTGGCTGAATGCTTAGGAACTCTCATCCTG        176
NM_001318144.2      CACATCCGCTACCGGCTGCTCCGACAGGCGCTGGCCGAGTGCCTGGGGACCCTCATCCTG        168
NM_016689.2         CACATCCGCTACCGGCTGCTTCGCCAGGCGCTGGCGGAGTGCCTGGGGACCCTCATCCTT        179
                    ** ****** **  * * ** ** ****   **** ** *** * ** ** ********

NM_213468.1         GTGATGTTTGGCTGTGGTTCATTGGCCCAGTTGAAACTAAGCGAAGGTTCTCATGGTCTC        236
NM_001318144.2      GTGATGTTTGGCTGTGGCTCCGTGGCCCAGGTTGTGCTCAGCCGGGGCACCCACGGTGGT        228
NM_016689.2         GTGATGTTTGGCTGTGGCTCCGTGGCTCAGGTGGTGCTCAGCCGTGGCACCCATGGTGGC        239
                    ***************** **  **** *** *    ** ***   **  * ** ***

NM_213468.1         TTTCTCACTGCAAACCTTGCTTTTGGGTTTGGTGCTACTCTTGGAATCTTGGTTTGCGGC        296
NM_001318144.2      TTCCTCACCATCAACCTGGCCTTTGGCTTTGCTGTCACTCTGGGCATCCTCATCGCTGGC        288
NM_016689.2         TTCCTCACCATCAACTTGGCTTTTGGCTTCGCTGTCACCCTTGGCATCTTGGTGGCTGGC        299
                    ** *****    *** * ** ***** ** * **  ** ** ** *** *  *    ***

NM_213468.1         CAGGTGTCAGGCGGACATTTAAATCCTGCTGTTACATTTGCTCTCTGCCTCTTGGGAAGA        356
NM_001318144.2      CAGGTCTCTGGGGCCCACCTGAACCCTGCCGTGACCTTTGCCATGTGCTTCCTGGCTCGT        348
NM_016689.2         CAGGTGTCTGGTGCCCACTTGAACCCCGCTGTGACCTTCGCAATGTGCTTCCTGGCACGA        359
                    ***** ** ** *  **  * ** ** ** ** ** ** **  * *** ** ***   *

NM_213468.1         GAAAAATGGAGAAAGTTTCCTGTGTACTTTCTGTTCCAAACACTCGGATCCTTCTTGGGT        416
NM_001318144.2      GAGCCCTGGATCAAGCTGCCCATCTACACCCTGGCACAGACGCTGGGAGCCTTCTTGGGT        408
NM_016689.2         GAGCCCTGGATCAAGCTGCCCATCTATGCACTGGCACAGACACTGGGGGCCTTCTTGGGC        419
                    **    ****  *** * **  * **    ***   ** ** ** **  **********

NM_213468.1         GCTGCTATTATCTTTGCCGAATACCATGACGCAATTTATGATTATGCTGGAGAATCAAAT        476
NM_001318144.2      GCTGGAATAGTTTTTGGGCTGTATTATGATGCAATCTGGCACTTCGCCGACAA------C        462
NM_016689.2         GCTGGGATTGTTTTTGGGCTGTACTACGATGCAATCTGGGCCTTTGCCAACAA------T        473
                    ****  **  * ****     **  * ** ***** *     *  **     *

NM_213468.1         GAATTGCTTGTACTGGGTGAAAAAGAAACAGCTGGGATTTTTGCTACATACCCAAGCAAA        536
NM_001318144.2      CAGCTTTTTGTTTCGGGCCCCAATGGCACAGCCGGCATCTTTGCTACCTACCCCTCTGGA        522
NM_016689.2         GAGCTTTTCGTCTCTGGCCCCAACGGCACAGCTGGAATCTTTGCCACCTATCCCTCTGGA        533
                     *  *  * **    **    ** *  ***** ** ** ***** ** ** **      *
```

```
NM_213468.1         TATCTCACCCCCCTAAATGGATTTTTTGACCAGGTGATAGGCACAGCATCCCTGATTGTG        596
NM_001318144.2      CACTTGGATATGATCAATGGCTTCTTTGACCAGGACCGGCCAGCATTGGTGGTGGGTGCC        582
NM_016689.2         CACTTGGACATGGTCAATGGCTTCTTTGATCAGTTCATAGGCACAGCCGCCCTTATTGTG        593
                     *  *        * ***** ** ***** ***          **       *   **

NM_213468.1         TGCATCCTGGCCATTGTGGACCCCTACAACAACCCGATCCCTCAAGGTCTTGAGGCCTTC        656
NM_001318144.2      CATCGTGTCCCCACTCCTGGGCTCCATTG----------------------CGGGTGTC        619
NM_016689.2         TGTGTACTGGCCATCGTTGACCCTTATAACAACCCTGTGCCCCGTGGCCTGGAGGCTTTC        653
                           *  ***     *  *   *                              **   **

NM_213468.1         ACAGTGGGATTCAGCGTCCTTATCATTGGTCTCTCCATGGGCTTCAATTCTGGCTATGCA        716
NM_001318144.2      TTCGTGTACCAGCTGATG------ATCGGCTGCCACCTGGAGCA-------GCCCCCACC        666
NM_016689.2         ACTGTGGGCCTGGTGGTCCTGGTCATTGGAACCTCCATGGGCTTCAATTCTGGCTATGCC        713
                       ***          *       ** **   *  * ***          * *    *

NM_213468.1         GTAAACCCAGCTCGAGATTTTGGACCTCGTCTTTTTACTGCCATGGCTGGTT-------G        769
NM_001318144.2      CTCCAACGAGGAAGAGAATGTGAAGCTGGCCCATGTGAAGCACAAGGAGCAGATCTGAGT        726
NM_016689.2         GTCAACCCTGCCCGTGACTTTGGACCTCGCCTCTTCACCGCCCTGG--C-----TGGCTG        766
                     *  * *  *   * ** * ** * ** * *  *     **    *

NM_213468.1         GGGTAGTGAAGTCTTCACAGCCAGGGATTATTGGTTTTTGGTGCCCATCTTTGCTCCGTT        829
NM_001318144.2      GGGCAGGGGCCATCTCCCCACTCCGCTGCCCTGGCCTTGAGCATCCACTGACTGTCCAAG        786
NM_016689.2         GGGCTCAGAAGTCTTCACGACTGGCCGGCACTGGTGGTGGGTACCCATTGTCTCCCCACT        826
                    ***    *      ** *  *          ***   *  *   ***        **

NM_213468.1         CATTGGAGCCGTTATTGGTGTGATTGTGTACCAGCTGATGGTGGGATGGCATGTGGAAGG        889
NM_001318144.2      GGCCACTCC--------------------------------CAAGAAGCCCCCTTCACGA        814
NM_016689.2         CCTGGGTTCCATCGCTGGTGTCTTCGTGTACCAGCTCATGATTGGTTGCCACCTGGAGCA        886
                           *                                    *  * *   *  *

NM_213468.1         AGAGGCACGAGATAAGAAAGCTAAAGCTAGAGAGGAGGTGATGAACCTCAATGACGTCG-        948
NM_001318144.2      TCCACCCTTTCAGGCTAAGGAGCTCCCT------------ATCTACCCTCACCCCACGAG        862
NM_016689.2         GCCCCCACCCTCCACCGAGGAAGAGAAT------------GTGAAGCTGGCCCACATGAA        934
                         *           * *       *             *  * *       *

NM_213468.1         CCAGCAAGGAATGATTTTCAGGATCTCCATCAGCCCTCAAC-ACCACAAATCCACAT---        1004
NM_001318144.2      ACAGCCCCTTCAGGATTT--------CCACTGGACCTTGCCCAAATAGCACCTT------        908
NM_016689.2         ACACAAGGAGCAGATCTGAGTGGGCAGCAGCCCCCCTCCCCCACTGTGCACTCTCCTGAG        994
                     **         *   *          **     ***   * *      *

NM_213468.1         AAGCTTCTGCTTAAGGATCTCTGGGTTCAATAAAAACCACTTGCACCACACAAAACTCCC        1064
NM_001318144.2      AGGCCACTGC-----------------CCCTAAGCTGG-GGT--GGAACCGGAATTTGGG        948
NM_016689.2         TGTCCACTGACTGTGTGGGGACCAGTCCCCGAAAGCCC-TTTGTGATGCCTCTCTCGGGC        1053
                       *  ***                  *   **        *      *
```

```
NM_213468.1         TCATCTGTTTCAGAGGTGATTCTCACAGTTTTGCAGATCTTCCACTGGACTTTGAATAAA        1124
NM_001318144.2      TC----------------AATACATCCTTTTG-----TCTCCCAAGGGAAGAGAATGGGC        987
NM_016689.2         TA----------------AACCGCTCCCTGTG-----TCCACCCCTG--CTGGATGGGCC        1090
                    *                 *      *  * *      **  **   *

NM_213468.1         AAGCGTTATACATTTCTATACTTTACAAATTCTGGGTTGT---TTTAACCAAATTTGGGT        1181
NM_001318144.2      AGCAGGTAT--GTGTGTGTG---TGCATGTGTGTGCATG--------------TGTGTG        1027
NM_016689.2         CTCCAGAAT--TTCTATGAACTCTGCCCATTAGGGCATTAGGTTCCCACCCACCTTTAAG        1148
                           **   * * *      * *   *    *  *               *

NM_213468.1         ACAATTTTGACTAACCCAATGGTTGGCTTAAAA--------------------------        1214
NM_001318144.2      CA----------------------------------------------------------        1029
NM_016689.2         CCAAGGTAGGATAGCAAATAAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG        1208


NM_213468.1         ATTTTATTTAAAAATTTAACCAACTGCTGGGTTATTGCTTTCCAACCCAGCATTATTTTA        1274
NM_001318144.2      TGTGTGTGCAGGGGTGTGTGTGTGTGGGGGGGG--TTCCCAGATATTCAGGGCAAGGGAC        1087
NM_016689.2         AGAGAGAGAATGAATGTGTACATGTGTGCTGTT--TTCTAAGCTGAATGATGCAAAGGCA        1266
                             *    * *       **    *    * *                 *

NM_213468.1         GATTTCCAACTTTATATCAGACATTTGATTGTTTTTCAAGCATGACTCTTATTTCTTTCT        1334
NM_001318144.2      CAGTCGGAAGGGATTCTGG-CTATTGGG-------GGAGCCCAGAGACAGGGGAAGGCAG        1139
NM_016689.2         AGGGACCAAGTTTTCAAAACAAACTGTA-------GCAG------CTCAGGGGAAGGGAG        1313
                           **            * *           *        *

NM_213468.1         TCCACAT-----GCTTTATTCTTCCATTAAGGGAACAAGGTGTTACTTCAACAGGACGAA        1389
NM_001318144.2      CC--TGTCCATCTGTGCATAAGGAGAGGAAAGTTCCAGGGTGTGTATGTTTCAGG-----        1192
NM_016689.2         CCCAGGGGAAGGGAGAAAGTGAGTCAGGAATGTGCCAGAGTGTGCATGCTTCAGG-----        1368
                     *               *       *  ** *   **  ****   *    ****

NM_213468.1         TTATATTAATTCACTCTTTTGAAAAGTTAGTAAAAGTTTAGATCTTTGAAATGAAACTGT        1449
NM_001318144.2      ------GGCTTCACATGGAGGAGC---T-----------------GCAGATAGATAT-G        1224
NM_016689.2         ------GACTCCTCCATGTGGAGG---T-----------------GGACCCAGAAGT-G        1400
                            * * *      **     *                   *     *  *

NM_213468.1         TGTTGTTACACCTATCAATATATACAGTATTTCATGCTTTTGTGCTACTA---ATATCTA        1506
NM_001318144.2      TGTTTCTGTGTATGTGTATGTCTGCCTTTTTTTCTAAGTGGGGGCTTCTACAGGCTTTTG        1284
NM_016689.2         AGTTTCTAAGTATGCGTGTGCCTACTGTTTTTTTTTTTTTTTTTGAAATG--GACTTCTA        1458
                     ***  *     *     *   * *  * ***  *   *         *      * *

NM_213468.1         A-------ATATGTTGCCATTCACGCTGAGAAAGTTGAGAAGTTACCCAAAAAAGCTACA        1559
NM_001318144.2      GGAAGTAGGGTGGATGTGGGTAGGGCTGGGAGGAGGGGGCCACAGCTTAGGTTTGGAGCT        1344
NM_016689.2         G----------GCTTGGGGAGGGGGAAGGGATAAGAAGGGTGTAGCTCACATCTGGAGCT        1508
                               **        *  * **       *      *  *     *   *
```

```
NM_213468.1         GTACTTTTTAAAAAAATTGTAAATATTTTGTATTGTCATTTTCATTTAAAGCATTTTTTT        1619
NM_001318144.2      CTGGATGTACATACATAAGTAGGAGCAGTGG--------------------GACGTGTTT        1384
NM_016689.2         A-------TGACCCTTGACTGGGGGCTGTGT--------------------AATATGTTT        1541
                             *       *       **                      *  * ***

NM_213468.1         TTGTTGTATTACATGTATGGATTACAACTTAACATGTACACTTACGTCTGATAGTGG---        1676
NM_001318144.2      CTGTCATAATGCAGGCATGAAGGGTGGAGTGAAGTCAGGTCATAAGTTTCATGTTTGCTT        1444
NM_016689.2         CTGTTATAAGATAGACATTGGGAGGGGCTGAAGTCCAGGTCGTAAGTTTCATAATTTGTT        1601
                     ***  **    *   **            *        * ** ** * **  *

