In Silico Study of the Transcriptional Factor Binding Site for Rhodopsin in Zebrafish (*Danio rerio*)

By

Nurul Amani Binti Ab Ghani

A report submitted in fulfillment of the requirements for the degree of Bachelor of Applied Science (Animal Husbandry Science) with Honours

Faculty of Agro Based Industry

UNIVERSITI MALAYSIA KELANTAN

2018

DECLARATION

I hereby declare that the work embodied in this report is the result of the original research and has not been submitted for a higher degree to any universities or institutions.

_____

Student

Name: Nurul Amani Binti Ab. Ghani

Date:

I certify that the report of this final year project entitled "*In Silico* Study of the Transcriptional Factor Binding Site for Rhodopsin in Zebrafish (*Danio rerio*)" by Nurul Amani Binti Ab Ghani, matric number F14A0292

has been examined and all the correction recommended by examiners have been done for the degree of Bachelor of Applied Science (Agriculture Technology) with Honours, Faculty of Agro-Based Industry, Universiti Malaysia Kelantan.

Approved by:

_____

Supervisor

Name: Dr. Hasnita Binti Che Harun

Date:

ACKNOWLEDGEMENT

First and foremost, I would like to express my gratitude to Almighty Allah S.W.T. for enabling me to complete my research within the time given. Moreover, I would like to thank Allah S.W.T. for giving me strength to overcome all obstacles in the completion of my research.

Special thanks to my research supervisors, Dr. Hasnita binti Che Harun and formerly Dr. Tan Sze Huey. It was a really great work and experience to be one of Dr. Tan Sze Huey's mentees. Without their assistance and dedicated involvement in every step throughout the process, this research would have never been accomplished.

I would like to express my deep gratitude to my research partner Jannatun Naimah binti Amir. She has been a good research companion who is always encouraging me to give out the best during conducting this research. She also helps me out when I was in difficulty while conducting this research. Without her kindness and encouragement this research would never have progressed as smoothly as it did.

With my deepest gratitude, none of this could have happened without my family, my parents and siblings, for giving me moral and financial support to complete my research. They gave me strength to complete this research. Every time I was ready to quit, my family did not let me and I am forever grateful. This thesis stands as a testament to your unconditional love and encouragement.

Lastly, I would like to show my gratitude to all my friends who are always by my side in completing this research. Last but not least, thanks to all people who involved in giving many ideas to finish this research successfully.

In Silico Study of the Transcriptional Factor Binding Site for Rhodopsin in Zebrafish (_Danio rerio_)

ABSTRACT

In silico study is a computational biological experiment simulation by means creating an artificial system that exhibit biological behaviour of the living system. *In silico* study had been used as one of the series of biological experiments, before conducting *in vivo* experiment. However, *in silico* have not much been used in the field area of researches in aquaculture species genetics. An important molecular protein in aqua genetic is rhodopsin. Rhodopsin is a photopigment presence in both animal and human. In zebrafish (*Danio rerio*) the presence rhodopsin is essential for its ability in vision. However, in some cases such as stress due to climate change for example severe hot season will cause bleaching of light exposure entering the retina. High exposure to light will cause destruction of rhodopsin. It is essential for the rhodopsin to be regulated in order to maintain or prevent it from lose its ability in vision and disrupt other biological processes. Thus, transcription factor binding site are essential to be studied to understand the regulation of rhodopsin. Zebrafish was used as model organism due to its established complete genome. This study was conducted by obtaining full genome sequence of rhodopsin in zebrafish at National Center for Biotechnology Information (NCBI) and using MatInspector software tool to obtain the suggested transcription factor binding site. From this study several transcription factor binding sites that was suggested from MatInspector was selected and discussed in further its regulation and its relation to the rhodopsin protein.

Keywords: *In silico*, rhodopsin, zebrafish (*Danio rerio)*, transcription factor binding site, protein.

Kajian *In Silico* Terhadap Tapak Pengikat Faktor Transkripsi untuk Rhodopsin di dalam Zebrafish (*Danio rerio*)

ABSTRAK

Kajian *in silico* adalah simulasi berkomputer eksperimen biologi yang mana sistem tiruan yang menggambarkan kelakuan sistem hidup biologi. Kajian *in silico* telah digunakan sebagai salah satu daripada siri eksperimen biologi sebelum melakukan eksperimen *in vivo*. Walau bagaimanapun, *in silico* tidak banyak digunakan dalam bidang penyelidikan genetik spesis akuakultur. Antara protin molekul yang penting dalam genetik akuakultur ialah rhodopsin. Rhodopsin adalah fotopigmen yang wujud di dalam genetik haiwan dan manusia. Di dalam zebrafish (*Danio rerio*), kewujudan rhodopsin adalah penting untuk keupayaannya dalam penglihatan. Walau bagaimanapun, dalam beberapa kes seperti tekanan akibat perubahan iklim misalnya musim panas yang teruk akan menyebabkan pendedahan cahaya yang tinggi memasuki retina. Pendedahan cahaya yang tinggi akan menyebabkan kemusnahan rhodopsin. Ia adalah penting bagi rhodopsin untuk dikawal selia bertujuan untuk mengekalkan atau menghalangnya daripada kehilangan keupayaannya dalam penglihatan dan mengganggu proses biologi yang lain. Oleh itu, tapak pengikat faktor transkripsi adalah penting untuk dikaji bagi memahami pengawalan rhodopsin. Zebrafish digunakan sebagai organisma model kerana genom lengkapnya telah ditetapkan. Kajian ini dijalankan dengan mendapatkan genom rhodopsin zebrafish yang lengkap di National Center of Biotechnology Information (NCBI) dan menggunakan alat perisian MatInspector untuk mendapatkan tapak pengikat faktor transkripsi yang dicadangkan. Melalui kajian ini, beberapa tapak pengikat faktor transkripsi yang dicadangkan dari MatInspector telah dipilih dan dibincangkan dengan lebih mendalam mengenai pengawalan dan hubungannya dengan protin rhodopsin.

Kata kunci: *In silico*, rhodopsin, zebrafish (*Danio rerio*), tapak pengikat faktor transkripsi, protin.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABREVIATION AND SYMBOLS

Ascl1a	Achaeta-scute family BHLH transcription factor 1a
Crx	Cone-rod homeobox
HMX1	Homeobox 1
Lin28a	Protein lin-28 homolog A
NCBI	National Center for Biotechnology Information
Otd	Orthodenticle
rho	Rhodopsin
Srf	Serum response factor
Stat	Signal Transducer and Activator of Transcription
Stat3	Signal transducer and activator of transcription 3
TFBS	Transcription Factor Binding Site

CHAPTER 1

INTRODUCTION

1.1 Research Background

In the late 1940's, the first computational biological experiment was carried out by John von Neumann who was a mathematician and computer architecture (Fernando, Goldstein & Husbands, 2015; New World Encyclopedia contributors, 2016). This computational biological experiment has been the astonishing solutions to broad complications regarding to the problems when conducting biological experiments which are involved with the use of animal as model organism. This computational biological experiment is called as *in silico* study, which work as alternative to the use of animal as model organism in various biological experiment (Sana Banu et al., 2015). Alongside with the tremendous growth of bioactivity databases, the computational biological experiment has been a powerful virtual screening tool to predict protein targets and for the development of new drug delivery systems (Geetha, Sivaram & Jayakumar, 2016; Koutsoukas et al., 2011). However, there are limited findings on *in silico* transcription factor binding site of molecular study (Mohd Ramli et al., 2012).

