Working title:
Analysis of action sequences in zebrafish.

Student: Jerneja Mislej

University of Lund

bif15jmi@student.lu.se

Project supervisor: Fredrik Ek

University of Lund

Fredrik.Ek@med.lu.se

Assistant project supervisor: Mauno Vihinen

University of Lund

Mauno.Vihinen@med.lu.se

Abstract

The master thesis project objective is to design, implement and preform the analysis of zebrafish larvae action sequences.

Zebrafish larvae movement was recorded with a high speed camera and processed by a tracking software. The results were even further summarized into different types of turns, defined by direction and angle. Different turns are represented with designated letters and the final sequence of turns stored in a Fasta file format. Such representation of information offers the possibility of using tools, algorithms and principals already highly developed in protein and DNA sequence analysis.

The projects work will revolve around the adaption of these approaches, with the aim of finding activity patterns in neurodegenerative and neuropsychiatric diseases.

I Basic information

Within the masters program of Bioinformatics at Lund University, Department of Biology, I will carry out a 30 credit master thesis project under the code BINP30.

The project will take place at the department of Experimental Medical Science under the supervision of Fredrik Ek from the Chemical Biology and Therapeutics research group and co-supervision of Mauno Vihinen from the Protein Structure and Bioinformatics research group.

The project will start in week 45 on 7th of November 2016 and will finish in week 23, on the 7th of June 2017.

II Project

I Introduction

The project will be concerned with the data collected from zebrafish larvae movement. There are several reasons that make zebrafish suitable as a model organism, especially when it comes to evaluation of new drugs. Since they belong to the subphylum Vertebrata as humans do, there is a substantial degree of relation between them, resulting in an average of approximately 70% sequence homology. It has been proven that zebrafish larvae respond to drugs in a similar way as more advance animals[1], since the structure and molecular pathways of the central nervous system are highly conserved between mammals and zebrafish[2].

At the department of Experimental Medical Science at Lund university, the Chemical Biology and Therapeutics research group has developed an infrastructure for behavioral analysis of zebrafish. With the infrastructure, the behavior of up to 192 zebrafish larvae can be recorded in parallel, using a high speed camera and a tracking software. Since zebrafish larvae have specific and distinct movement patterns (swim bouts), that can be used to model behavioral and cognitive aspects of neurological diseases[3][4], the recordings from the high speed camera were processed and summarized into a sequence of turns. Each turn is classified according to the direction and angle. All together, eight different turns have been categorized and coded with a designated letter. Additional codons are present in the action sequence, these codons are used to determine the direction of the turns, joint turns and separation between bouts. The final action sequence is stored in a Fasta format file, to enable meta information to be stored along with the corresponding sequence.

Simple example of such an action sequence in a Fasta format:

```
>ZF Recording 2016-04- 01; tW: 300s; Start: 122917; End: 123417; Recording:
1-1; Individual: 9; Drug: Clozapine 1 mM PTZ 2.5 mM
bLsbLibLibLsbLsLsLsmmbLiLsbLgLsbLgLsbLsLsmbLsbLgbLgRsbLibRgbLgbLgbLgLs
bLgbLgbLsLsmbLsbLsLsmbLgbLgbLgbLgbLsLsmbLgbLsRsbLsbLsLsmbLgbLsbLgbLgbL
gbLgbRsbLgbLsbLsLsmbLsbLgbLgRsbLgLsbLgbLsbLsbLebRebLsbLsbLsbLibLsbLsbL
gbLgbLgbLsbLsbLgbLgbLgbRebRiRsmbRiRsRsmmbRgLsbLibRebRhRsmbLibLjbLebReb
```

II Project aims

Project aims to analyze the zebrafish larvae action sequences, stored in Fasta format files. As such coding of zebrafish larvae movement and storage of action sequences in Fasta file format is a novelty, the projects aims include and start with familiarization with such an approach, along with the exploration and assessment of feasibility of analysis. Further on, the project will also aim to consider the broader research behind the master thesis project, which is to find activity patterns in neurodegenerative and neuropsychiatric diseases, focusing on drug effects. This will serve as guidance in the design and implementation of the action sequence analysis. Since researchers would like to identify the changes from exposure to drugs that effect the major neurotransmitter system in the brain, the master thesis project will aim to design and implement action sequence similarity algorithms, extract action sequence key features and motifs and preform the final statistical analysis.

III Methods

Zebrafish larvae action sequences are stored in Fasta format files. These files contain the description line and the corresponding sequences consisting of codons representing various swim bouts. Such information representation is similar to the one of protein and DNA sequences. Since the area of protein and DNA sequence analysis is highly developed and researched, the master thesis project methods will be based on it, focusing on protein sequence analysis. For this reason, the master thesis project will be co-supervised and supported by the Protein Structure and Bioinformatics research group of Lund university. As the zebrafish larvae action sequences substantially differ from the protein sequences, the methods will attempt to adapt the existing tools and algorithms or implement new ones, on the existing principles. One of such principles is the protein sequence alignment or similarity search, which has many different solutions and algorithms, which commonly depend on a scoring(substitution) matrix[5]. These algorithms could easily be adapted to assess similarity between zebrafish larvae action sequences by using a special scoring matrix designed and developed in collaboration with the researchers conducting the zebrafish larvae movement experiments, who have the appropriate knowledge of the relation and biological meaning of different turns. Other methods include sequence feature extraction, like codon frequency, frequency of codon pairs or even further, codon patterns or motifs. The results obtained from the action sequence analysis will be further analyzed using various statistical methods, in order to address the main research questions.

IV Timeline

- 7.11.2016 - 7.12.2016: Project starts, the student gets acquainted with the topic and available data.

- 8.12.2016 - 8.1.2017: Preparing and processing Fasta files containing zebrafish larvae action sequences.

- 9.1.2017 - 9.3.2017: Design and implementation of the action sequence analysis methods.

- 10.3.2017 - 1.5.2017: Statistical analyses.

- 2.5.2017 - 19.5.2017: Completion and closure.

- 20.5.2017 - 7.6.2017: Master thesis report and presentation preparation.

References

[1] Ek F, Malo M, Andersson MA, Wedding C, Kronborg J, Svensson P, Waters S, Petersson P, Olsson R (2016) Behavioral Analysis of Dopaminergic Activation in Zebrafish and Rats Reveals Similar Phenotypes
ACS Chemical Neuroscience. May 2016, Vol. 7, No. 5: 633-646

[2] Stewart AM, Braubach O, Spitsbergen J, Gerlai R, Kalueff AV. (2014) Zebrafish models for translational neuroscience research: from tank to bedside.
Trends Neurosci. 2014;37(5):264-78. doi: 10.1016/j.tins.2014.02.011.

[3] Kasparek T, Rehulova J, Kerkovsky M, Sprlakova A, Mechl M, Mikl M (2013) Mind the fish: zebrafish as a model in cognitive social neuroscience.
Front Neural Circuits. 2013 Aug 8;7:131. doi: 10.3389/fncir.2013.00131.

[4] Kasparek *et. al* (2012) Cortico-cerebellar functional connectivity and sequencing of movements in schizophrenia
BMC Psychiatry 2012, 12:17 http://www.biomedcentral.com/1471-244X/12/17

[5] Pearson WR (2013) An Introduction to Sequence Similarity ('Homology') Searching
Curr Protoc Bioinformatics. Chapter 3: Unit3 1 doi: 10.1002/0471250953.bi0301s42