NM_213468.1         --AGAGATCCTGACATTTATTTGACATTGCAGTTTTTAGACTGTTAGATCTGTTCATTTC        1734
NM_001318144.2      TTGTTTTGTTTTGTTTTTAATGTATGTAGCAGATGTTACAGTCTTAGGGATCCGGGA--T        1502
NM_016689.2         TTTTAAATAT--ATAAATATATACATACATATATGTTACAGCCCTAGGAATAGGGGT--G        1657
                               **          *  * *** *    ***   *

NM_213468.1         ATAACATGTTAATTTTGAAAGTGAA----TGTTTTAAATCAACTGAT---TGTTATCCTA        1787
NM_001318144.2      GGGAGACCCCACTTTAGAAAGGGTCGTCACTCCTTTAATCCTCTACTCAACAATGTACTC        1562
NM_016689.2         GGAAACTCCACTTTTTAAAAGGGGTTTCCTTTCTTTAATCCTCCAATCAACAATGTACTG        1717
                       *       ***  **** *         ** ****  *   *     * * **

NM_213468.1         ATAAATTTTTTTAAAGAAGAAAAAAAAAAAAAAAAAA---------      1824
NM_001318144.2      TTTTACTTTTATATTAAAAAAAAT-AAAATAAATATGTGCCTAAAA      1607
NM_016689.2         TTGCCTTTTATATATAAAAAAGAATAAAACGTATACATGCTACAGG      1763
                     *    ***       ** ** *  ****   * *
```

Appendix C

Multiple sequence alignment, *ctsl1*. The first line in each bundle of sequences is for humans (*Homo sapiens, HS*). The second line is for Mice (*Mus musculus, MM*). The third line is for Zebrafish (*Danio rerio, DR*). The asterisks under each bundle of sequences signify DNA bases are identical. Program: Clustal Omega.

Cathepsin L1 (ctsl1)

```
XM_006517080.2      GTGCGGCGTTAGAAGCCACGTGCTGATCTAAATTCTATGGACACCCGTCCAGCCAGACCC        60
NM_001257971.2      ------------------------------------------------------------        0
XM_004283041.2      ------------------------------------------------------------        0

XM_006517080.2      TGGAGAGAACCAAGGCACCACGCTGGCGTGGACCCTCCAAAGTGGATCAGCCTCGCCCCA        120
NM_001257971.2      ------------------------------------------------------------        0
XM_004283041.2      ------------------------------------------------------------        0

XM_006517080.2      AGGCTTGCACCACAGGACAGGTTACGACCCGGCGGCGTCACGCGCCCGGACTCCCGCAGG        180
NM_001257971.2      ------------------------AGAACC-GCGACCT-----CCGCAACCTTGAGCGGC        30
XM_004283041.2      ------------------------------------------------------------        0

XM_006517080.2      CTCCGC-------CCCGAGGCAGGCATAGCCAATGACGGGGCGGGGGCGGGCCCTGTCGG        233
NM_001257971.2      ATCCGTGGAGTGCGCCTGCGCAGCTACGACCGCAGCAGGAAAGCGCCGCCGGCCAGGCCC        90
XM_004283041.2      ----------------------------------------------------CGGAGCC        7
                                                                         * *

XM_006517080.2      GGCTGTAGCCTGAGAGGCTTTAAAGCCTGAGCCCGGCGCTGTTCCTCCAGATTCTCGGAC        293
NM_001257971.2      AGCTGTGGCCGGACAG-------------------GGACTGGAAGAGAGGACGCGGTCGA        131
XM_004283041.2      TTAAATGGGAGGACAT-------------------GCGCCCACAGCCCAATCTCCGACGC        48
                         * *   ** *                   *  *             *

XM_006517080.2      CTCGGCGACCTCCGGGGATCCGAGTTTGCAGACTTCTTGTGCGCAGCTAGCCGCCTCAGG        353
NM_001257971.2      GTAGGTGTGCACCA-GCCCTGGCAACGAGAGCGTCTACCCCGAACTCTGCTGGCCTTGAG        190
XM_004283041.2      CCCTGGA---------TCCATAGTTGTTGGCTGCGCCTGCGCAGCTGCGCCTGCCGCAGG        99
                        *                                                ***    *

XM_006517080.2      TGTTTGAACCATGAATCTTTTACTCCTTTTGGCTGTCCTCTGCTTGGGAACAGCCTTAGC        413
NM_001257971.2      GTTTTAAAACATGAATCCTACACTCATCCTTGCTGCCTTTTGCCTGGGAATTGCCTCAGC        250
XM_004283041.2      CTTTTGAAACATGAATCCTTCACTCATCCTGACTGCCCTTTGCTTGGGAATAGCCTCAGC        159
                      *** ** ******** *  **** *  *  *** * * *** ******  **** ***

XM_006517080.2      TACTCCAAAATTTGATCAAACCTTTAGTGCAGAGTGGCACCAGTGGAAGTCCACACACAG        473
NM_001257971.2      TACTCTAACATTTGATCACAGTTTAGAGGCACAGTGGACCAAGTGGAAGGCGATGCACAA        310
XM_004283041.2      GGCTTCAAAACTTGATCTCAGTTTAAATGCACAATGGAAATTGTGGAAGGCAACACACAG        219
                      **  ** * ******  *  **    *** * ***     ******* * *  ****

XM_006517080.2      AAGACTGTATGGCACGAATGAGGAAGAGTGGAGGAGAGCGATATGGGAGAAGAACATGAG        533
NM_001257971.2      CAGATTATACGGCATGAATGAAGAAGGATGGAGGAGAGCAGTGTGGGAGAAGAACATGAA        370
XM_004283041.2      GAAACTATATGGCTTGAATGAAGAAGGATGGAGGAGAACAGTGTGGGAGAAGAATATGAA        279
                     * * * ** ***  ****** ****  ********* *  * ********** ****
```

```
XM_006517080.2      AATGATCCAGCTACACAACGGGGAATACAGCAACGGGCAGCACGGCTTTTCCATGGAGAT        593
NM_001257971.2      GATGATTGAACTGCACAATCAGGAATACAGGGAAGGGAAACACAGCTTCACAATGGCCAT        430
XM_004283041.2      AATGATTGACCTGCACAATCAGGAATACAGCCAAGGGAAACATGGCTTCACCATGGCAAT        339
                     *****  * ** *****   *********  * *** * **  ****  * ****  **

XM_006517080.2      GAACGCCTTCGGTGACATGACCAATGAGGAATTCAGGCAGGTGGTGAATGGCTACCGCCA        653
NM_001257971.2      GAACGCCTTTGGAGACATGACCAGTGAAGAATTCAGGCAGGTGATGAATGGCTTTCAAAA        490
XM_004283041.2      GAATGCCTTTGGTGACATGACCAGTGAAGAATTCAGGCAGGTGATGAATGGTCTTCGAAA        399
                    *** ***** ** ********** *** *************** *******    *   *

XM_006517080.2      CCAGAAGCACAAGAAGGGGAGGCTTTTTCAGGAACCGCTGATGCTTAAGATCCCCAAGTC        713
NM_001257971.2      CCGTAAGCCCAGGAAGGGGAAAGTGTTCCAGGAACCTCTGTTTTATGAGGCCCCCAGATC        550
XM_004283041.2      CCAGAAGCGCAAGAAGGGGAAAGTGTTCCCAGTACCCGTCTTTGGTTTGATTCCCTCATC        459
                    **  **** ** ********   * ** *  * ***  *  *   *  *   ***   **

XM_006517080.2      TGTGGACTGGAGAGAAAAGGGTTGTGTGACTCCTGTGAAGAACCAGGGCCAGTGCGGGTC        773
NM_001257971.2      TGTGGATTGGAGAGAGAAAGGCTACGTGACTCCTGTGAAGAATCAGGGTCAGTGTGGTTC        610
XM_004283041.2      CATGGACTGGAGAGAGAAAGGCTATGTAACTCCTGTGAAGAATCAGGGTATGTGTGGTTC        519
                      **** ******** ** ** *  ** ************** *****   *** ** **

XM_006517080.2      TTGTTGGGCGTTTAGCGCATCGGGTTGCCTAGAAGGACAGATGTTCCTTAAGACCGGCAA        833
NM_001257971.2      TTGTTGGGCTTTTAGTGCTACTGGTGCTCTTGAAGGACAGATGTTCCGGAAAACTGGGAG        670
XM_004283041.2      TTGTTGGGCTTTTAGTGCAACTGGTGCCCTTGAAGGACAGATGTTCCGGAAAACTGGCAA        579
                    ********* ***** **  * ***   ** ****************  ** ** ** *

XM_006517080.2      ACTGATCTCACTGAGTGAACAGAACCTTGTGGACTGTTCTCACGCTCAAGGCAATCAGGG        893
NM_001257971.2      GCTTATCTCACTGAGTGAGCAGAATCTGGTAGACTGCTCTGGGCCTCAAGGCAATGAAGG        730
XM_004283041.2      ACTTATTTCACTGAGTGAGCAGAACCTGGTGGACTGCTCTCGGCCTCAAGGCAATGAGGG        639
                     ** ** *********** ***** ** ** ***** ***    *********** * **

XM_006517080.2      CTGTAACGGAGGCCTGATGGATTTTGCTTTCCAGTACATTAAGGAAAATGGAGGTCTGGA        953
NM_001257971.2      CTGCAATGGTGGCCTAATGGATTATGCTTTCCAGTATGTTCAGGATAATGGAGGCCTGGA        790
XM_004283041.2      TTGCAATGGTGGCCTAATGGATAATGCCTTCCAGTATGTTAAGGACAACGGAGGCCTGGA        699
                     ** ** ** ***** ******  *** ********  ** **** ** ***** *****

XM_006517080.2      CTCGGAGGAGTCTTACCCCTATGAAGCGAAGGACGGATCTTGTAAATACAGAGCCGAGTT        1013
NM_001257971.2      CTCTGAGGAATCCTATCCATATGAGGCAACAGAAGAATCCTGTAAGTACAATCCCAAGTA        850
XM_004283041.2      CTCAGAGGAATCCTATCCATATTTTGGAAAGGATGAATCCTGCCACTACAGACCCCAGAG        759
                    *** ***** ** ** ** ***   *  *  ** * *** **  * ****   ** **

XM_006517080.2      CGCTGTGGCTAATGACACAGGGTTCGTGGATATCCCTCAGCAAGAGAAAGCCCTCATGAA        1073
NM_001257971.2      TTCTGTTGCTAATGACACCGGCTTTGTGGACATCCCTAAGCAGGAGAAGGCCCTGATGAA        910
XM_004283041.2      TTCTGCTGCCAACGACACTGGTTTCGTGGACATTCCTAAGCAAGAGAAGGCCCTTATGAA        819
                      ***  ** ** ***** ** ** ***** ** *** **** ***** ***** *****
```

```
XM_006517080.2      GGCTGTGGCGACTGTGGGGCCTATTTCTGTTGCTATGGACGCAAGCCATCCGTCTCTCCA       1133
NM_001257971.2      GGCAGTTGCAACTGTGGGGCCCATTTCTGTTGCTATTGATGCAGGTCATGAGTCCTTCCT       970
XM_004283041.2      GGCAGTGGCAACTGTGGGCCCCATCTCTGTTGCTATAGATGCAAGCCATCCAACCTTCCA       879
                    *** ** ** ******** ** ** ********** ** *** * ***    *  ***

XM_006517080.2      GTTCTATAGTTCAGGCATCTACTATGAACCCAACTGTAGCAGCAAGAACCTCGACCATGG       1193
NM_001257971.2      GTTCTATAAAGAAGGCATTTATTTTGAGCCAGACTGTAGCAGTGAAGACATGGATCATGG       1030
XM_004283041.2      GTTCTATAAAGCAGGCATTTATTATGACCCACACTGCAGCAGTGAAGACCTGGATCATGG       939
                    ********    ****** ** * *** **  **** *****  *  ** * ** *****

XM_006517080.2      GGTTCTGTTGGTGGGCTATGGCTATGAAGGAACAGATTCAAATAAGAATAAATATTGGCT       1253
NM_001257971.2      TGTGCTGGTGGTTGGCTACGGATTTGAAAGCACAGAATCAGATAACAATAAATATTGGCT       1090
XM_004283041.2      TGTTCTGGCGGTTGGCTATGGGTTTGAAGGAGCAGACTCAGATAACAATAAATACTGGCT       999
                     ** ***  *** ***** ** * **** *  **** *** **** ******** *****

XM_006517080.2      TGTCAAGAACAGCTGGGGAAGTGAATGGGGTATGGAAGGCTACATCAAAATAGCCAAAGA       1313
NM_001257971.2      GGTGAAGAACAGCTGGGGTGAAGAATGGGGCATGGGTGGCTACGTAAAGATGGCCAAAGA       1150
XM_004283041.2      TGTCAAGAACAGCTGGGGTGAAGATTGGGGCATGGACGGCTACATAAAGATGGCCAAAGA       1059
                     ** **************    ** ***** ****  ****** * ** ** ********

XM_006517080.2      CCGGGACAACCACTGTGGACTTGCCACCGCGGCCAGCTATCCTGTCGTGAATTGATGGGT       1373
NM_001257971.2      CCGGAGAAACCATTGTGGAATTGCCTCAGCAGCCAGCTACCCCACTGTGTGAGCTGGTGG       1210
XM_004283041.2      CCGGGACAACAACTGTGGAATCGCCACCATGGCCAGTTATCCTACTGTCTAAACCTCAAA       1119
                    ****   *** * ****** * *** *    ***** ** **    **

XM_006517080.2      AGCGGTAATGAGGACTTATGGACACTATGT----------CCAAAGGAATTCAGCTTAAA       1423
NM_001257971.2      ACGGTGATGAGGAAGGACTTGACTGGGGATGGCGCATGCATGGGAGGAATTCATCTTC-A       1269
XM_004283041.2      GGCAGGATTGAGAAC---------------AGAACATCCAGAGGAAGAATTTTCTTTAAA       1164
                          *    * *                          * *****    **  *