In aquaculture industry, the growth and adaptation to environmental changes of aquaculture species are very important. This is because nowadays, the climate changes all over the world are unpredictable. In order to maintain the growth and develop a stable

adaptation to the sudden climate changes, regulations of rhodopsin is one of the ways. This is because rhodopsin able to regulates physiological, behavioural and biochemical activities in vertebrates through photoreceptors. Rhodopsin is a very interesting gene to be discussed and explore. Generally rhodopsin is a photopigment or light absorbing molecule that responsible in vision of fish (Magnoli et al., 2012). The level of rhodopsin regulations need to be in precise level as transgenic mice which there is specific level of rhodopsin regulation, if not the retinal will undergo degeneration (Sohocki et al., 2008). Retinal degeneration will lead to loss of visual function and disrupt other biochemical processes in the vertebrate body (Tan et al., 2001; Olsson et al., 1992). There are limited findings on *in silico* transcription factor binding site (Mohd Ramli et al., 2012) and the transcription factor binding site for a specific protein in zebrafish. Rhodopsin is selected due to its important role in development, maturity and functioning of the sensorial organ. Thus, this project will provide a set of data of transcription factor binding site for rhodopsin in zebrafish. This study also promotes further research on *in vivo* experiment by using the results of this *in silico* study.

1.2 Problem Statement

In silico study is a powerful tool that requires minimum time and expenses to study gene regulation. However, its application is mainly restricted to human disease study. *In silico* approach should be applied on other animal studies in order to reduce the usage of experimental animal.

1.3 Objectives

The objective of the present study was to identify the transcription factor binding sites of rhodopsin in zebrafish.

1.4 Scope of Study

A set of transcription factor binding site for rhodopsin protein data in zebrafish will be determined by conducting an *in silico* study.

CHAPTER 2

LITERATURE REVIEW

2.1 *In silico* study

In silico study is a biological experiment that carried out entirely in computer. It has become solutions to broad complications that have been through in previous biological experiment. Previous biological experiment had gone through several complications in obtaining information about living system, measuring systems with arbitrary precision and there are various measures that need to stress for before, during and after experiment practices (Fernando et al., 2015).

Since the first *in silico* study, many studies have been conducted using *in silico* method. One of the study is "*In silico* Models of Acute Inflammation in Animals" by (Vodovotz et al., 2006). Their study aimed to compare the use of animal models studies and mathematical model (Vodovotz et al., 2006). The experiment was conducted by constructing a mathematical model that resemble to the global tissue dysfunction and the action of cytokines and cells of the critical inflammatory response of the animal models. The comparison was made in other study which rats were affected by surgical trauma and were put under stress with extraction of various inocula of *Escherichia coli*. In this study, the researchers proposed that the mathematical model could be translated to human inflammation which bring several insights dynamics of acute inflammation or

trauma can be tested *in vivo* study, thus minimise the use of animal model (Vodovotz et al., 2006).

There have been vast previously experimental study of severe inflammation, using different animal models (Bier & McGinnis, 2008; Garrido, Figueiredo et al., 2004). However, the use of animal as model organism in laboratory experiment brings out to ethical value issues (Guither, 1998). There is one phrase which is, "humane endpoint" that is always used when using animals as model organism in experimental laboratory work. Humane endpoint refers to halting an experiment at any point in which discontinues the suffering of animal model. However, determining a humane endpoint could frustrate researchers on his study due to difficulty in deciding the animal suffering point and experiment end point (Nemzek et al., 2004).

2.2 Zebrafish (*Danio rerio*)

Zebrafish or scientifically known as *danio rerio* is a tropical freshwater fish which a very practical model organism to investigate the gene function and vertebrate development. There are various benefits of using zebrafish as a model organism which is the genome of zebrafish has been completely sequenced through the *Danio rerio* Sequencing Project, the upgrowth of all dominant organs is notably fast, the stages of zebrafish growth are recorded in hours alternately with mouse which are recorded in days, a sexually mature female zebrafish able to lay hundreds of eggs in one breeding chance which is permitted for extra specimens to be studied. However, there are several

disadvantages of using zebrafish are zebrafish has late sexual maturity which is 3-4 months instead of 2 months for mouse, zebrafish is a very sensitive animals which require a genuine house care and require well-controlled environment (Uchida, 2012).

Besides that, zebrafish is being used as model organism due to their unique combination of genetic tractability and morphological characteristics. In addition, zebrafish has diverse testament that proved the key role of this model organism in the research of circadian clock which is associated to the extra retinal photoreceptors-melatonin production (Magnoli et al., 2012). Precisely, it is the most convenient model organism as it not requires large area and provide ample sample for research. Most importantly, zebrafish has full sequenced genome which allows scientists to create mutations up to 14000 genes in purpose to study their functions.

Other than that, adult zebrafish has the potential to reform any neuronal cell type in a successful way which is lost during retinal damage (Nelson et al., 2012). Furthermore, in studying vision of an organism, fish is the best model organism to investigate the evolution of vertebrate colour vision. This is because fish have the most advanced colour vision due to diverse photic environments in water (Tsujimura et al., 2010).

According to Tsujimura et al. (2010), vertebrates have vast groups of cone visual cells in the retina. Every group has different absorption spectrum and contains visual opsins which are mainly come from various repertoires of colour sensors thus enabling

them to adapt the diverse photic environment in water. The pineal gland of zebrafish plays important endocrine and neuronal roles which include the regulation of the zebrafish circadian clock through the melatonin secretion (Laurà et al., 2012).

2.3 Rhodopsin

Light is one of the most essential signals for life in many biological processes and mechanisms. For instance, photosynthesis may not able to carry out and lead to the death of plants and subsequently will lead to death of other organisms. The responsible protein that involve in the capturing light energy and information is rhodopsin, the photoreceptor protein presence in most living organisms (Kato, Inoue, Kandori, & Nureki, 2016). Many not know that rhodopsin is a very interesting photoreceptor protein to be discussed for due to its ion-translocating rhodopsins which involves both pumps and channels mechanism and its important role as capturing the light energy and information. Thus, making it the most studied photoreceptor protein recently (Kato et al., 2016). A study conducted by Nagloo, Hart and Collin (2017), repeated outcomes has prove that one of the crucial factors that contribute to increase in survival and growth of aquaculture species is visual environment which involves the visual system and visual ecology of the species.

However, the disability or degeneration of photoreceptor proteins will cause retinal neurodegenerative diseases that subsequently results in cone cell blindness and death (Hao et al., 2012). Continuous exposure to the intense light will causing the rod and cone cell death according to established protocols (Nelson et al., 2012).

In the past centuries, the vertebrate has undergone evolution of some of their cells and molecular biology. As consequence to that, gene duplication has arisen due to this vertebrate evolution resulting in production of several opsin families. Opsin which is one of the important visual pigments which mainly divided into two large classes which are; non-visual opsin and visual opsin. Visual opsin comprises rhodopsin (rh1) and four cone opsin group; rhodopsin-like (rh2), short-wavelength sensitive 1(sws1), short-wavelength sensitive 2 (sws2) and long-wavelength sensitive (lws). Non-visual opsin includes pinopsin, melanopsin and exorhodopsin. Visual opsin can be found at photoreceptor cells of the retina and its main role is to initiate the visual transduction cascade (Menon, Han, & Sakmar, 2001; Tsujimura et al., 2010). In this study, rhodopsin of opsin family will be focused in finding its transcription binding site.

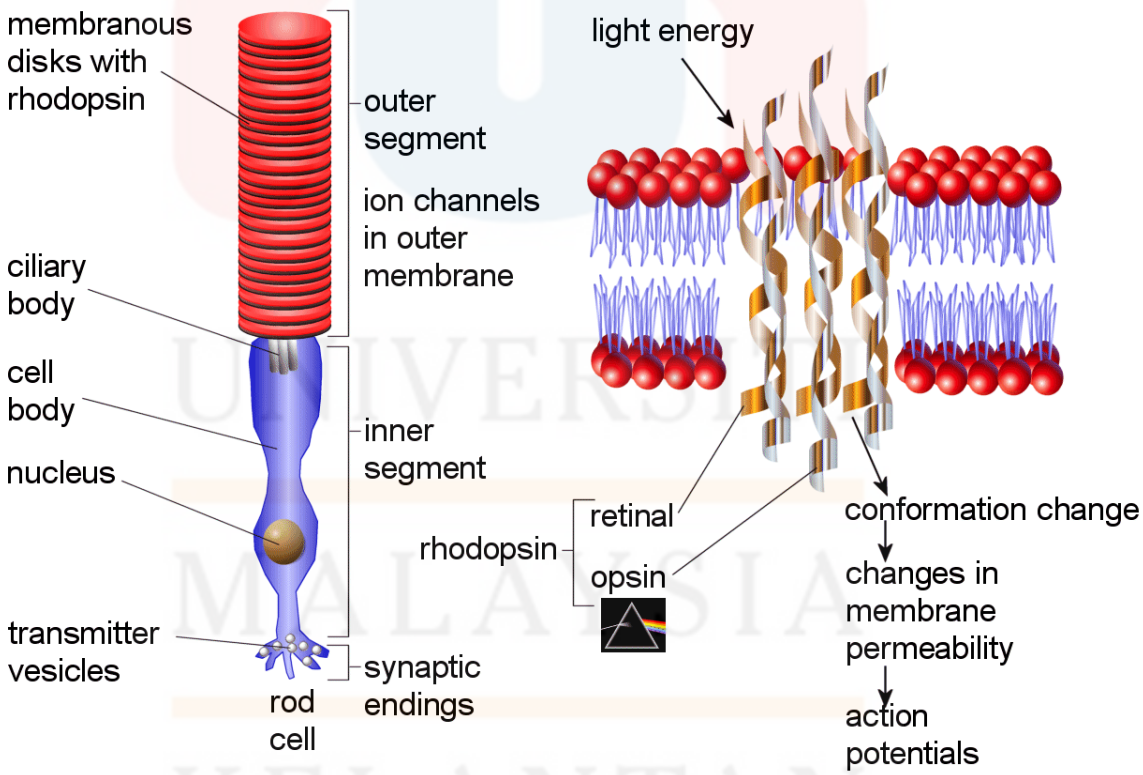


Figure 1.1: The rhodopsin structure. Picture was generated from (Biology, 2010).

Rhodopsin is a photopigment or light-absorbing molecule that related to the vision of fish (Magnoli et al., 2012; “rho - Rhodopsin Danio Rerio,” n.d.). Rhodopsins can be divided into two classes which are microbial rhodopsins and animal rhodopsins (Gehring, 2014; Kato et al., 2016). Microbial rhodopsins consists of heterogenous group which are archeal light-driven ion pumps, sensory rhodopsins, halorhodopsins and bacterial rhodopsins from uncultured marine bacteria (fungal opsins, rhodopsins from cynobacteria and dinoflagelates, and channelopsins from green algae. Animal rhodopsins consists of G-protein-coupled receptors (GPCRs) (Gehring, 2014). A reviews, stated that rhodopsin can be found in the outer segment of pinealocytes (Bailey & Cassone, 2004; Laurà et al., 2012; Mano, Kojima, & Fukada, 1999; Okano et al., 2000; Vigh, Jotio, & Szkl, 1998). The pinealocytes are true cone like photoreceptor which has outer segment, inner segment and a basal pole containing the synaptic pedicles, arranged around a central lumen of the pineal gland (Allwardt & Dowling, 2001; Falco, 1999; Laurà et al., 2012). Specifically, the outer segment of pinealocytes is build up by the lamellar stereocylium membranes which arise from a typical linking sterocylium characterized by 9+2 microtubules organization (Magnoli et al., 2012). According to Schihida and Matsuyama, rhodopsin is a G-protein coupled receptors (GPCRs) which contains apoprotein, opsin and chromophore (Magnoli et al., 2012).

2.3.1 Relation of rhodopsin and melatonin

When light hits rhodopsin, it will excite and play its function as photon absorption, trigger the cytoplasmic signalling cascade implying G proteins and their effectors give

rise in a decrease of intracellular cGMP concentration proceeding to the hyperpolarization of the photoreceptor cell (Terakita, 2005).

Meanwhile, melatonin is being released to the cerebrospinal fluid immediately after the synthesis and will foster the regulation of behaviours and physiological functions for instance, locomotors activity, rest and food intake, skin pigmentation, growth and reproduction (Falco, 1999). These two proteins, rhodopsin and melatonin are correlated to each other in playing role of individual protein as both of them are generated from pineal gland.

2.3.2 Expression of rhodopsin in zebrafish

A study conducted by Magnoli et al., (2012), stated that the levels of rhodopsin at 10 dpf reached values higher twice than those found in adult stage. In contrast in Day-20, Day-30 and Day-40, the levels of the rhodopsin decrease significantly. The level of rhodopsin, continuously to decrease until it reaches negative parameter at Day-50. However, it begins to arise at Day-90. A study has been reported a finding of second rhodopsin-like gene (*rh 1 -2*) which, it only expressed at matured stage of zebrafish. Meanwhile, visual pigment rhodopsin (*rh -1*) is found at early stage of zebrafish life (Morrow, Lazic, & Chang, 2011).

2.4 Transcription factor binding site

Transcription factor is a protein that regulates gene expression and bind to DNA. Meanwhile, transcription factor binding site is a region where transcription factor binds. Transcription factor binding sites are a subset of DNA binding sites and it is a part of either promoter or enhancer region of a gene. An enhancer works best in binding the DNA, which they stimulate transcription at a higher rate than promoter does (Shannon, 2017).

Identification of transcription factor binding site (TFBS) will help the further researchers on practicing the *in vivo* study. Transcription factor binding site (TFBS) is fundamental for the transcriptional control of gene expression in normal embryonic development, the creation and maintenance of tissue particular protein synthesis and the response to particular cellular signalling pathways. It is also vital for RNA splicing occur automatically after identifying TFBS thus produce corresponding protein (Latchman, 1997).

There is one study that was successfully found a single enhancer region near the two red opsin genes of zebrafish. Consequently, it plays important role in their differential expression patter. This study proposed that both of the red opsin genes react in presence of enhancer in restricted manner. However, enhancer or transcription factor of rhodopsin or rh-1 have not been found. Thus, in this study will provide a set of enhancers or transcription factor for rhodopsin (Tsujimura et al., 2010).

CHAPTER 3

MATERIALS AND METHODS

3.1 Materials

This study does not require laboratory materials and expensive tools in carrying out *in silico* study. The most important tools that are required in this study are laptops and good internet access. The laptop is already owned by student and Universiti Malaysia Kelantan (UMK) has provided free internet access.

This *in silico* study requires the use of website and software tool. The website is no need to purchase and can be access online in free of charge. However, MatInspector software tool have a free trial for 30 days and after that purchase is required to further access it. The website that are being used, which is NCBI (https://www.ncbi.nlm.nih.gov/nuccore) and the software tool used is MatInspector software tool (https://www.genomatix.de/online_help/help_matinspector_help.html).