XM_006517080.2      ACTGACCAAACCCTTATTGAGTCAAACCATGGTACTTGAATCATTGAGGATCCAAGTCAT       1483
NM_001257971.2      GTCTACCAGCCCC-CGCTGTGT-CGG-ATACACACTCGAATCATTGAAGATCCG-AGTGT       1325
XM_004283041.2      ACTAACCAGACCT-TACTGTGT-GGGATGAAACACTTGAATCGTTGAAGATCCAAGTTGT       1222
                        ****  **     ** **           *** ***** **** *****      *

XM_006517080.2      GATTTGAATTCTGTTGCCATTTTTACATGGGTTAAATGTTACCACTACTTAAAACTCCTG       1543
NM_001257971.2      GATTTGAATTCTGTGATATTTTCACACTGG--TAAATGTTACCTCTATTTTAA-TTACTG       1382
XM_004283041.2      GATTTGAATTCTGTGACATTTTATGCTGGT--GAAATGTTACCACTCCTTT-A-ATACTA       1278
                    **************     ***      *    ********** **  **  *  * **

XM_006517080.2      TTATAAACAGCTTTATAATATTGAAAACTTAGTGCTTAATTCTGAGTCTGGAATATTTGT       1603
NM_001257971.2      CTATAAATAGGTTTATATTATTGATTCACT---------TACTGACTTTGCATT-T-T-C       1430
XM_004283041.2      TTGTAAATAGGTTTGTATGATCGACTTGCC---------TCAAGAATTTTGAATTTTC-A       1328
                     * **** ** *** **  ** **               *   ** * *  * * *



XM_006517080.2      TTTATATAAAGGTTGTATAAAACTTTCTTTACCTCTTAAAAATAAATTTTA---------       1654
NM_001257971.2      GTTTTTAAAAGGATGTATAAATTTTTACCTGTTTAAATAAAATTTAATTTCAAATGTAGT       1490
XM_004283041.2      TGTTTTAAAAGGATGTATAAAGTTTTACCTTTAAAATAAAACAATTTCAGAGTA------       1382
                      * *  ***** ********  ***   *        ***

XM_006517080.2      ---------------------------GCTCAGTGTGTGTGTA-----------------       1670
NM_001257971.2      GGTGGGGCTTCTTTCTATTTTTGATGCACTGAATTTTTGTGTAATAAAGAACATAATTGG       1550
XM_004283041.2      ------------------------------------------------------------       1382


XM_006517080.2      --------------      1670
NM_001257971.2      GCTCTAAGCCATAA      1564
XM_004283041.2      --------------      1382
```

Appendix D

Multiple sequence alignment, *cfl1*. The first line in each bundle of sequences is for humans (*Homo sapiens, HS*). The second line is for Mice (*Mus musculus, MM*). The third line is for Zebrafish (*Danio rerio, DR*). The asterisks under each bundle of sequences signify DNA bases are identical. Program: Clustal Omega.

Cofilin 1 (cfl1)

```
NM_213641.2     -----------------------------------GCAAAAGGGCGCGCCCGAGTGT   22
NM_005507.3     ------------------------------------------------------------  0
NM_007687.5     GCCGGAAGGCCGCCCCGCCCCTCATTGTGCGGCTCCTACTAAACGGAAGGGGCCGGGAGA   60


NM_213641.2     TGTA---TCAAGTGTGCGAGACCGAGGGACTGCATCACAG-CTCATTCAATTCCAATCCA   78
NM_005507.3     ------GTTCAGTCGGGTCCCGGCAGCGGCTGCAGCGCTCTCGTCTTCTGCGGCTCTCGG   54
NM_007687.5     GGCCGCGTTCAGTCGGATCCCGGCAGCAGCTGCGGCGGC--TCAGGTCTTCAGTTGCTCT   118
                       *  ***  *        **   ****  *          **


NM_213641.2     GACGCCGACTAGCACAGAGACACAGCCATGGCCTCCGGAGTTACAGTGGAAGAGACTGTG   138
NM_005507.3     TGCCCTCTCCTTTTCGTTTCCGGAAACATGGCCTCCGGTGTGGCTGTCTCTGATGGTGTC   114
NM_007687.5     TTCTGTCTCCCTTTCGTTTCCGGAAACATGGCCTCTGGTGTGGCTGTCTCTGATGGTGTC   178
                  *     *     *     *  *  ********* ** **  * **    **   ***


NM_213641.2     TTGACGGTCTTCAACGAGATGAAGGTCCGCAAGGCTCATTGCAATGAGGAAGAGAAGAGC   198
NM_005507.3     ATCAAGGTGTTCAACGACATGAAGGTGCGTAAGTCTTCAACGCCAGA---GGAGGTGAAG   171
NM_007687.5     ATCAAGGTGTTCAATGACATGAAGGTTCGCAAGTCTTCAACACCAGA---AGAAGTGAAG   235
                 * * *** ***** ** ******** ** *** **         **    **   **


NM_213641.2     AAGAGGAAGAAGGCCGTCATGTTCTGCCTGAGCGACGACAAGAAGCACATCATCATGGAG   258
NM_005507.3     AAGCGCAAGAAGGCGGTGCTCTTCTGCCTGAGTGAGGACAAGAAGAACATCATCCTGGAG   231
NM_007687.5     AAACGCAAGAAGGCGGTGCTCTTTTGCCTGAGTGAGGACAAGAAGAACATCATCCTGGAG   295
                **  * ******** **  * ** ******** ** ********* ******** *****


NM_213641.2     CAAGGCCAGGAGATCCTGCAGGGCGATGAA------------GGAGATCCCTACCTTAAG   306
NM_005507.3     GAGGGCAAGGAGATCCTGGTGGGCGATGTGGGCCAGACTGTCGACGACCCCTACGCCACC   291
NM_007687.5     GAGGGCAAGGAGATCCTGGTAGGAGATGTGGGGCAGACTGTGGACGACCCCTACACCACT   355
                 * *** ***********   ** ****             *  ** ******   *


NM_213641.2     TTTGTGAAGATGCTTCCTCCCAATGATTGTCGCTACGCTCTCTATGATGCCACGTATGAG   366
NM_005507.3     TTTGTCAAGATGCTGCCAGATAAGGACTGCCGCTATGCCCTCTATGATGCAACCTATGAG   351
NM_007687.5     TTTGTCAAGATGCTGCCAGACAAGGACTGCCGCTATGCACTCTATGATGCAACCTATGAG   415
                ***** ******** **    ** ** ** ***** ** ********** ** ******


NM_213641.2     ACTAAAGAGACCAAGAAAGAGGACCTGGTTTTCATCTTCTGGGCCCCAGAGAGCGCTCCA   426
NM_005507.3     ACCAAGGAGAGCAAGAAGGAGGATCTGGTGTTTATCTTCTGGGCCCCCGAGTCTGCGCCC   411
NM_007687.5     ACCAAGGAGAGCAAGAAGGAGGACCTGGTGTTCATCTTCTGGGCCCCCGAGAATGCACCC   475
                ** ** **** ****** ***** ***** ** ************** ***   ** **


NM_213641.2     CTTAAAAGCAAGATGATCTACGCCAGCTCCAAGGATGCCATCAAGAAGAAGTTCACAGGT   486
NM_005507.3     CTTAAGAGCAAAATGATTTATGCCAGCTCCAAGGACGCCATCAAGAAGAAGCTGACAGGG   471
NM_007687.5     CTCAAGAGCAAAATGATCTATGCCAGCTCCAAGGATGCCATCAAGAAGAAGCTGACAGGA   535
                ** ** ***** ***** ** ************** ************** * *****
```

```
NM_213641.2      ATTAAGCACGAGTGGCAAGTGAATGGTATGGACGATATCAAGGACCGAAAGACCCTTGCT   546
NM_005507.3      ATCAAGCATGAATTGCAAGCAAACTGCTACGAGGAGGTCAAGGACCGCTGCACCCTGGCA   531
NM_007687.5      ATCAAGCATGAATTACAAGCTAACTGCTACGAGGAGGTCAAGGACCGCTGCACCCTGGCA   595
                 ** ***** ** *  ****  **  *    ** **  **********    ***** **

NM_213641.2      GAGAAGCTCGGGGGCGCATCGGTGGTGTCTCTGGAGGGAAAGCCTCTAACCGATTGAGGC   606
NM_005507.3      GAGAAGCTGGGGGGCAGTGCCGTCATCTCCCTGGAGGGCAAGCCTTTGTGAGCCCCTTCT   591
NM_007687.5      GAGAAACTAGGTGGCAGCGCCGTCATTTCCCTGGAGGGCAAGCCTTTGTGAGCCACCTCC   655
                 ***** ** ** ***    * **  * ** ******** ****** *    *

NM_213641.2      TGACACATTTCAGGGGTTTAGCCGTTCATTCCGACATGGGTAGGGCAGATGGGCACAGCA   666
NM_005507.3      GGCCCCCTGCCTGGAGCATCTGG-CAGCCCCACACCTGCCCTTGGGGGTTGCAGGCTGCC   650
NM_007687.5      AGCCCCCTGCCTGGAGCATCTTAACAGCCCCAGACCTGCTCTTGGGTGTTGCAGGCTGCC   715
                  * * * *  * ** *  *           *  ** **     **  * **    * **

NM_213641.2      CACCAC-TGTTCGTGGGCCATGGGTGGGTTAATGCGGGT--GGGGAGGGTCGGGCAAAAG   723
NM_005507.3      CCCTTCCTGCCAGACCGGAGGGGCTGGGGGGATCCCAGCAGGGGGAGGGCAATCCCTTCA   710
NM_007687.5      CCTTTCCTGCCAGACCGGAGGGGCTGGGGGGATCCCAGCAGGGGGAGGGCTATCCCTTCA   775
                 *    * **   *   *    ** ****   ** *  *   ********     *

NM_213641.2      TGACAGTTTCCAAACTCCACACGACGAAAGATGTAGGCTGTCATTCCAGTTCACACATAC   783
NM_005507.3      CCCCAGTTGCCAAACAGACCCCC---CACCCCCTGGATTTTCCTTCTCCCTC---CATCC   764
NM_007687.5      CCCCAGTTGCCAAACATCCCTCC---CACCCCCTGGACCGTCCTTCTCCCTC---CATCC   829
                    ***** ******    * *     *     * *    ** ***    **   *** *

NM_213641.2      AATGAACAAGAACAAAAACACAAAAAGAGAAACAAAAAAAAAGATCTTAAAAGACAAAAA   843
NM_005507.3      CTTGACGGTTCTGGCCTTCCCAAACTG----------CTTTTGATCTTTTGATTCCTCT-   813
NM_007687.5      -CTGACGGTTCTGGCCTTCCCAAACTG----------CTTTTGATCTTCTGATTCCTCT-   877
                   ***             * ****  *             ******   *  *

NM_213641.2      TGAAGGATGATGATTATAAAAACAAAGTAATATAACTGTAAATGTGTACTGGCAGGTTTT   903
NM_005507.3      --------TGGGCTGAAGCAGACCAAGTTCCCCCCAGGCACCCCAGTTGTGGGGGAGCCT   865
NM_007687.5      --------TGGGTTGACGCAGACCAAGTCCCGTCCTAGGCACCCAGTTTGGGGGGAGCCT   929
                           * * *   * ** ****         *       **   **  *    *

NM_213641.2      CCTTCTTTCTGGGTTCCATTAGATTGTCAATACGAAAGACTGAGACAGCAGCCATTACAA   963
NM_005507.3      GTATTTTTTTTAACAACATCC--------CCATTCCC------CACCTGGTCCTCCCCCT   911
NM_007687.5      GTATTTTTTTTTTTAACGACA--------CCCC--TA------CTCCGTATCCCTCCCCA   973
                    * *** *      *                           *     **    *

NM_213641.2      TTCAGCTATACAACATACTGTTTGTTCTTAAGATTTAGTTTTTTTTTTTT--TTTTTTTT   1021
NM_005507.3      TCCCATGCTGCCAACTTCTAACCGCAATAGTGACTCTGTGCTTGTCTGTTTAGTTCTGTG   971
NM_007687.5      TCCCATGCTGCCAACTTCTAACCACAATAGTGACTCTGTGCTTGTCTGTTTAGTTCTGTG   1033
                 * *     * * *  * **       *   ** *  **  ** * * * **   ** * *
```

```
NM_213641.2      GCCAAACGTCTAGTTAAGTTTATTTTCTAAGTTTGTGTGTGTTGCGACTAGGTATAGTAC   1081
NM_005507.3      TATAAATGGAATGTTGTGGAGATGACC------CCTCCCTGTGCCGGCTGGTTCCTCTCC   1025
NM_007687.5      TGTAAATGAAATGTGGAAATGACCCTC----------CCTGCCCCAGCTGGCTGCCCTCC   1083
                    *** *    **       *    *           **   *  ** * *    * *

NM_213641.2      AAAGTCACACATTGTTATCGGACCATTCTGGGAACACGATCTACTCCCCTTTCCCTCTTA   1141
NM_005507.3      CTTTTCCC---CTGGTCACGGCTACTCATGGAAGCAGGACCAG-TA--------------   1067
NM_007687.5      CCTTTCCT---TTGATCTTGACCACTCATGGAAGCAGGACCAG-TA--------------   1125
                     **      ** *   *     *  *** * ** ** *   *

NM_213641.2      TTTTATTTTTGAGCAAGTTATTTTAATGATTCCATCTGATTTAAATCTTTCCAAACGTTC   1201
NM_005507.3      ----------------------AGGGACCTTCG-----ATTAAAAAAAAAAAAAGACAA-T   1099
NM_007687.5      ----------------------AGGGACCTTCA-----ATTTAAAACAAAACAAAACA-A   1157
                                       ***      *** ***       * *

NM_213641.2      AGTCAAAATCATGGCACTGGTGTCAAGAAATTGTACACATTCCTTCTACCTGTATAATCT   1261
NM_005507.3      AATAAAA------AGGCTCATTAATGGGATGTGTTTTT-------CAAGGTTGGGACACA   1146
NM_007687.5      AAAAACAATAAAAAGGCTAATTAACAAAAAAAAAAAAA-------A--------------   1196
                 *   * *         **  *       *

NM_213641.2      GGGACCTAGTGGATGAGCGGTCTGTTGAAGAATTTGAATTGAAGGACACGTATGAAGTTT   1321
NM_005507.3      GGAGACTTCAGGATTGGGGGTGCACTGGGATAATGGTGCTATCTCTCTCCTCTGTACTCC   1206
NM_007687.5      ------------------------------------------------------------   1196

NM_213641.2      AGATGGGGAGAGCGATATATTTGGGCAGCCTGTGTACAGAGCTTTGATGGATATTGGTCA   1381
NM_005507.3      ACACC---------------CAACCAGCCCCTGCACACCCTTATCT-GGCTCCAA----   1245
NM_007687.5      ------------------------------------------------------------   1196

NM_213641.2      GTGGATGTTTTGTATCGTCTCCTTATCTCTAATAAAAAAGCACTAAATAACTCTGCTGTG   1441
NM_005507.3      ------------------------------------------------------------   1245
NM_007687.5      ------------------------------------------------------------   1196

NM_213641.2      TTACAACTTTTTTTTTTGTTTCATGTTGGGTGAATTATGTGATAAGGGCTGTAACTTGTC   1501
NM_005507.3      ------------------------------------------------------------   1245
NM_007687.5      ------------------------------------------------------------   1196

NM_213641.2      TTTTTTGGCCCTGGTTGAATCTAATTTTACTCTTAGGTTTAAGTTAAGTCCATATGAGTT   1561
NM_005507.3      ------------------------------------------------------------   1245
NM_007687.5      ------------------------------------------------------------   1196
```

```
NM_213641.2      GATCTTTTATATTGGTTTATAGAAATCTGCTTTACATTTTGTGTCTGGATATTGTGCATT   1621
NM_005507.3      ------------------------------------------------------------   1245
NM_007687.5      ------------------------------------------------------------   1196

NM_213641.2      TGTTGCTGTTTTATCAGTGTTCAGCATGGTTTATGCACTGGTATATGTGATGGTTTGCAA   1681
NM_005507.3      ------------------------------------------------------------   1245
NM_007687.5      ------------------------------------------------------------   1196

NM_213641.2      CATTCATAAAGTTTCAAATGGATCCAAA   1709
NM_005507.3      ---------------------------   1245
NM_007687.5      ---------------------------   1196
```

Appendix E

Multiple sequence alignment, *fapb2*. The first line in each bundle of sequences is for humans (*Homo sapiens, HS*). The second line is for Mice (*Mus musculus, MM*). The third line is for Zebrafish (*Danio rerio, DR*). The asterisks under each bundle of sequences signify DNA bases are identical. Program: Clustal Omega.

Fatty Acid Binding Protein 2 (fabp2)

```
NM_007980.3     CTTAGAACTGGCTGCCTCTGCCTCCGGAGAGCAGCGATTAAAAGTGTCAGCCATCATTAC   60
NM_000134.4     ------------------------------------------------------------   0
NM_131431.1     ------------------------------------------------------------   0

NM_007980.3     CTGGCCCTAATTCTTGAAATAAAAATGCCCACATGCTGTAGTTGAAGGCAGAGTAGGAAT   120
NM_000134.4     ------------------------------------------------------------   0
NM_131431.1     ------------------------------------------------------------   0

NM_007980.3     GATTATCAGATTTTAATTCAGTTGAATCCCAGCAGTAGATTCAAAGAAAGCATGGGAAGA   180
NM_000134.4     ------------------------------------------------------------   0
NM_131431.1     ------------------------------------------------------------   0

NM_007980.3     GAAACCAAAGGGGCCTGGCATGTGAGGCGGTTAGGTTATCTCCTGAACTTCGAACTTCCA   240
NM_000134.4     ------------------------------------------------------------   0
NM_131431.1     ------------------------------------------------------------   0

NM_007980.3     CATCACAGTATGAATTGGTTCGAAGATAAGAAATAGAATAAATTCTCTCTAGTGGACAGG   300
NM_000134.4     ------------------------------------------------------------   0
NM_131431.1     ------------------------------------------------------------   0

NM_007980.3     ACTGGACCTCTGCTTTCCTAGAGACACACACAGCTGAGATCATGGCATTCGACGGCACGT   360
NM_000134.4     -----ATCTCTAGCTGCCTAGAGGCTGACTCAACTGAAATCATGGCGTTTGACAGCACTT   55
NM_131431.1     ----------------------------TCTGTCATCATCATGACCTTCAACGGGACCT   31
                                            *      ****** * **  ** * ** *

NM_007980.3     GGAAAGTAGACCGGAACGAGAACTATGAAAAGTTCATGGAGAAAATGGGCATTAATGTGA   420
NM_000134.4     GGAAGGTAGACCGGAGTGAAAACTATGACAAGTTCATGGAAAAAATGGGTGTTAATATAG   115
NM_131431.1     GGAAAGTCGACCGCAATGAGAACTACGAGAAGTTCATGGAACAAATGGGCGTCAACATGG   91
                **** ** ***** *  ** ***** ** **********  *******  * **  *

NM_007980.3     TGAAGAGGAAGCTTGGAGCTCATGACAATCTGAAACTGACAATCACACAGGATGGAAATA   480
NM_000134.4     TGAAAAGGAAGCTTGCAGCTCATGACAATTTGAAGCTGACAATTACACAAGAAGGAAATA   175
NM_131431.1     AGAAAAGGAAACTGGCTGCCCATGACAACCTGAAGATCACCCTGGAGCAGACCGGAGACA   151
                 *** ***** ** *  ** ********  ****  * **  *    **    *** * *

NM_007980.3     AATTCACAGTCAAAGAATCAAGCAACTTCAGAAACATTGATGTTGTGTTTGAGCTCGGTG   540
NM_000134.4     AATTCACAGTCAAAGAATCAAGCACTTTTCGAAACATTGAAGTTGTTTTTGAACTTGGTG   235
NM_131431.1     AGTTCAACGTGAAGGAAGTCAGCACTTTCCGCACACTGGAAATTAACTTTACTCTGGGCG   211
                * ****  ** ** ***   ****  **  * *   * **  **   ***   ** ** *
```

```
NM_007980.3      TAAACTTTCCCTACAGTCTAGCAGACGGAACGGAGCTCACTGGGGCCTGGACCATTGAGG   600
NM_000134.4      TCACCTTTAATTACAATCTAGCAGACGGAACTGAACTCAGGGGGACCTGGAGCCTTGAGG   295
NM_131431.1      TCACCTTTGACTATTCTCTGGCAGACGGCACTGAGCTCACAGGATCCTGGGTCATAGAGG   271
                 * * ****   **   *** ******** ** ** ****  **  *****  * * ****

NM_007980.3      GAAATAAACTTATTGGGAAATTCACACGTGTAGACAATGGAAAGGAGCTGATTGCTGTCC   660
NM_000134.4      GAAATAAACTTATTGGAAAATTCAAACGGACAGACAATGGAAACGAACTGAATACTGTCC   355
NM_131431.1      GAGACACGCTCAAGGGGACTTTCACACGCAAGGACAACGGAAAGGTACTAACAACAGTCA   331
                 ** * *  ** *  ** *  **** ***    ***** ***** *  ** *   * ***

NM_007980.3      GAGAGGTTTCTGGTAATGAACTAATCCAGACCTACACATATGAAGGAGTTGAGGCCAAGC   720
NM_000134.4      GAGAAATTATAGGTGATGAACTAGTCCAGACTTATGTATATGAAGGAGTAGAAGCCAAAA   415
NM_131431.1      GGACTATCGTTAATGGTGAACTTGTACAGAGCTATAGCTATGATGGAGTCGAGGCCAAGA   391
                 *    *      *  ******  * ****  **    ***** ***** ** *****

NM_007980.3      GATTCTTTAAGAAGGAATAAGTCAACTTCTCAGAGCCTGGAGCAACGCTGAAGAGCTAAG   780
NM_000134.4      GGATCTTTAAAAAGGATTGAGCATTATTCTTGGCGCACAGTCCAAAATACAAATTGGACA   475
NM_131431.1      GGATTTTCAAGAGGGCTTAAACTGTTATCTGCACTCCCAATGTCAAACTTACAACTGAAG   451
                 *  * ** ** * **  * *       ***     *        *     *      *

NM_007980.3      CTGATGTCAGATTTCTT--TCTCCATCATGCTAATGCC--------------AGGCT---   821
NM_000134.4      GAAGATCTATATTGTAC-----CAGAACTATTTATTTCACCCCATCAAGTATAAGGTTAC   530
NM_131431.1      TGGGACA-ATATGAAACTTTATAGTGTTTGAATATTTAACCTGA----AAATAGCTTTAA   506
                        * **               *    **                *   *

NM_007980.3      -CATTCGT--------CATCCTATCAGCA----CTGGTCTCCAGCCTTGTCAAAGCTAAA   868
NM_000134.4      TGATTGATTGGT----CCTTTTATAAACATTGGTATATTTCCATTCATGCCAAAGCAAAA   586
NM_131431.1      TGATTGGCAGCCCTAAAGTTTACAAAACATTTATGTAAATCGTGTCATTTGCA-------   559
                   ***            *      * **         **    * *    *

NM_007980.3      GAAGTAAAAGCTAATTAAAAGAACTTCATTTGTTTTATGATCCTTAAGCTATACATGAAC   928
NM_000134.4      GAAGTAAAAGCTAATTAGGATTTAA---TTTGTTTTATATTCTCTAAGATATATATTTAC   643
NM_131431.1      --------------------CCTGT---GAGTTTTGATATTTATGATAATAAAGCTATCT   596
                                          *** **  *    *   ** *  *

NM_007980.3      TAGTCTTTTAAAAGAAAATAAATCCTGTTCTCACACAAAAAAAAAAAAA----------   977
NM_000134.4      TA--------AAAGAATTTGTGACATTTTAAAAAACAAAAATAAATATTGCGTCCATGTT   695
NM_131431.1      TA--------A-GGCCCTTTTGTGAATGGAAAAAAAAAAAAAAAAAAAAAAAAAAA--   645
                 **        *  *    *            * * ***** *** *

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      GCTTTATATGTAGCCTTGCCTTTTAAAAGAAAAAGTATGTGAATATGAATTGACAGACTG   755
NM_131431.1      ------------------------------------------------------------   645
```

```
NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TTTTCGTAGAGAGAGGGTCTTACTCTTTCACTCAGGCTGGAATGTAGTGGAGAGATCATA   815
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      GCTCACTGTAACCTCAAACTCCTGGACTCATGCAATCTTCCTGCCTCAGGCTTCTGAGTA   875
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      GCTAGGACTATGGGTACATTCCACAGTGCCCAGCTAATTTTTGTTTTGTTTTCTTTTTAT   935
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TTTTTTTAGAGATGGGGTCTTGCTATATTGCCCAGGCTGGTCTTGAACCCCTGGCCTCAA   995
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      GCAATCCTCCTGCCTCAGCCTCTCAAGTTGTTTTTTTTCTTTACATTTGATAAACTAAAA   1055
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      GCATAGGCTGCATATGAGTCTTTAACATCTTGAACTGGTTGTGAATAATTTTCTGGCACT   1115
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      GGTTGTAAGTAATATCTATTATTATAAAAATAATATATGCTCAACCAGAAAACTTAGAAA   1175
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TAAGAAACACAAATGTAAAATAAGTATTTCCATAACTCATAATCCAGAGATAATTGCCAT   1235
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TCTGATTTTGATAGATATCCTCTCAGCTCTCTTCCCTGGGGGCAGATATTTCCCAATACA   1295
NM_131431.1      ------------------------------------------------------------   645
```

```
NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TACCACTTTGAATAGGATGATAGGAAATAAATGATGTACTACATTAAATTAAATTATTGT   1355
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      ATTACATTTTTGTACACATCAGTCATTCCCACGCTTGGCTGAAAATCAGGATCATCTGAG   1415
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      AAACTTAAACAATTTCTGCATTCTTAATCTCCACTGTTATTCTATTATATCAGAATCGCT   1475
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      AATAGAACCAAGAATTCTAGAAAATTTCCTGGTGATTCTGATGCAGCCTGTCCACTAACT   1535
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TGTCTTTGAGAACAATGGAGATATCAGTTATCAATGTTATTTTTAACCACCCCCTTCTTT   1595
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TTTGTTGTTGTGGGTTTTTTCACATAAACACATACATTGCCAATTTTCCAGGGCTCAAAA   1655
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      AGTTTCCTTTTATTCATATTTTCACATGACGTAAAATTTTATGTGCTTCACATAATTGTA   1715
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TTTTAGCAGGGTACATATTAGGGGATGGAGAGGGGGTACAATTTTTAAGCATTTGCAGCT   1775
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      GCAACTCTATAGACTTTTGACAAATTCTTTTTCACACTGATGTAGATAACAGCTTAACAT   1835
NM_131431.1      ------------------------------------------------------------   645
```