3.2 Methods

The *in silico* study begin by accessing the NCBI website to obtain rhodopsin protein sequence. Then, the TFBS of rhodopsin will be obtained from MATCH™ tool (https://www.genomatix.de/online_help/help_matinspector/matinspector_help.html). Through the suggested TFBS or enhancers proposed, the rhodopsin level will be induced.

CHAPTER 4

RESULTS

The numbers of suggested TFBS for rhodopsin in zebrafish obtained from MatInspector software tool were 421 TFBS. MatInspector also provided additional information for the suggested TFBS which were Sequence name, Accession number, Gene symbol, Gene ID, Matrix family, Detailed family information, Matrix, Detailed matrix information, Tissue, opt., Start position, End position, Anchor position, Genomic start pos, Genomic end pos, Chromosome, Strand, Core similarity, Matrix similarity, Matrix similarity –opt., Evidence #, Evidence and Sequence. Example of the first three suggested TFBS obtained from the MatInspector is as Table A.1 below.

4.1 Selected Transcription Factor Binding Site

Among the 421 TFBS suggested by the MatInspector, this study will further focus on five TFBS selected. The selected TFBS is based on its approximate function with the rhodopsin function and its close regulation that can induce rhodopsin protein. In this study, the detailed family information and detailed matrix information of the selected TFBS will further discuss in the discussion section. The detailed family information is the full name for the matrix family, whilst the detailed matrix information is the full name for the individual matrix.

Table 4.1: The selected TFBS of rhodopsin in zebrafish.

Detailed Family Information	Detailed Matrix Information	Start Position	End Position	Anchor Position	Sequence
Homeodomain transcription factors	H6 family homeobox 1 / NKX5-3	69	87	78	acctgctTTATtgttgcaa
Bicoid-like homeodomain transcription factors	Cone-rod homeobox-containg transcription factor / otx-like homeobox gene	150	166	158	tttatTAATcaagggtt
Heat shock factors	Heat shock factor 1	327	251	339	gagacaactgcaGGAAattactaag
Signal transducer and activator of transcription	Signal transducer and activator of transcription 3	332	350	341	taatTTCCtgcagttgtct
Serum response element binding factor	Serum response factor	428	446	437	gttcgctTATGcgggacct

CHAPTER 5

DISCUSSION

Based on the findings obtained, there are five potential TFBS that can regulate rhodopsin selected. However, some of the TFBS were not mentioned in the subsection due to its similar matrix family with the TFBS mentioned earlier, though it is still discussed under its matrix family. These TFBS are briefly describes its role and relation in regulating the rhodopsin protein in zebrafish.

5.1 Homeodomain transcription factors (H6 family homeobox 1 / NKX5-3)

Throughout early embryonic development, a large family of homeobox genes are responsible in the formation of many body structures (Homeoboxes, 2017). One of the homeobox transcription factors are homeobox 1 (hmx1), which responsible in eye development which belongs to the H6 family of homeobox proteins ("HMX1 H6 family homeobox 1 [Homo sapiens (human)]," 2017). The genes were revealed in the entire eye at 10 somite-of-stage (ss) and later at 18 ss in the dorsal part. Hmx1 plays role as a dimer and dimerization will occur when there is conserved SD1 (Marcelli, Boisset, & Schorderet, 2014). Other than that, this gene act as important task in the development of craniofacial such as eye and ear (Binss et al., 2009).

5.2 Bicoid-like homeodomain transcription factors (Cone-rod homeobox-containing transcription factor / otx-like homeobox gene)

Cone-rod homeobox (Crx) is an Otx-like homeodomain transcription factor which plays a crucial role in the development and maintenance of photoreceptors. Those photoreceptors need a precise regulated gene expression which conciliated by a network of photoreceptor transcription factors centered on Crx. The role of Otx was first discovered when the studies on Drosophila orthodenticle (Otd). Otd is a paired-type homeodomain protein which is crucial for the formation of eye, antenna and anterior brain in the fly. Otd also plays an important role in photoreceptor development where it regulates the expression of opsin genes (Hennig, Peng, & Chen, 2008).

Otx1 and Otx2 transcription factors are belonged to the orthodenticle (Otd) family which play crucial role on proper brain development during early and later events. The functions plays by Otx1 are involved in pituitary functions, corticogenesis and sense organ development. Meanwhile, Otx2 is responsible in the early development of brain for proper anterior neural plate specification and organisation of primitive steak (Acampora, Gulisano, & Simeone, 2000).

5.3 Heat shock factors (Heat shock factor 1)

Heat shock factors are inducible transcriptional regulators of genes encoding molecular chaperones and other stress proteins. It is also important for all organisms as

it plays a major role in order for an organism to survive when exposed to acute stress. Another important roles of this genes are it is responsible for the life-enhancing pathways and normal development (Åkerfelt, Morimoto, & Sistonen, 2010).

5.4 Signal transducer and activator of transcription (Signal transducer and activator of transcription 3)

Signal transducer is used during signal transduction process where it involves in transmission of molecular signals. These molecular signals will be transmitted from a cell's exterior into its interior. It will then transmit within the cells by the secretion of hormones and other chemical factors. One of the molecular signals is sensory signals which are received from external environment for example light, taste, sound, smell and touch. The transmission of signals within the cells should be transmitted effectively by ensuring the cells communicate persuasively among them. It will result the potential of an organism to function normally towards the molecular signals ("Signal Transduction," 2003, "Signal Transduction," 2017).

There are seven family members of Stat has been identified which are Stat1, Stat2, Stat3, Stat4, Stat6, Stat5a and Stat5b. These Stat proteins play a major role in innate and acquired immunity and also involved in a wide range of biological processes. Each Stat protein plays specific functions in the biological process of an organism. For example, Stat1 involved in the interferon γ receptor α-chain, Stat2 forms a complex with Stat1 after the stimulation of the interferon α/β receptor, Stat4 and Stat6 are responsible

for the interleukin-12 and interleukin-4 receptors, Stat3 involve in vast biological processes including differentiation of cells of the granulocytic lineages and macrophages, whilst Stat5a and Stat5b appear the distinctive overlapping for granulocyte colony-stimulating factor, prolactin, erythropoietin, stem cell factor and growth hormone (Aaronson & Horvath, 2002; Benekli et al., 2003).

Signal transducer and activator of transcription 3 (Stat3) is one of the transcription factors that play a major role in regenerating retinal damage which may due to laser injury, heat lesion and intense light. Stat3 functions by involving in the stem cell maintenance and tissue development. It stimulates both the bone morphogenic protein-Smad pathway and the Notch signalling pathway by promoting neuronal stem cell renewal to drive astrogliogenesis. When an extracellular ligand binds a transmembrane receptor and continues with the phosphorylating tyrosine residues on the bound receptor, the Stat3 signalling pathways are triggered. The gene was regulated when the phosphorylated Stat3 dimer pass into the nucleus and binds DNA sequences (Nelson et al., 2012).

According to Tsareva et al. (2007), Stat3 bring about the most complex spectrum of cellular responses including survival, apoptasis and proliferation, depending on the tissue context. Stat3 is expressed only when the retinal is damaged and it is not presence in the undamaged retina (Nelson et al., 2012). A study was conducted by Nelson et al. (2007), to investigate the effect of expression Stat3 or Ascl1a to the photoreceptor cell death. The result of this study stated that the expression of Ascl1a and Stat3 proteins are unrestrictedly and correspond to each other during Müller glia

proliferation at the beginning of the neuronal regeneration of retina damage in adult zebrafish. This is because the Ascl1a is expressed in the proliferating Müller glia, whilst Stat3 is expressed in all Müller glia. Expression of Stat3 in the Müller glia along with the presence of Ascl1a and Lin28a are required in the regenerating of retinal damaged in the adult zebrafish by reentering the Müller glia cell cycle.