```
NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TTACATAGTTCTTACTACTTGCCACACACTGTTCTAAGTGGTATATACATATGACATATA   1895
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TATTAAAATGTAGGAAGAAAAATGTTTGAGTACCTGATGAAAATGAATAGAGAGTATGTA   1955
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      ATCTTTGAAAGCTGAATACTGCGTGTTCTCACTTATAAGTGGGAGCTAAATAATGTATAC   2015
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      ACATGGACACACCAGAGTGTAGAATAATAGACACTGGAGACTTGGAAGAGTTGGAGGGTT   2075
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      GGCAGGGGGCGATGATAATTTACTTAATGGGTACAATGTACATTATTCAGGGGGTGATGG   2135
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      ------------------------------------------------------------   977
NM_000134.4      TTACAGCCCAGACTTTACCACTATGCAATATATCAATGTAACAAAACTATACTTGTACTC   2195
NM_131431.1      ------------------------------------------------------------   645

NM_007980.3      --------------------------------        977
NM_000134.4      CTTAAATTTATATAAAAATACCTTAAATCTA         2226
NM_131431.1      --------------------------------        645
```

Appendix F

Multiple sequence alignment, *lipf*. The first line in each bundle of sequences is for humans (*Homo sapiens, HS*). The second line is for Mice (*Mus musculus, MM*). The third line is for Zebrafish (*Danio rerio, DR*). The asterisks under each bundle of sequences signify DNA bases are identical. Program: Clustal Omega.

Triacylglycerol lipase (lipf)

```
NM_213404.1         ----ACGCTTCACACATT---TTAGGTTGCTGAGAGCTTAACCAAAGTTGTTGGAAAGAT        53
NM_001198828.2      AAATACTAACCAGCCAGAGAAACAGAATCCTAACTATTTCTGAGGAAACTGCAGGTCCAA        60
NM_026334.3         -------AGGCAACCAGAGAAGCAG----AATCCTGTTTCTGAAGGCACTGGCACTACAA        49
                             **  **       **           **                  *

NM_213404.1         GTTATGGCTGGTGTTGTCTTTGGTAATACTTACATCTGGACTGGTTCAGTATGGACAGTC        113
NM_001198828.2      AATGTGGCTGCTTTTAACAATGGCAAGTTTGATATCTGTACTGGGGACTACACATGGTTT        120
NM_026334.3         GATGTG---GCTGCTATTAGTAACAAGTGTGCTATCTGCATTTGGAGGTGCACATGGCCT        106
                      * **   * *  *     *   **   *   ***** * * *

NM_213404.1         GGCTCTGGATGGGACAGGATTGGACCCCGAAGTTAACATGAACATTAGTGAAATTATCAG        173
NM_001198828.2      GTTTGGAAAATTACATCCTGGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTAC        180
NM_026334.3         ATTTGGAAAACTGGGTCCCAAAAACCCTGAAGCAAACATGAATGTTAGTCAGATGATAAC        166
                       *    *              *** ****  *  *****  ***** * ** ** *

NM_213404.1         ACACTGGGGTTACCCTGCTGAAGAATTTGAGGTGGTCACTGAAGACGGATACATCCTGAG        233
NM_001198828.2      TTATTGGGGATACCCAAATGAAGAATATGAAGTTGTGACTGAAGATGGTTATATTCTTGA        240
NM_026334.3         TTACTGGGGATATCCAAGTGAGGAATATGAAGTTGTTACTGAAGATGGCTACATTCTGGG        226
                      * ***** ** **   *** **** *** ** ** ******** ** ** ** **

NM_213404.1         CATCAACCGAATCCCACATGGAGTCAAAAACAAAAACGA---AGAAGTCAAGCCTGTAGT        290
NM_001198828.2      AGTCAATAGAATTCCTTATGGGAAGAAAAATTCAGG-----------------------        276
NM_026334.3         GGTCTATAGAATTCCTTATGGGAAGAAAAATTCTGAGAATATCGGCAAGAGACCTGTGGC        286
                      ** *  **** **  ****    *****

NM_213404.1         GTTTCTCCAACACGGTCTTCTCGCTGCTGGAAGTAACTGGGTGACAAACCTGCCCAACAA        350
NM_001198828.2      ------------------------------------------------------------        276
NM_026334.3         ATATTTGCAGCATGGTTTGATTGCATCAGCCACAAACTGGATTACAAATCTGCCAAACAA        346


NM_213404.1         CAGTCTGGGATTCGTTTTAGCTGATGCTGGATTTGATGTATGGATCGGGAACAGTCGCGG        410
NM_001198828.2      ---------------GAATACAGATGCTGGTTATGATGTGTGGCTGGGCAACAGCAGAGG        321
NM_026334.3         CAGCCTGGCCTTCATTCTAGCAGATGCTGGCTATGATGTGTGGCTGGGGAACAGTCGAGG        406
                                  * ******** * ****** *** * ** *****  * **

NM_213404.1         GAACACCTGGTCCTGCAAACATGTCAGCCTTGACCCAAGACAGAAAGAATACTGGAAGTT        470
NM_001198828.2      AAACACCTGGGCCAGAAGAAACTTGTACTATTCACCAGATTCAGTTGAATTCTGGGCTTT        381
NM_026334.3         GAATACATGGTCCCGGAAAAATGTATACTATTCACCAGACTCAGTTGAATTCTGGGCTTT        466
                     ** ** *** ** * * * *  *   *  *   ***         **** ****   **

NM_213404.1         CAGTCATGATGAAATGGCTAAAAAGGATCTTCCTGCAGTGATAAACTTCATCACAAAGAT        530
NM_001198828.2      CAGCTTTGATGAAATGGCTAAATATGACCTTCCAGCCACAATCGACTTCATTGTAAAGAA        441
NM_026334.3         CAGCTTTGATGAAATGGCTAAATATGACCTTCCAGCCACCATAGACTTCATTGTACAGAA        526
                    ***   *************** * ** ***** **    **  *******   * ***
```

```
NM_213404.1         GACTGGCCAGGAACAAATCTTTTACGTGGGCCATTCTCAGGGAACAACAATAGCATTCAT        590
NM_001198828.2      AACTGGACAGAAGCAGCTACACTATGTTGGCCATTCCCAGGGCACCACCATTGGTTTTAT        501
NM_026334.3         AACTGGACAAGAGAAGATACACTATGTTGGTCACTCTCAGGGCACCACTATCGGTTTTAT        586
                     ***** **  *  *  *    ** ** ** ** ** ***** ** ** ** *  ** **

NM_213404.1         GGCTTTCTCCACAATGCCTGAGCTGGCCAGCAAAATCAAGATGTTCTTTGCACTGGCTCC        650
NM_001198828.2      TGCCTTTTCCACCAATCCCAGCCTGGCTAAAAGAATCAAAACCTTCTATGCTCTAGCTCC        561
NM_026334.3         TGCCTTTTCTACCAATCCTGCTCTGGCTAAAAAAATCAAGAGGTTTTATGCATTAGCTCC        646
                     ** ** ** ** *  **    ***** *  * ****** *  ** * ***  * *****

NM_213404.1         TGTTGCCACAGTGGGGATGACCAAAAGCCCGATGACCAAACTGTCAGTAATTCCAGAGTT        710
NM_001198828.2      TGTTGCCACTGTGAAGTATACAAAAAGCCTTATAAACAAACTTAGATTTGTTCCTCAATC        621
NM_026334.3         AGTTGCTACTGTGAAGTATACAGAAAGTCCCTTTAAAAAGATTTCACTTATTCCTAAGTT        706
                     ***** ** ***  *   **  **** *   * *  **  *   * *  ****  * *

NM_213404.1         TCTGATTTGGGATCTCTTTGGTCAAAAAGACTTTTTTCCTCAGAATGAACTGATCAAGTT        770
NM_001198828.2      CCTCTTCAAGTTTATATTTGGTGACAAAATATTCTACCCACACAACTTCTTTGATCAATT        681
NM_026334.3         TCTTCTCAAGGTGATATTTGGTAACAAAATGTTCATGCCCCACAACTACTTAGATCAATT        766
                     **  *   *    * ****** * ***   **    ** ** **     *     * **

NM_213404.1         TTTCGCCACTGAGTTCTGCAGCAGGAAGCCTTTGAGTGTGTTGTGTGGAAACGTCTTCTT        830
NM_001198828.2      TCTTGCTACTGAAGTGTGCTCCCGTGAGATGCTGAATCTCCTTTGCAGCAATGCCTTATT        741
NM_026334.3         TCTTGGTACGGAAGTGTGCTCACGGGAGCTGCTAGATCTTCTCTGCAGCAACGCTTTATT        826
                    * * *  ** **  * ***    *  **    *   * *  * **  * ** *  ** **

NM_213404.1         TCTTCTCTGCGGTTTTGATGAGAAGAACCTGAACATGAGTAGAACACCTGTGTATACGGC        890
NM_001198828.2      TATAATTTGTGGATTTGACAGTAAGAACTTTAACACGAGTCGCTTGGATGTGTATCTATC        801
NM_026334.3         CATCTTCTGTGGATTTGACAAGAAAAACTTAAATGTGAGTCGCTTTGATGTGTATCTAGG        886
                      *  * ** ** *****    ** *** * **   **** *      *******

NM_213404.1         ACACTGTCCAGCAGGAACCTCCGTCCAGAACATGATTCACTGGGCTCAGGCTGTGAAATC        950
NM_001198828.2      ACATAATCCAGCAGGAACTTCTGTTCAAAACATGTTCCATTGGACCCAGGCTGTTAAGTC        861
NM_026334.3         GCATAATCCAGCAGGAACATCTACTCAAGACCTTTTCCACTGGGCACAGCTTGCTAAATC        946
                     **   ************ **    **  ** *  * ** *** * ***  **  ** **

NM_213404.1         CAGTAAGCTCATGGCATACGATTATGGAAGAGCTGG---AAATATGGCACATTATAATCA        1007
NM_001198828.2      TGGGAAATTCCAAGCTTATGACTGGGGAAGCCCAGTTCAGAATAGGATGCACTATGATCA        921
NM_026334.3         TGGGAAGCTTCAAGCCTATAACTGGGGAAGTCCATTACAGAACATGTTACACTACAATCA        1006
                      * **  *    ** **  * *  *****  *       ** * *   ** **  ****

NM_213404.1         GTCGACCCCTCCACTTTACAACATCCAGGACATGATGGTACCCACAGCTGTGTGGTCTGG        1067
NM_001198828.2      GTCCCAACCTCCCTACTACAATGTGACAGCCATGAATGTACCAATTGCAGTGTGGAACGG        981
NM_026334.3         GAAAACGCCTCCCTACTATGATGTGTCAGCCATGACCGTGCCAATTGCAGTGTGGAACGG        1066
                    *      *****    **  *  *    * *****  ** ** *  ** ******   **
```

```
NM_213404.1         TGGACAGGACACTCTGGCAGATCCCAAGGATGTGGCTTTGCTGCTAACACAGATCCCTAA        1127
NM_001198828.2      TGGCAAGGACCTGTTGGCTGACCCCCAAGATGTTGGCCTTTTGCTTCCAAAACTCCCCAA        1041
NM_026334.3         TGGCCATGACATCCTGGCTGATCCCCAAGATGTCGCAATGCTGCTTCCCAAACTCCCCAA        1126
                    ***  * ***    **** ** *** * ***** *   *  ****  *  *  **** **

NM_213404.1         ATTGGTGTATAAACGCGACATTAAGCACTGGGAGCACCTGGATTTTATCTGGGGGATGGA        1187
NM_001198828.2      TCTTATTTACCACAAGGAGATTCCTTTTTACAATCACTTGGACTTTATCTGGGCAATGGA        1101
NM_026334.3         CCTTCTGTACCATAAGGAGATTCTTCCCTACAATCACCTGGACTTCATCTGGGCGATGGA        1186
                      *  * **  *    ** ***     *   * *** **** ** *******  *****

NM_213404.1         CGCCCCACAAGAGATGTATGAGAAAATGATCGAAATAATGAGGGAGAATGTTTGAAGAGA        1247
NM_001198828.2      TGCCCCTCAAGAAGTTTACAATGACATTGTTTCTATGATATCAGAAGATAAAAAGTAGTT        1161
NM_026334.3         TGCGCCTCAAGAGGTTTACAATGAGATAGTTACCATGATGGCAGAAGACTAACAGAATTA        1246
                     ** ** *****  * **  *  * **  *    ** **    **  *

NM_213404.1         CATACAGTATTAACTTTGTTTAGACTATCAGTCCAAATCTGATTTTTGCGCATGTCCAAG        1307
NM_001198828.2      CTGGATTTAAAGAATTAT--CCGTTTGTTTTTCC-AAAATACTTTATTCTCTCATACA--        1216
NM_026334.3         TCTACATTCATTCCAGAA--CCCTTCCT------------------CCTTTCATACA--        1283
                           *                 *                   *    * **

NM_213404.1         TAAAATTAGATGTGATCAGGTCAAGAAAAAAAAAAGGTGTTTTCCAACTAAGCACAAATCC        1367
NM_001198828.2      -TAGTAT---TTTCATAATGTTTGACATGCAGTGCTTCTTTC------------------        1254
NM_026334.3         -CAGTTT---TAA---------ATAAATGTAATACATGTTAT------------------        1312
                      *   *   *              *   *     * **

NM_213404.1         TCCGATTTGGACTGAGCGAGTGAAAAAAGCATTTGATTGAGCAGCAGGTTTTATCAGCAT        1427
NM_001198828.2      TGTAATTTTGACTT-------TAGAAATATATTGGCATCAACAAA---------------        1292
NM_026334.3         TGTAATTTTGGCTT-------TAGAAATATATTGGCATCAACAAAGTT------------        1353
                    *   **** * **         * ***   *** *  * * **

NM_213404.1         TCTTTGAGCCTTAACTAATTCTTTTCTACCTCTTTGTTTGTTCTTGCTGTAAGCAGATGA        1487
NM_001198828.2      ------------------------------------------------------------        1292
NM_026334.3         ------------------------------------------------------------        1353

NM_213404.1         ACCATTTGTTTGAGGTGCTGGTAACCCAATCAACATTTTACGAATGTAAATTTGATCTTA        1547
NM_001198828.2      ------------------------------------------------------------        1292
NM_026334.3         ------------------------------------------------------------        1353

NM_213404.1         AAAACCAATGGAAGTCTTTCTGTGCCATCATACAGAGACTGTTTAACACTCAAGATCAGG        1607
NM_001198828.2      ------------------------------------------------------------        1292
NM_026334.3         ------------------------------------------------------------        1353
```

```
NM_213404.1         TGTTGTTTTTAATGTCCTCTAAAGTAAATGCACTTGTTTAATGTTTTAACTGTAGCTTTT        1667
NM_001198828.2      ------------------------------------------------------------        1292
NM_026334.3         ------------------------------------------------------------        1353

NM_213404.1         ACTAACCACTACATCTTCAAACTGTAGCATTTCATCATTTATTTTAATGTTTTGGTAGGA        1727
NM_001198828.2      ------------------------------------------------------------        1292
NM_026334.3         ------------------------------------------------------------        1353

NM_213404.1         TTGTTTAAATCATGTGACATATCCCAATAAACACAGAAGACATGAAAAAAAAAAAAAAAA        1787
NM_001198828.2      ------------------------------------------------------------        1292
NM_026334.3         ------------------------------------------------------------        1353

NM_213404.1         AAAAAAAAAAAA        1800
NM_001198828.2      ------------        1292
NM_026334.3         ------------        1353
```

Appendix G

Multiple sequence alignment, *vil1*. The first line in each bundle of sequences is for humans (*Homo sapiens, HS*). The second line is for Mice (*Mus musculus, MM*). The third line is for Zebrafish (*Danio rerio, DR*). The asterisks under each bundle of sequences signify DNA bases are identical. Program: Clustal Omega.

Villin 1 (vil1)

```
NM_200238.1     GAACCAGCAGCACCTTTACCTGGCTGCTTCCTTCTCTATTGCTCCGGTGGCAGGTATTTC   60
NM_007127.3     ----------------------ACAATTCCCTGAGATCTCCCAGGTGGCAGCTGCCTC   36
NM_009509.2     ----------------GCTTGCCACAACTTCCTAAGATCTCCCAGGTGGTGGCTGCCTC   43
                                        **  * * *****  * *   **

NM_200238.1     CCTGAGTCTTAAGCTGGCACCATGCCACAGGAGGTCAAGACTGACGTTCCCAAAGTCCTT   120
NM_007127.3     CCCAAGACAGGCTCACTCACCATGAC---CAAGCTGAGCGCCCAAGTCAAAGGCTCTCTC   93
NM_009509.2     TTCCAGACAGGCTCGTCCACCATGAC---TAAACTGAATGCCCAAGTCAAAGGCTCTCTC   100
                    ** *     *   ******* *     *  * *   *  * **          **

NM_200238.1     AATAAAACCACACCAGGACTTCAGATATGGAGAGTGGAGAACATGGAGCTTGTGCCCTGT   180
NM_007127.3     AACATCACCACCCCGGGGCTGCAGATATGGAGGATCGAGGCCATGCAGATGGTGCCTGTT   153
NM_009509.2     AACATCACCACTCCCGGGATACAGATATGGAGGATCGAGGCTATGCAGATGGTACCTGTT   160
                ** *  ***** ** **  * **********  * ***   *** ** * ** **   *

NM_200238.1     CCATCTAAAACATTTGGACAGTTTTTTGAGGGAGACAGCTATGTAATTCTCTACACTCAC   240
NM_007127.3     CCTTCCAGCACCTTTGGAAGCTTCTTCGATGGTGACTGCTACATCATCCTGGCTATCCAC   213
NM_009509.2     CCTTCCAGCACCTTTGGAAGCTTCTTCGATGGTGACTGCTATGTAGTCCTGGCTATCCAC   220
                ** ** *  ** ******   ** ** ** ** *** ****  *  * **    *  ***

NM_200238.1     AAAACCAGCAACAACTTCTCCTATGACATCCATTATTGGCTGGGAAAGTCCACCTCTCAG   300
NM_007127.3     AAGACAGCCAGCAGCCTGTCCTATGACATCCACTACTGGATTGGCCAGGACTCATCCCTG   273
NM_009509.2     AAGACCAGCAGCACTCTCTCCTATGATATCCACTACTGGATTGGCCAGGACTCGTCCCAG   280
                ** **   ** **   * ******** ***** ** *** * **  **  * * ** * *

NM_200238.1     GATGAGATGGGAGCAGCGGCGATTTACACCACGCAGATGGACGACCATCTAGGGGGTGTT   360
NM_007127.3     GATGAGCAGGGGGCAGCTGCCATCTACACCACACAGATGGATGACTTCCTGAAGGGCCGG   333
NM_009509.2     GATGAGCAGGGGGCAGCTGCCATCTACACCACACAGATGGATGACTACCTGAAGGGCCGG   340
                ******  *** ***** ** ** ******** ******** ***   **   ***

NM_200238.1     GCTGTGCAACATCGTGAGACCCAGGGCCATGAGAGTGCCACCTTTCAAGGATACTTCAAA   420
NM_007127.3     GCTGTGCAGCACCGCGAGGTCCAGGGCAACGAGAGCGAGGCCTTCCGAGGCTACTTCAAG   393
NM_009509.2     GCTGTCCAGCACCGCGAGGTTCAAGGCAACGAGAGCGAGACTTTCCGGAGCTACTTCAAG   400
                ***** ** ** ** ***   ** *** * ***** *   * ** *   * ********

NM_200238.1     CAGGGCATCATATACAAAAAAGGTGGAGTTGCATCTGGGATGAAACAGGTGGAGACGAAC   480
NM_007127.3     CAAGGCCTTGTGATCCGGAAAGGGGGCGTGGCTTCTGGCATGAAGCACGTGGAGACCAAC   453
NM_009509.2     CAAGGCCTTGTGATCCGGAAAGGGGGAGTGGCTTCCGGCATGAAGCACGTAGAAACAAAC   460
                ** *** *  *   *   ***** ** ** ** ** ** ***** ** ** ** ** ***

NM_200238.1     ACATACAACATACGCAGACTTCTGCATGTGAAGGGAAATAAGCATGTGGTTGCAGGAGAG   540
NM_007127.3     TCCTATGACGTCCAGAGGCTGCTGCATGTCAAGGGCAAGAGGAACGTGGTAGCTGGAGAG   513
NM_009509.2     TCCTGTGATGTCCAGCGACTGTTGCACGTCAAGGGCAAGAGGAATGTGCTGGCTGGAGAG   520
                 * *   *  * *   * **  **** ** ***** ** * * * *** * ** ******
```

```
NM_200238.1     GTGGAAATGAGCTGGAACAGCTTCAATAAAGGCGATGTGTTCCTGCTGGACCTGGGGAGT   600
NM_007127.3     GTAGAGATGTCCTGGAAGAGTTTCAACCGAGGGGATGTTTTCCTCCTGGACCTTGGGAAG   573
NM_009509.2     GTGGAAATGTCCTGGAAGAGTTTCAACAGAGGGGATGTCTTCCTGCTGGACCTTGGGAAG   580
                ** ** ***  ****** ** *****   *** ***** ***** ******** ****

NM_200238.1     CTGATCATTCAGTGGAATGGGCCTAAGAGTAACCGAATGGAGAGATTAAGGGGAATGAAC   660
NM_007127.3     CTTATCATCCAGTGGAATGGACCGGAAAGCACCCGTATGGAGAGACTCAGGGGCATGACT   633
NM_009509.2     CTTATTATCCAGTGGAATGGGCCAGAGAGTAACCGCATGGAGAGACTTCGGGGCATGGCC   640
                ** ** ** ********** **  * ** * *** ******** *  **** ***

NM_200238.1     CTAGCCAAGGACATCCGCGATCGAGAGAGAGGAGGCCGAGCGCAAGTGGCCGTAGTGGAA   720
NM_007127.3     CTGGCCAAGGAGATCCGAGACCAGGAGCGGGGAGGGCGCACCTATGTAGGCGTGGTGGAC   693
NM_009509.2     TTGGCCAAAGAGATCCGAGACCAGGAACGGGGTGGACGTACCTACGTAGGTGTGGTGGAC   700
                 * ***** ** ***** ** *  **  * ** ** **  *  * ** *  ** *****

NM_200238.1     GGAGATGATGAGCAATCTTCAGAGGAAGCCATGAAACTGATGAAGCAATCCCTGGGCGAG   780
NM_007127.3     GGAGAGAATGAATTGGCATCCCCGAAGCTGATGGAGGTGATGAACCACGTGCTGGGCA--   751
NM_009509.2     GGGGAGAAGGAAGGGGACTCCCCACAGCTGATGGCAATTATGAACCACGTGCTGGGCC--   758
                ** **  * **       **     *    ***    * ***** **    ******

NM_200238.1     AGGCCGAAAAACATCAAAGAAGCCATCCCAGATGAGATCGTTGATGAGAAACTCAAAACT   840
NM_007127.3     -AGCGCAGGGAGCTGAAGGCGGCCGTGCCCGACACGGTGGTGGAGCCGGCACTCAAGGCT   810
NM_009509.2     -CACGCAAAGAACTGAAGGCTGCTATTTCTGACTCAGTGGTGGAGCCGGCCGCTAAGGCT   817
                   *  *   *  * ** *  **  *  * **     * ** **   *      **  **

NM_200238.1     GCCATTAAGCTCTATCATATTTCAGATGCGCAGGGGAGCCTTGTAGTGCAGGAAGTGGCA   900
NM_007127.3     GCACTCAAACTGTACCATGTGTCTGACTCCGAGGGGAATCTGGTGGTGAGGGAAGTCGCC   870
NM_009509.2     GCACTCAAGCTGTACCATGTGTCTGACTCCGAAGGAAAACTGGTGGTTAGAGAAGTTGCT   877
                **  * ** ** ** *** * ** **  *  * ** *  ** ** **    ***** **

NM_200238.1     GTAAAGCCCCTCACACAAGACCTGTTGAAAACTGAGGACTGCTATCTGCTGGATCAAGGC   960
NM_007127.3     ACACGGCCACTGACACAGGACCTGCTCAGTCACGAGGACTGTTACATCCTGGACCAGGGG   930
NM_009509.2     ACTCGGCCACTCACACAAGACCTGCTCAAGCATGAGGACTGTTACATCCTGGACCAGGGA   937
                     *** ** ***** ****** * *     ******** **  * ***** ** **

NM_200238.1     GGAATAAAGATCTTCATCTGGAAGGGCAAAAAGGCATCCAAAGCTGAAAGGACAGAGTCT   1020
NM_007127.3     GGCCTGAAGATCTACGTGTGGAAAGGGAAGAAAGCCAATGAGCAGGAGAAGAAGGGAGCC   990
NM_009509.2     GGCCTGAAGATCTTTGTGTGGAAGGGAAAAAATGCCAACGCTCAGGAGAGGAGCGGAGCC   997
                **  * *******   * ***** ** ** ** **          ** * **  *   *

NM_200238.1     CTAAAGATGGCAGAAGCCTACGTAAAAGCAAAAGGATATCCTGTTTCCACCTACATAGAG   1080
NM_007127.3     ATGAGCCATGCGCTGAACTTCATCAAAGCCAAGCAGTACCCACCAAGCACACAGGTGGAG   1050
NM_009509.2     ATGAGCCAGGCCTTGAACTTCATCAAAGCCAAGCAGTACCCACCGAGCACGCAGGTTGAG   1057
                 * *     **      ** * * ***** **    ** **      ***  *  * ***
```

```
NM_200238.1      ACGGTGAGCGAGGGAGCAGAATCGTCAGTGTTCAAGCAACTCTTCCAGAAATGGACAGAC   1140
NM_007127.3      GTGCAGAATGATGGGGCTGAGTCGGCCGTCTTTCAGCAGCTCTTCCAGAAGTGGACAGCG   1110
NM_009509.2      GTGCAGAATGACGGGGCCGAGTCCCCCATCTTCCAACAACTCTTCCAGAAGTGGACAGTG   1117
                   *  **  ** ** ** ** **  *  * **  * ** ********** *******

NM_200238.1      AAGGGGCAGACAGTTGGAATGGGCACCACACACAATCCGGGGAAAATTGCGAAAGTTGAG   1200
NM_007127.3      TCCAACCGGACCTCAGGCCTAGGCAAAACCCACACTGTGGGCTCCGTGGCCAAAGTGGAA   1170
NM_009509.2      CCCAACCGGACCTCAGGCCTCGGCAAAACCCACACTGTGGGCTCTGTGGCTAAGGTGGAA   1177
                       * ***    **  * ****  ** **** *  ***     * ** ** ** **

NM_200238.1      CAAGTCAAGTTTGATGCCACCTCTATGCATGCCAGACCTGACATAGCTGCACAGCAGAAA   1260
NM_007127.3      CAGGTGAAGTTCGATGCCACATCCATGCATGTCAAGCCTCAGGTGGCTGCCCAGCAGAAG   1230
NM_009509.2      CAGGTGAAGTTTGATGCTCTGACCATGCATGTACAACCTCAGGTGGCTGCCCAGCAGAAA   1237
                 ** ** ***** *****     * *******     *** *  * ***** ********