5.5 Serum response element binding factor (Serum response factor)

Serum response factor (Srf) is a single highly conserved transcription factor during evolution (Davis et al., 2008; “UniProt: the universal protein knowledgebase,” 2017; Vogel & Gerster, 1999). Srf was identified its mechanism as binding upon a serum stimulating HeLa cells to a motif called the serum response element (Sre) which is a short odf dyad symmetry located 300bp to the 5’ of the transcription initiation in the c-fos promoter (“UniProt: the universal protein knowledgebase,” 2017; Vogel & Gerster, 1999). It controls the expression of variety genes including cell growth and signalling, muscle differentiation, actin cytoskeletal organization and neuronal circuitry (Davis et al., 2008).

CHAPTER 6

CONCLUSION AND RECOMMENDATION

6.1 Conclusion

Based on findings obtained, there were five possible TFBS that can regulate rhodopsin in zebrafish. The five possible TFBS were Homeodomain transcription factors (H6 family homeobox 1 / NKX5-3), Bicoid-like homeodomain transcription factors (Cone-rod homeobox-containing transcription factor / otx-like homeobox gene), Heat shock factors (Heat shock factor 1), Signal transducer and activator of transcription (Signal transducer and activator of transcription 3), and Serum response element binding factor (Serum response factor). Some of the TFBS was discussed indirectly in the subsection of Chapter 5 due to it belongs to the same matrix family with the TFBS mentioned earlier.

The most effective TFBS that can regulate rhodopsin in zebrafish efficiently is Bicoid-like homeodomain transcription factors (Cone-rod homeobox-containing transcription factor / otx-like homeobox gene). Crx is an Otx-like homeodoamain transcription factor plays a crucial role in the development of photoreceptors. Photoreceptors are regulated by various networks of photoreceptor transcription factors which centered on Crx.

6.2 Recommendation

As a recommendation, *in silico* study should be done more broadly in the field of molecular biology experiment before conducting *in* vivo study. This is because most molecular study will involve experimental animal as model organism in their research. The use of experimental animal as model organism in molecular experiment, will lead to the animal suffering due to uncertain determination of humane endpoint. However, this study will be more precise when further analysis is implemented in *in vivo* and study. This is because the suggested TFBSs will be further confirmed its regulation on the rhodopsin protein in zebrafish. Furthermore the best regulated TFBS can be confirmed its regulation while conducting the wet lab.

It is beneficial knowledge if there is more study on the regulation of TFBS of rhodopsin in zebrafish. Rhodopsin is not restricted to its relation in vision. It is also crucial in the growth development, other biological and chemical processes in an individual organism.

REFERENCES

Aaronson, D. S., & Horvath, C. M. (2002). Roadmap for those who dont know JAK-STAT. *Science*, *296*(May), 1653–5. https://doi.org/10.1126/science.1071545

Acampora, D., Gulisano, M., & Simeone, A. (2000). Genetic and molecular roles of Otx homeodomain proteins in head development. *Gene*, *246*(1–2), 23–35. https://doi.org/10.1016/S0378-1119(00)00070-6

Åkerfelt, M., Morimoto, R. I., & Sistonen, L. (2010). Heat shock factors: integrators of cell stress, development and lifespan, *11*(AuguST), 545–555. https://doi.org/10.1038/nrm2938

Allwardt, B. A., & Dowling, J. E. (2001). The pineal gland in wild-type and two zebrafish mutants with retinal defects. *Journal of Neurocytology*, *30*(6), 493–501. https://doi.org/10.1023/A:1015689116620

Bailey, M. J., & Cassone, V. M. (2004). Opsin photoisomerases in the chick retina and pineal gland: Characterization, localization, and circadian regulation. *Investigative Ophthalmology and Visual Science*, *45*(3), 769–775. https://doi.org/10.1167/iovs.03-1125

Benekli, M., Baer, M. R., Baumann, H., & Wetzler, M. (2003). Signal transducer and activator of transcription proteins in leukemias. *Blood*, *101*(8), 2940–2954. https://doi.org/10.1182/blood-2002-04-1204.Supported

Bier, E., & McGinnis, W. (2008). Model organisms in the study of development and disease. *Molecular Basis of Inborn Errors of Development*, 25–48. Retrieved from http://books.google.com/books?hl=en&lr=&id=wGoj9RtTcVIC&oi=fnd&pg=PA25&dq=Model+Organisms+in+the+Study+of+Development+and+Disease&ots=YDD9Pwoz3k&sig=NJpFt15Zrs4K-BOMSZyoXEeLBGE

Binss, D., Dimmer, E., Huntley, R., Barrell, D., O'Donovan, C., & Apweiler, R. (2009). DNA binding. Retrieved December 10, 2017, from https://www.ebi.ac.uk/QuickGO/help

Davis, J. L., Long, X., Georger, M. A., Scott, I. C., Rich, A., & Miano, J. M. (2008). Expression and comparative genomics of two serum response factor genes in zebrafish. *The International Journal of Developmental Biology*, *52*(4), 389–396. https://doi.org/10.1387/ijdb.072514jd

E, T., Q, W., Ab, Q., X, X., NM, Q., NS, P., … MR, A.-U. (2001). The relationship between opsin overexpression and photoreceptor degeneration. Retrieved December 12, 2017, from https://www.ncbi.nlm.nih.gov/pubmed/11222515

Falco, J. (1999). CELLULAR CIRCADIAN CLOCKS IN THE PINEAL, *58*(98).

Fernando, C. T., Goldstein, R. A., Husbands, P., & Stekel, D. J. (2015). - " In silico biology", *480*, 477–480.

Garrido, A. G., Figueiredo, L. F. P. De, & Silva, M. R. E. (2004). Experimental models of sepsis and septic shock : an overview. *Acta Cirúrgica Brasileira*, *19*(2), 82–88. https://doi.org/10.1590/S0102-86502004000200001

Geetha, P., Sivaram, A. J., Jayakumar, R., & Mohan, C. G. (2016). Integration of In Silico Modelling, Prediction by Binding Energy and Experimental Approach to Study the Amorphous Chitin Nanocarriers for Cancer Drug Delivery. *Carbohydrate Polymers*. https://doi.org/10.1016/j.carbpol.2016.01.059

Gehring, W. J. (2014). The evolution of vision, *3*(February). https://doi.org/10.1002/wdev.96

Guither, H. D. (1998). *Animal Rights: History and Scope of a Radical Social Movement* (First Edit). Southern Illinois University Press. Retrieved from http://www.sonoma.edu/users/w/wallsd/pdf/Animal-Rights-Movement.pdf

Hao, H., Kim, D. S., Klocke, B., Johnson, K. R., Cui, K., Gotoh, N., … Swaroop, A. (2012). Transcriptional Regulation of Rod Photoreceptor Homeostasis Revealed by In Vivo NRL Targetome Analysis, *8*(4). https://doi.org/10.1371/journal.pgen.1002649

Hennig, A. K., Peng, G.-H., & Chen, S. (2008). Regulation of photoreceptor gene expression by Crx-associated transcription factor network. *Brain Research*, *1192*, 114–133. https://doi.org/10.1016/j.brainres.2007.06.036

HMX1 H6 family homeobox 1 [Homo sapiens (human)]. (2017). Retrieved December 10, 2017, from https://www.ncbi.nlm.nih.gov/gene/3166

Kato, H. E., Inoue, K., Kandori, H., & Nureki, O. (2016). Prospects & Overviews The light-driven sodium ion pump: A new player in rhodopsin research, 1274–1282. https://doi.org/10.1002/bies.201600065