NM_200238.1      ATGGTGGATGATGGGTCAGGAGAAGCAGAGGTGTGGAGGATAGAGGACAATGAGCTCGCC   1320
NM_007127.3      ATGGTAGATGATGGGAGTGGGGAAGTGCAGGTGTGGCGCATTGAGAACCTAGAGCTGGTA   1290
NM_009509.2      ATGGTGGATGACGGGAGTGGGGAAGTGCAGGTGTGGCGCATCGAGGACTTGGAGCTGGTG   1297
                 ***** ***** ***   ** ****   ******** * ** *** **   ***** *

NM_200238.1      CCTGTGGACAGGAAGTGGCTCGGTCACTTCTACGGTGGGGACTGTTATCTCATTCTCTAC   1380
NM_007127.3      CCTGTGGATTCCAAGTGGCTAGGCCACTTCTATGGGGGCGACTGCTACCTGCTGCTCTAC   1350
NM_009509.2      CCTGTGGAGTCCAAGTGGCTGGGCCATTTCTACGGTGGTGACTGCTACCTGCTGCTCTAC   1357
                 ********    ******** ** ** ***** ** ** ***** ** **  * ******

NM_200238.1      AAATATGAAGTCAGCAGCAAATTTCACTACATTCTCTACATGTGGCAAGGCCGCCATGCA   1440
NM_007127.3      ACCTACCTCATCGGCGAGAAGCAGCATTACCTGCTCTACGTTTGGCAGGGCAGCCAGGCC   1410
NM_009509.2      ACCTACCTCATAGGGGAAAAGCAACACTACTTGTTATATATCTGGCAGGGCAGCCAGGCC   1417
                 *  **     *  *    **    ** *** *  * **  * ***** *** **** **

NM_200238.1      TCCACAGCAGAGCTGACAGCCTGTGCATATCAGGCTGTGATTCTGGACCAGAAATATAAT   1500
NM_007127.3      AGCCAAGATGAAATTACAGCATCAGCTTATCAAGCCGTCATCCTGGACCAGAAGTACAAT   1470
NM_009509.2      AGCCAGGATGAAATTGCAGCCTCGGCGTATCAAGCCGTCCTGTTGGACCAGAAGTACAAT   1477
                   *   *  **  *  **** *  ** ***** ** **  *  ********** ** ***

NM_200238.1      GGAGAACCGGTGCAGGTCCGCGTTCCTATGGGAAAGGAGCCCATGCATCTAATGGCCATA   1560
NM_007127.3      GGTGAACCAGTCCAGATCCGGGTCCCAATGGGCAAGGAGCCACCTCATCTTATGTCCATC   1530
NM_009509.2      GACGAGCCAGTACAGATCCGGGTCACGATGGGCAAGGAGCCGCCTCACCTCATGTCTATC   1537
                 *  ** ** ** *** **** **  * ***** ********    ** ** *** * **

NM_200238.1      TTCAAGGGCAAAATGATTGTTTATGAGGAAGGAAGTTCCAGGGAAGGATCCTCCCAATCT   1620
NM_007127.3      TTCAAGGGACGCATGGTGGTCTACCAGGGAGGCACCTCCCGAACTAACAACTTGGAGACC   1590
NM_009509.2      TTCAAGGGCCGCATGGTGGTTTATCAGGGAGGCACCTCCCGAAAGAACAACTTGGAGCCT   1597
                 ********    *** * ** **  *** *** *  *** *         **   *  *
```

```
NM_200238.1      CGGCCATCAGTTAGATTGTTTCATGTCCATGGTACGAACGAGTTCAACACACGTGCCACT   1680
NM_007127.3      GGGCCCTCCACACGGCTGTTCCAGGTCCAGGGAACTGGCGCCAACAACACCAAGGCCTTT   1650
NM_009509.2      GTGCCCTCTACGAGGCTATTTCAGGTCCGAGGGACCAATGCTGATAACACCAAGGCTTTT   1657
                   *** **     *  * ** ** ****  ** **    *     *****    **   *

NM_200238.1      GAGGTGCCTCCACGCTCATCTTCTCTGAACTCCAATGATGTGTTTGTGCTCAGCACGGAT   1740
NM_007127.3      GAGGTCCCAGCGCGGGCCAATTTCCTCAATTCCAATGATGTCTTTGTCCTCAAGACCCAG   1710
NM_009509.2      GAGGTGACAGCCCGGGCCACGTCCCTCAACTCCAATGATGTCTTCATACTCAAGACTCCG   1717
                 *****  *  * **  *    *  ** ** ********** **  * ****  **

NM_200238.1      AAGTGCTGCTACCTGTGGTATGGCAAGGGCTGCAGTGGAGATGAGAGAGAGATGGCCAAA   1800
NM_007127.3      TCTTGCTGCTATCTATGGTGTGGGAAGGGTTGTAGCGGGGACGAGCGGGAGATGGCCAAG   1770
NM_009509.2      TCCTGCTGCTACCTGTGGTGTGGGAAGGGCTGCAGTGGGGATGAGAGGGAGATGGCCAAG   1777
                    ******** ** **** *** ***** ** ** ** ** *** * **********

NM_200238.1      TCTCTAGCTGACATTATTTCTGAAAGGGAGAAGCAAGTCATAGCTGAAGGCCAGGAACCA   1860
NM_007127.3      ATGGTTGCTGACACCATCTCCCGGACGGAGAAGCAAGTGGTGGTGGAAGGGCAGGAGCCA   1830
NM_009509.2      ATGGTTGCTGATACCATCTCTCGGACGGAGAAACAAGTGGTAGTAGAGGGGCAGGAGCCA   1837
                     * ***** *  ** **    * ****** *****  * *  ** ** ***** ***

NM_200238.1      GCTGACTTCTGGGTTAATCTAGGTGGAAAGTCCCAGTATGCTAGCAACAAAAGACTTCAG   1920
NM_007127.3      GCCAACTTCTGGATGGCCCTGGGTGGGAAGGCCCCCTATGCCAACACCAAGAGACTACAG   1890
NM_009509.2      GCCAACTTCTGGATGGCTCTGGGCGGGAAGGCGCCCTACGCCAACACCAAGAGGCTGCAG   1897
                 **  ******** *    ** ** ** *** * *  ** ** * ** *** ** ** ***

NM_200238.1      GATGAAAACATCAGTATCACGCCCCGGTTGTTTGAGTGCTCCAACCAGACAGGACGCTTC   1980
NM_007127.3      GAAGAAAACCTGGTCATCACCCCCCGGCTCTTTGAGTGTTCCAACAAGACTGGGCGCTTC   1950
NM_009509.2      GAGGAAAACCAAGTCATCACTCCTCGGCTCTTCGAGTGCTCCAACCAGACCGGACGCTTT   1957
                 ** ******      ***** ** *** * ** ***** ****** **** ** *****

NM_200238.1      ATAGCAACAGAGATCACAAACTTCAACCAGGATGACCTGGATGAGGATGACGTGATGTTG   2040
NM_007127.3      CTGGCCACAGAGATCCCTGACTTCAATCAGGATGACTTGGAAGAGGATGATGTGTTCCTA   2010
NM_009509.2      CTGGCCACAGAGATCTTTGACTTCAATCAGGATGACCTGGAGGAGGAGGATGTGTTCCTA   2017
                  * ** *********    ******* ********* **** ***** ** *** *  *

NM_200238.1      CTGGACATCTGGGATCAGGTGTACCTGTGGATCGGAAAAGGTGCCAATGACACAGAGAAG   2100
NM_007127.3      CTAGATGTCTGGGACCAGGTCTTCTTCTGGATTGGGAAACATGCCAACGAGGAGGAGAAG   2070
NM_009509.2      TTGGATGTCTGGGACCAGGTCTTCTTCTGGATAGGGAAACATGCCAATGAGGAAGAGAAG   2077
                  * **  ******* ***** * * * ***** ** ***  ****** **    ******

NM_200238.1      CGTGAAGCGGTGGTTACAGCACAGGAGTATCTGAAGTCCCACCCTGCTGGCCGTGACCTG   2160
NM_007127.3      AAGGCCGCAGCAACCACTGCACAGGAATACCTCAAGACCCATCCCAGCGGGCGTGACCCT   2130
NM_009509.2      AAGGCTGCAGCTACAACTGTACAAGAATACCTCAAGACCCACCCTGGAAACCGAGACCTT   2137
                    *  ** *     ** * *** ** ** ** *** **** **       ** ****
```

```
NM_200238.1      GATACACCTGTTCTTGTGGTGAAACAAGGTTTTGAACCACCCACCTTTACAGGATGGTTC   2220
NM_007127.3      GAGACCCCCATCATTGTGGTGAAGCAGGGACACGAGCCCCCCACCTTCACAGGCTGGTTC   2190
NM_009509.2      GAGACCCCTATCATCGTGGTGAAGCAGGGACACGAGCCCCCCACCTTCACAGGCTGGTTC   2197
                 ** ** **  *  * ******* ** **    ** ** ******* ***** ******

NM_200238.1      CACGCCTGGGACCCACAAAAGTGGAGTGAAGGGAAATCTTACGACCAGCTTAAAGCAGAG   2280
NM_007127.3      CTGGCTTGGGATCCCTTCAAGTGGAGTAACACCAAATCCTATGAGGACCTGAAGGCGGAG   2250
NM_009509.2      CTGGCTTGGGATCCCTTCAAGTGGAGTAACACCAAATCCTATGATGACCTTAAGGCAGAG   2257
                 *  ** ***** **    ********* *    ***** ** **  * ** ** ** ***

NM_200238.1      CTTGGGGATGCAACTGATGTTATCAAAATTACTGTGGATCTGACCCAACCATCCTCAAAC   2340
NM_007127.3      CTTGGCAACTCTAGGGACTGGAGCCAGATCACTGCTGAGGTCACAAGCCCCAAAGTGGAC   2310
NM_009509.2      CTGGGAAACTCTGGGGACTGGAGCCAGATTGCTGACGAGGTTATGAGCCCGAAAGTGGAC   2317
                 ** **  *  *    **    * * * **  ***  **  * *     **        **

NM_200238.1      CAAACCAACTCAAGTAACTCAACTCAAGGAGGTTCGATGCCCCGTCCTGTAACACAAACC   2400
NM_007127.3      GTGTTCAATGCTAACAGC---------------AACCTCAGTTCTGGGCCTCTGCCCATC   2355
NM_009509.2      GTTTTCACTGCCAATACC---------------AGTCTGAGTTCTGGGCCCCTGCCCACC   2362
                      **   * *  * *                  *      *         *  * *

NM_200238.1      TTTCCTGCAGAGAAGCTGGTGAACGTTCAGACAGAAGACCTGCCTGAGGGAGTCGACCCC   2460
NM_007127.3      TTCCCCCTGGAGCAGCTAGTGAACAAGCCTGTAGAGGAGCTCCCCGAGGGTGTGGACCCC   2415
NM_009509.2      TTCCCCCTGGAGCAGCTGGTAAACAAGTCTGTAGAGGATCTCCCTGAGGGTGTGGACCCC   2422
                 ** **    *** **** ** ***        *** ** ** ** ***** ** ******

NM_200238.1      ACGAGAAAAGAGGACTACCTCTCAGACGATGACTTTGCCCTGGCCATGGGGATATCTCGG   2520
NM_007127.3      AGCAGGAAGGAGGAACACCTGTCCATTGAAGATTTCACTCAGGCCTTTGGGATGACTCCA   2475
NM_009509.2      AGCAGGAAGGAGGAGCACCTGTCCACCGAAGACTTCACTAGGGCCTTGGGCATGACTCCA   2482
                 *  ** ** *****  **** **    ** ** **  *   **** * ** **  ***

NM_200238.1      ATGAATTTCTATGCCATGCCATCCTGGAAGCAACTGAATCTAAAGAAAGAAAAGGGCCTG   2580
NM_007127.3      GCTGCCTTCTCTGCTCTGCCTCGATGGAAGCAACAAAACCTCAAGAAAGAAAAAGGACTA   2535
NM_009509.2      GCTGCCTTCTCTGCCCTGCCTCGATGGAAGCAACAAAACATCAAGAAAGAAAAAGGACTG   2542
                       **** ***  ****    **********  **  * ********** ** **

NM_200238.1      TTTTAGGCACAAATGTGCACAAATACATCCCCGTGCCACAAAACAGACTTGCT------C   2634
NM_007127.3      TTTTGAGAAGAGTAGCTGTGGTTGTAAAGCAGTACCCTA---CCCTGATTGTAGGGTCTC   2592
NM_009509.2      TTTTGAGAATTGAAGCTCTCTGGCTGTCCAGCAGCCCCTAC-CCTGCCTTCAAGGGCTTT   2601
                 ****  * *     *                    **      *    **

NM_200238.1      TCATGTTAAACTGTATGAATTCTAGTAGAGGTGCCATGCAAAAACAACTAAATTTCCTTT   2694
NM_007127.3      ATTTTCTCACCGATATTAGTCCTACACCAATTG-AAGTGAAATTTTGCAGATGTGCCTAT   2651
NM_009509.2      ---GTGCCGCCATTACTGGTTTTAGTCCTGTGGCAGATGAAAATGTCCAATTGTACCTGT   2658
                          *  **    *  **        *      ***     *     * *** *
```

```
NM_200238.1      TAGAGTTTTTAAGATGTTTA-CATATAATTCCTTTATTTAAACATTTCTCTTTTTTTGCA   2753
NM_007127.3      GAGCACAA---ACTTCTGTGGCAAATGCCAGTTTTGTTTAATAATGTACCTATTCCTTCA   2708
NM_009509.2      GAGCCACA---GT--GTGACAATTCCTTTTGTTTATAATAGTAATTTGCCCATTCCTTCA   2713
                  **            *              **    **   ** *  *  **  * **

NM_200238.1      CATGCAAACTGA-------TTTATC-----TGAAATTGACAT-ATTTAAAGTCAAC----   2796
NM_007127.3      GAAAGATGATACCCCAAAAGGAGCCTATGGTCCTCATTTCAACTTCTAAGGTCGCTAGAT   2768
NM_009509.2      GACGCATGCCACAGACCCATGGAAATCTTGTAGAGTTTTCTT-TTCTTAGATGGACAGCT   2772
                  *   *                        *     *  *    * * *  *

NM_200238.1      ---------------ATTATAAAGCAGAATTAATTACCTTA-GAGTATAACTTTTAAACA   2840
NM_007127.3      TGTTTCTATCCTGAGGTATTGCATCAATTTTAATACTCCTATAGTTTTCTCTTCTTAGAA   2828
NM_009509.2      AAGTACT-------------CCAGGAGACATTAGCGTCTGG-GGGTTTCTCT--------   2810
                                    *  *    * *    *       * *  **

NM_200238.1      GGAAAATGGATGATATTGGGAA-AATGAAGTCTGCAGATCTATCTTTTTTTTTTTACACA   2899
NM_007127.3      GAGCACAAACACTCCATGGAACATTAGAGTTCTGAGGCACTACCCTAGCTTGTCCTCTAT   2888
NM_009509.2      -GGCACCGTCACTCA---------------CTCAGGATCTTATCCTGATCTTACC----   2849
                      *                       **   *  **        *  *

NM_200238.1      -------GATTTGTATTGAT-GTAGCATGGCGA-----ACTTCCAGTTTAATTTCCTTTC   2946
NM_007127.3      CATGACTCATTTTTATCTATGGCAGGTAGGCTGAAGCACTTTGCAGGTTTACATCTTCCC   2948
NM_009509.2      -------CTCCTCACACTCAAAAGGGAGGGCTAAGGCCAAAGCTGGGCTTACAGCTCTAA   2902
                           *           *   ***             *  * *   *

NM_200238.1      CTTTTTACTTGCTTTTCCTTAAAGCATCTCAAGAGGAAACATTTAAGCACTGACAAAAAG   3006
NM_007127.3      CAGAGTAACAGCTTTTCCTTTTC------ACATA------------TACTTTCCTTACT   2989
NM_009509.2      C----CCAGAGCCTTTGCAAAGC-----TCCACA------------GACTCCTCAAATG   2940
                 *        ** *** *            * *             ***     *

NM_200238.1      CCACGACTCACTGAAGCGGCTTAAGCACTACAATAAAACTTAATATCTTTGAACTAAGAT   3066
NM_007127.3      GCCTTACTCAGTGGGTAAGTTAAAGGGCTGAAGG--------------------AGAGT   3028
NM_009509.2      ACAACACCAGGAACATGGGTTTGCTACGTGAAGT--------------------CCAAT   2979
                  *   **           * *       *  *                        *

NM_200238.1      GCAAGTTGTATTTGAGATTAATGATTTATAGTGAGTATTAATTATAAAAGACACCCGGAG   3126
NM_007127.3      TGAATGGTCCACAAGACTACCCTCTTAAGAGGTTTCACAAATTCCAAA-CAGTACCAGTG   3087
NM_009509.2      CAGAAGCCAATAGGTGATTTTCTCTTAAAACTGGTTATCC-------A-GTGTCCCCAGG   3031
                    *            *      ** * *      *          *      **   *

NM_200238.1      ---------------------------------ACTACTGCACACTGTCTAATGTGTAAT   3153
NM_007127.3      AGAGCAGCACTTCCACTGGGGCTAGGCTTGAGACCTAAAGGCAAGTATGAAATGCATATG   3147
NM_009509.2      A-----ACATGTCCCTTTAAACAAATAAATCAAACTA-------ATATGAGGTTAATAAA   3079
                                                  ***       * *    *   **
```

```
NM_200238.1     TTCCATTGGTATTAATCATTTACTGTGTACAAAAGGCAGACCACCTGTCTAATGGAA---   3210
NM_007127.3     CTACTTCACTCCCTCTCCCAACCCTTAATAATGAGGCAAAGCAAGAGCCTAGTGAAGGCC   3207
NM_009509.2     GG-CTTTAATGTCTC--TCACACATTAAAACAAAACAAAA-C-----------------   3117
                   * *   *           *  *       *   * * *

NM_200238.1     --AAGTCA----ATAAA-----------CATGAAAAGACATTATGATGGAAAAAAAAAA   3252
NM_007127.3     AATGCTAGGTTTACAAACTTACCCAGAAGCCTCTGCAAAGCTTCACAGGCTCCTCAGATG   3267
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     AAAAAAAAAAAAAAAAAAAA--A-------------------------------------   3273
NM_007127.3     AAAATAACAGGAATCAATGGGGACTACGGCCAGACACTGGTTTGCCATTCTGTTCCTTTT   3327
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     AAGAAGTAACAGTGCTGCAAGGAAGTCCATGTCAGAAAGCCAACAGAAGGTGATTTCCAC   3387
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     AACTTTGAACAGGTTGTTACAAGTATCAGCAAGAATGTGTCCTTTTCAGAAATAACAGTC   3447
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     AAATCAAAGAAGGTTAATAAAGGCTTTAATTTCATACACACAAAAAAACTCTATGCATAA   3507
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     TTTAAAAAGGAAACAAAAACAAAGAAAAACCGTAAAGGATACAGAGGAACAGTTCTGCTA   3567
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     AAACACAGATAAAAGTGCCGCTCCATACAAAACATAAAGAATCAGAATCAAAAGTCACTC   3627
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     TGAACATAAAGAAAAAAAAATCATCTCACAAATAATGTGGCCACAGCTGCCAGAAAACCTG   3687
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     GTAGTGGCTCAATTAGGCAAAGTGTAGGAATCTCATTTTTGTTTTTCTCTCCTTAAGTTT   3747
NM_009509.2     ------------------------------------------------------------   3117
```

```
NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     AAAGAAACAACAATGACAATAGGCCAGAGAAGTTAGGGAGGGAAAGAAAAGCTCAAAGGG  3807
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     AGGGAAACCTGGGGACAAGAGGTGTGCACACCCACATGTGGTCTCACTCTTCACACAGGC  3867
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     CCACTATTTTTGAAGTAGACCAGTTTAGTTGACTGTTCTTCTTTGTTCTGGCATCTGACT  3927
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     GGACCAACCTGGAACCTGGTCCAGACCCTCACCCACTCTATTCTTATGCCAATGGACATA  3987
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     CCTATACTTTGAACCTCTGTACTTTTAAGAAAAGTCCAATGTTACAAAATCAAATGCTTA  4047
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     TATTCAGACTGGCACACTTTTTAAATAAAAACTCCATACACCTCAGACATATAGCACACA  4107
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     TGGAGACAACTTACTAATTGTGTGTAAGTATGATACAATGAATGAGACTGCCTGAAGTCT  4167
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     AGTAATCAAAGCATGCCATAAGGTGAATGATTGTGGTTAAACACAGCAAAATAATTGTCA  4227
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     CAAAACTTTCAAGGCCTAACAAATTAGAATTTTCCAATAAAAAATATATATTTTTTCAGA  4287
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     TGTTAATAAGACATATCAGTAGAGACAAAATTAGGATTTTGAAGTAATGCAATAAAAAGA  4347
NM_009509.2     ------------------------------------------------------------  3117
```

```
NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     TGTTGGAGGGCAGAAGTCTATTTAGTTTTTGTATACACTTGCAAGAGTGCATTACTCAGT  4407
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     ATAAAGCAAAATGGGGAGGAAAAAGACATCCATCCATTTTATTGGAACACTTTTATGTGA  4467
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     CTTGAATCTGGTGTTAGGTTGTTGATTTTTCTAAAAATCTCCTATATATACAAAATCCAT  4527
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     ATGTACTTGGAGATCCAGCTGTTGCCCCCTGTTTAAAACAAAAGACCACCTCGGGGGGTC  4587
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     AATTAAATTAAAAAGGCCCTCCAACCACCCTAAATGGGATAACTAAGAGTATCTACTGCA  4647
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     GTCATTTCAGAGGACAGAGAAGGAAAATATTTTAATTTGCTTTAATATAACCTCTTTTCA  4707
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     GTAGATCACAAATGAGTTTACAAACTACTTTTTTTTCTCTTTAATTTAGGTGTTTGCAGA  4767
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     TAATTTTCATTATATCCGTAGCTGTATGTGTGTATAGTTACATAATGGTAACTACACACG  4827
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     ATACAGAAGAATCAGTAAATTCATGGATTATTTTGCTGAGGTTTTAAATTTTAAAGCCTC  4887
NM_009509.2     ------------------------------------------------------------  3117

NM_200238.1     ------------------------------------------------------------  3273
NM_007127.3     TGTTTCAGAATTTTATACTTGATCAAGGAGAAAAATAAATGTGTAGTCTAACATTTGCTT  4947
NM_009509.2     ------------------------------------------------------------  3117
```

```
NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     TCTGGAGTTAATTAACTGATCTGTAAGAACCACTGCATATGTCTTAAAATGTAAACATAT   5007
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     TTACATTTGTTGTATTTGTTATTGAGCCTTTAAGGTTAGGCCCAGAATGAACAGACCATA   5067
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     GCAAGTAAACAAATAATTTTTAGAATCAAAGTATTAATAGAAGACCAGTTCATGGATTTG   5127
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     CTTATTCTATCCTGCTGAGACAAAACTCATGAGTGTGCACACACATGTGAATATATCCCT   5187
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     ACGAAACAGTCTATCTTCTCATAGGCTTAAATTATAGTCATGGCTATTAAAGAAATTAAC   5247
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     AGCATCCAGCCACATGCAACTTTCCAACCTTCAGTACTATTAGGTGATTAAAATCAACAA   5307
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     ATATGAAGTTTAGTTCATTTTTCCCTTAAAATTCCAACAAAGATCAAAAGGTTGTATCTA   5367
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     GAACTAATTGCCATCAAGTTCCAATTCAATGTCATTTAAAGTAATGTAACCACACATTTG   5427
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     GTATTTTCAATAGGAAGGTTAATTAGGTATTGTGCAAACTGCCTGCAACGGAAGCACTCA   5487
NM_009509.2     ------------------------------------------------------------   3117


NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     GTCCTTATCTAACTTGTCCCTTCCTGGCCCCAACATGCTAACTGCCCCATCCCCAATTCT   5547
NM_009509.2     ------------------------------------------------------------   3117
```

```
NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      GGGCAAACTGATCACAATGTGCAAAAAATAATATATTATCTATTTATAATTTTAAAATAT   5607
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      ATACAGCCAGCTCAAAAACAACAACAAACCCCATCAACATAGTCCAGCTGAAATCTCCAC   5667
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      TGGTAGTCAAAGAAGTAGATTAAAGGAGTAAAGGAAGGAGAAGGCTGATAGGGCCACAAG   5727
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      AATGGACAGACATCAAGTAAAATTTGAGTCCCAAACATGCAAGTACATGAGCAGTGAGAT   5787
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      AACTTATATATTCCAATAACCTATTTTCAACAACTCCTGCCCAAGACATGAAGATCAAAT   5847
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      CAGGTTTCTGAGGTAAGTGTACTTCTAAACCATACACACATGAAAGCTATGAAAGCAATT   5907
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      CAAATGAGGCTGCTTCATGAGGCAATCTAGACTTATGGCATTATTTCTACTTTTCTCCAT   5967
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      GTTTTAATTGCAGCCTGCACTTTAAATCTTTCCCAATAATTTTTAACAGTGCCTCTCAAA   6027
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      TGCAAAGACACGTAAAAGAATTAATAACTAGCCAAAGAATTTTATCACCGCTTCACTCAT   6087
NM_009509.2      ------------------------------------------------------------   3117

NM_200238.1      ------------------------------------------------------------   3273
NM_007127.3      TTATAAGAAAACCAATTATTTCCAAGCAAAATCAAACCAAACCAAAGAGGCTCCTGGTAG   6147
NM_009509.2      ------------------------------------------------------------   3117
```

```
NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     AAATGAAACAAAACTTTAAAGCTAGTTTTAAAACAAATATTTTCCTCTGCTCTAAACTAC   6207
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     TCTGGCGTTTTCTACCACTCTACCATTTTGGAACATTCATTACAATAAGGTATATAGGTA   6267
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     GATGGTAGGAGGCAAAGCATTTATCAGTAGTTGAGCAAAACTGCTGAGGCCATTTATAAT   6327
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     TCAAAGAATGAAAACTTAGAATAGTTTACTACATTAGAATACATCCAAGTTCCAAGAGTA   6387
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ------------------------------------------------------------   3273
NM_007127.3     GGACTGGAGCTCTCTTAAGGCAATCTTTCAGAAAAGATGTTGCATCTTCTGTGATGATGG   6447
NM_009509.2     ------------------------------------------------------------   3117

NM_200238.1     ----------------------------------------------------  3273
NM_007127.3     TTTGTGTTTTTTTTGTTTTTGTTTTCGTTTTTTTTAATGAGATGGAGTCTCAC  6500
NM_009509.2     ----------------------------------------------------  3117
```

Vita

Victor Jean Nasr was born in Kobayat, Lebanon, to Jean and Wafaa Nasr. He graduated from Freedom High School in June 2010. The following autumn, he entered Valencia Community College to study Biology. In the fall of 2012, he entered University of Central Florida to study Biology, and in June 2014 he was awarded the Bachelor of Science in Biology. In the fall of 2017, he was accepted into Appalachian State University's Master of Science degree program. The M.S. degree was awarded in May 2022.