Koutsoukas, A., Simms, B., Kirchmair, J., Bond, P. J., Whitmore, A. V, Zimmer, S., … Bender, A. (2011). From in silico target prediction to multi-target drug design : Current databases , methods and applications. *Journal of Proteomics*, *74*(12), 2554–2574. https://doi.org/10.1016/j.jprot.2011.05.011

Latchman, D. S. (1997). Transcription factors: An overview. *International Journal of Biochemistry and Cell Biology*, *29*(12), 1305–1312. https://doi.org/10.1016/S1357-2725(97)00085-X

Laurà, R., Magnoli, D., Zichichi, R., Guerrera, M. C., De Carlos, F., Suárez, A. Á., … Germanà, A. (2012). The photoreceptive cells of the pineal gland in adult zebrafish (Danio rerio). *Microscopy Research and Technique*, *75*(3), 359–366. https://doi.org/10.1002/jemt.21064

Magnoli, D., Zichichi, R., Laurà, R., Cristina, M., Campo, S., Carlos, F. De, … Martino, P. U. G. (2012). Rhodopsin expression in the zebrafish pineal gland from larval to adult stage. *Brain Research*, *1442*, 9–14. https://doi.org/10.1016/j.brainres.2012.01.021

Magnoli, D., Zichichi, R., Laurà, R., Guerrera, M. C., Campo, S., De Carlos, F., … Germanà, A. (2012). Rhodopsin expression in the zebrafish pineal gland from larval to adult stage. *Brain Research*, *1442*, 9–14. https://doi.org/10.1016/j.brainres.2012.01.021

Mano, H., Kojima, D., & Fukada, Y. (1999). Exo-rhodopsin: A novel rhodopsin expressed in the zebrafish pineal gland. *Molecular Brain Research*, *73*(1–2), 110–118. https://doi.org/10.1016/S0169-328X(99)00242-9

Marcelli, F., Boisset, G., & Schorderet, D. F. (2014). A dimerized HMX1 inhibits EPHA6/epha4b in mouse and zebrafish retinas. *PLoS ONE*, *9*(6). https://doi.org/10.1371/journal.pone.0100096

Menon, S. T., Han, M., & Sakmar, T. P. (2001). Rhodopsin: structural basis of molecular physiology. *Physiological Reviews*, *81*(4), 1659–1688.

Mohd Ramli, S. Z., Samsuddin, N. A., Abdullah, J. M., Wan Ismail, W. F. N., Yusof, A. H., Alias, A., … Shamsuddin, S. (2012). *Expression analysis of transcription factors CTCF (CCCTC-binding factor), BORIS (Brother of the Regulator Imprinted Sites ...* Biomedical Research. Retrieved from https://www.researchgate.net/profile/Shaharum_Shamsuddin/publication/295624290_Expression_analysis_of_transcription_factors_CTCF_CCCTC-binding_factor_BORIS_Brother_of_the_Regulator_Imprinted_Sites_and_YB-1_Y-box_binding_factor_1_in_osteosarcoma_and_glioma

Morrow, J. M., Lazic, S., & Chang, B. S. W. (2011). A novel rhodopsin-like gene expressed in zebrafish retina. *Visual Neuroscience*, *28*(4), 325–335.

https://doi.org/10.1017/S0952523811000010
Nagloo, N., Hart, N. S., & Collin, S. P. (2017). Seriola aquaculture industry. *Aquaculture*, *474*(September 2016), 130–137. https://doi.org/10.1016/j.aquaculture.2017.03.042
Nelson, C. M., Gorsuch, R. A., Bailey, T. J., Ackerman, K. M., Kassen, S. C., & Hyde, D. R. (2012). Stat3 defines three populations of müller glia and is required for initiating maximal müller glia proliferation in the regenerating zebrafish retina. *The Journal of Comparative Neurology*, *520*(18), 4294–4311. https://doi.org/10.1002/cne.23213
Nemzek, J. A., Xiao, H. Y., Minard, A. E., Bolgos, G. L., & Remick, D. G. (2004). Humane Endpoints in Shock Research. *Shock*, *22*(1), 17–25. https://doi.org/10.1097/01.shk.0000133596.50443.5e
New World Encyclopedia contributors. (2016). John von Neumann. Retrieved December 2, 2017, from http://www.newworldencyclopedia.org/p/index.php?title=John_von_Neumann&oldid=999197
Okano, K., Okano, T., Yoshikawa, T., Masuda, A., Fukada, Y., & Oishi, T. (2000). Diversity of Opsin Immunoreactivities in the Extraretinal Tissues of Four Anuran Amphibians, (October 1998), 136–142.
Olsson, J. E., Gordon, J. W., Pawlyk, B. S., Roof, D., Hayes, A., Molday, R. S., … Dryja, T. P. (1992). Transgenic Mice with a Rhodopsin Mutation (Pro23His): A Mouse Model of Autosomal Dominant Retinitis Pigmentosa. *Neuron*, *9*, 815–830.
rho - Rhodopsin Danio Rerio. (n.d.). Retrieved March 25, 2017, from http://www.uniprot.org/uniprot/P35359
Sana Banu, S., Chandhana, K., Thahajeb, S., Roopa Vani, V., Vishnu Prriya, N., & V Krishnareddy, B. (2015). SRI VENKATESWARA COLLEGE OF PHARMACY , RVS NAGAR , TIRUPATHI ROAD , CHITTOOR- 517127 . (A . P) " ICMR SPONSORED ONE DAY NATIONAL SEMINAR " ON " ALTERNATIVE TO ANIMAL USE IN RESEARCH , EDUCATION AND TOXICITY INVESTIGATION " ON 22 nd SEP 2015. In *Impact of In-Silico Predictive Pharmacology and Toxicology Studies on Usage Experimental Animals Used in the Drug Discovery and Development* (p. 30). Retrieved from http://www.ijabpr.com/images/Sri-Venkateswara-College.pdf
Shannon, C. (2017). Transcription Factor Binding Site: Prediction and concept. Retrieved March 16, 2017, from http://study.com/academy/lesson/transcription-factor-binding-site-prediction-lesson-quiz.html
Signal Transduction. (2003). Retrieved December 12, 2017, from https://encyclopedia2.thefreedictionary.com/Signal+transducer
Signal Transduction. (2017). Retrieved December 12, 2017, from https://www.tocris.com/cell-biology/signal-transduction
Sohocki, M. M., Daiger, S. P., Bowne, S. J., Rodriquez, J. A., Northrup, H., Heckenlively, J. R., … Sullivan, L. S. (2008). Prevalence of Mutations Causing Retinitis Pigmentosa and Other Inherited Retinopathies. *National Institutes of Health*, *17*(1), 42–51. https://doi.org/10.1002/1098-1004(2001)17
Terakita, A. (2005). The opsins. *Genome Biology*, *6*(3), 213. https://doi.org/10.1186/gb-2005-6-3-213
Tsujimura, T., Hosoya, T., & Kawamura, S. (2010). A single enhancer regulating the differential expression of duplicated red-sensitive opsin genes in Zebrafish. *PLoS Genetics*, *6*(12), 1–10. https://doi.org/10.1371/journal.pgen.1001245
Uchida, S. (2012). *Annotating New Genes from In Silico Screening to Experimental Validation*. Woodhead Publishing Limited.
UniProt: the universal protein knowledgebase. (2017). *Nucleic Acids Research*, *45*(D1), D158–D169. https://doi.org/10.1093/nar/gkw1099
Vigh, B., Jotio, M., & Szkl, A. (1998). The pineal organ as a folded retina :

Immunocytochemical localization of opsins *, 653–659.
Vodovotz, Y., Chow, C. C., Bartels, J., Lagoa, C., Prince, J. M., Levy, R. M., … Clermont, G. (2006). in Silico Models of Acute Inflammation in Animals. *Shock*, *26*(3), 235–244. https://doi.org/10.1097/01.shk.0000225413.13866.fo
Vogel, A. M., & Gerster, T. (1999). A zebrafish homolog of the serum response factor gene is highly expressed in differentiating embryonic myocytes. *Mechanisms of Development*, *81*(1–2), 217–221. https://doi.org/10.1016/S0925-4773(98)00245-7

APPENDIX A

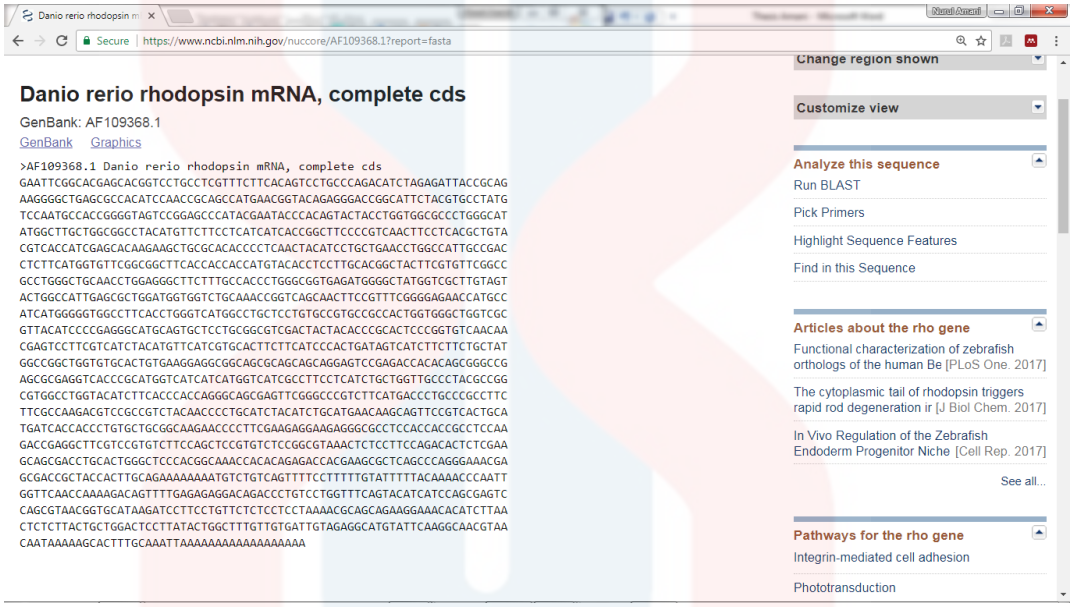


Figure A.1: The complete genome sequence of rhodopsin in *Danio rerio*. Picture was generated from (Nucleotide [Internet], 2004).

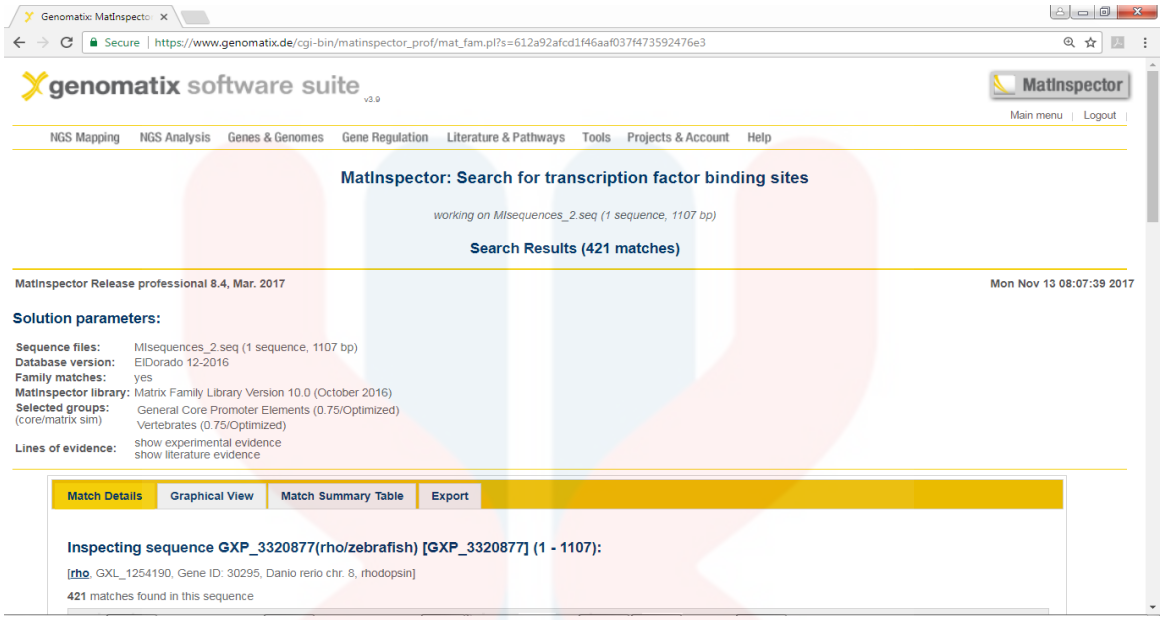


Figure A.2: The results of transcription factor binding site for rhodopsin in zebrafish 1. Picture was generated from (K, 2005).

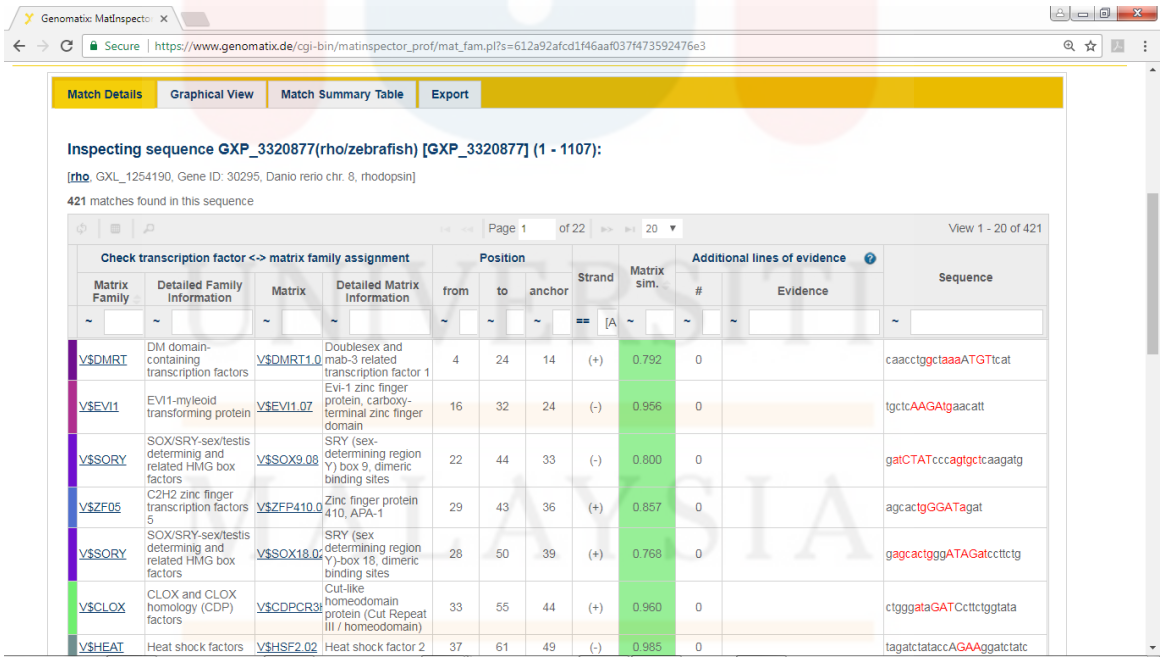
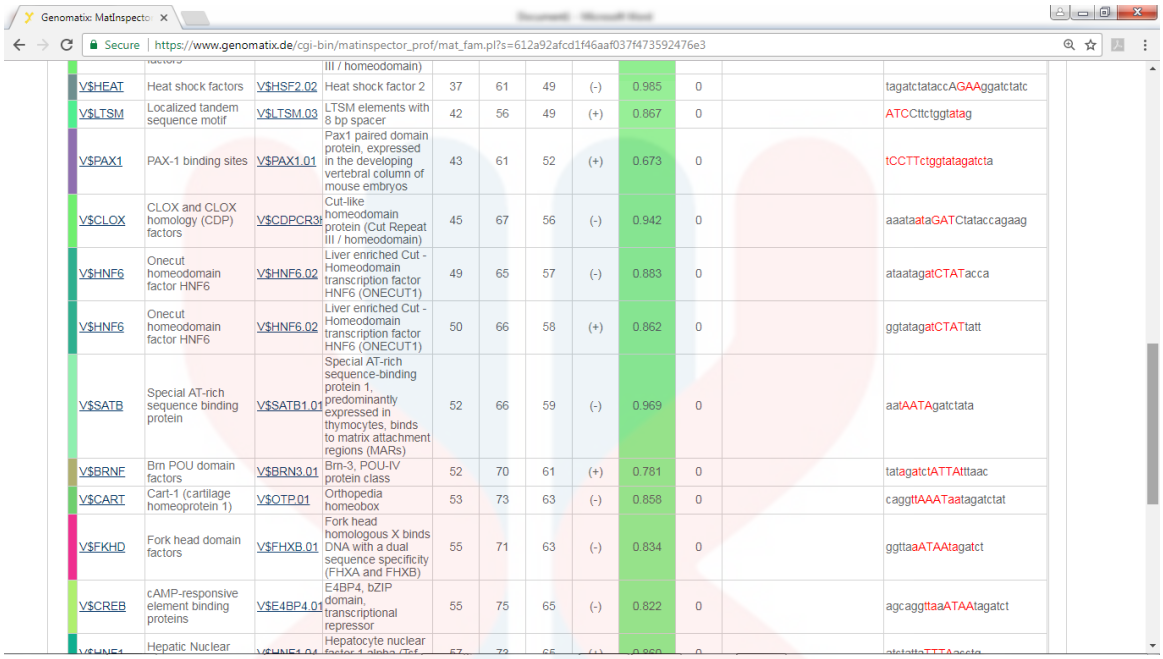


Figure A.3: The results of transcription factor binding site for rhodopsin in zebrafish 2. Picture was generated from (K, 2005).



			III / homeodomain)								
V$HEAT	Heat shock factors	V$HSF2.02	Heat shock factor 2	37	61	49	(-)	0.985	0		tagatctataccAGAAggatctatc
V$LTSM	Localized tandem sequence motif	V$LTSM.03	LTSM elements with 8 bp spacer	42	56	49	(+)	0.867	0		ATCCttctggtatag
V$PAX1	PAX-1 binding sites	V$PAX1.01	Pax1 paired domain protein, expressed in the developing vertebral column of mouse embryos	43	61	52	(+)	0.673	0		tCCTTctggtatagatcta
V$CLOX	CLOX and CLOX homology (CDP) factors	V$CDPCR3	Cut-like homeodomain protein (Cut Repeat III / homeodomain)	45	67	56	(-)	0.942	0		aaataatagATCtataccagaag
V$HNF6	Onecut homeodomain factor HNF6	V$HNF6.02	Liver enriched Cut - Homeodomain transcription factor HNF6 (ONECUT1)	49	65	57	(-)	0.883	0		ataatagatCTATacca
V$HNF6	Onecut homeodomain factor HNF6	V$HNF6.02	Liver enriched Cut - Homeodomain transcription factor HNF6 (ONECUT1)	50	66	58	(+)	0.862	0		ggtatagatCTATtatt
V$SATB	Special AT-rich sequence binding protein	V$SATB1.01	Special AT-rich sequence-binding protein 1, predominantly expressed in thymocytes, binds to matrix attachment regions (MARs)	52	66	59	(-)	0.969	0		aatAATAgatctata
V$BRNF	Brn POU domain factors	V$BRN3.01	Brn-3, POU-IV protein class	52	70	61	(+)	0.781	0		tatagatctATTAtttaac
V$CART	Cart-1 (cartilage homeoprotein 1)	V$OTP.01	Orthopedia homeobox	53	73	63	(-)	0.858	0		caggttAAATaatagatctat
V$FKHD	Fork head domain factors	V$FHXB.01	Fork head homologous X binds DNA with a dual sequence specificity (FHXA and FHXB)	55	71	63	(-)	0.834	0		ggttaaATAAtagatct
V$CREB	cAMP-responsive element binding proteins	V$E4BP4.01	E4BP4, bZIP domain, transcriptional repressor	55	75	65	(-)	0.822	0		agcaggttaaATAAtagatct
V$HNF1	Hepatic Nuclear		Hepatocyte nuclear								

Figure A.4: The results of transcription factor binding site for rhodopsin inzebrafish. Picture was generated from (K, 2005).

Table A.1: Example of the first three suggested TFBS from the MatInspector software tool.

Seq. name	Accession no.	Gene symbol	Gene ID	Matrix Family	Detailed Family Information	Matrix	Detailed Matrix Information	Tissue	Opt.	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Chromosome	Strand	Core sim.	Matrix sim.	Mat. sim. – opt.	Evidence #	Evidence	Sequence
GXP_3320877 (rho/zebrafish)	GXP_332087 7	rho	3029 5	V$DMRT	DM domain-containing transcription factors	V$DMRT1.01	Doublesex and mab-3 related transcription factor 1	Embryonic Structures Endocrine System Germ Cells Ovary Testis Urogenital System	0.77	4	24	14	5367131 8	5367129 8	8	+	0.934	0.792	0.022	0		caacctggctaaaATGTtcat

GXP_320877 (rho/zebrafish)	GXP_32087 7	rho	3029 5	V$EVI1	EVI1-myleoid transforming protein	V$EVI1.07	Evi-1 zinc finger protein, carboxy-terminal zinc finger domain	Adipose Tissue Antibody-Producing Cells Blood Cells Bone Marrow Cells Brain Central Nervous System Connective Tissue Erythrocytes Hematopoietic System Immun	0.9	16	32	24	5367130 6	5367129 0	8	-	1	0.956	0.056	0		tgctcAAGAtgaacatt

								e System Nervous System Neurons Thymus Gland														
GXP_320877 (rho/zebrafish)	GXP_32087	rho	3029	V$SORY	SOX/SRY-sex/testis determinig and related HMG box factors	V$SOX9.08	SRY (sex-determining region Y) box 9, dimeric binding sites	Bone and Bones Cartilage Connective Tissue Embryonic Structures Endocrine System Leydig Cells	0.69	22	44	33	5367 1300	5367 1278	8	-	0.75	0.8	0.11	0		gatCTATcccagtgctcaagatg

	Nervous System Neuroglia Skeleton Testis Thyroid Gland Urogenital System