Software requirements speciﬁcation (SRS)

Team name: ZTTZX

**Contents**

1. Introduction

1.1 Purpose

1.2 Document conventions

1.3 Intended audience

1.4 Project Scope

1.5 Contact information/SRS team members

1.6 References

2. Overall Description

2.1 Product perspective <- context, overview

2.2 Product functions \*\*\*

2.3 User classes and characteristics

2.4 Operating environment

2.5 Design/implementation constraints \*\*\*

2.6 Assumptions and dependencies

3. External Interface Requirements

3.1 User interfaces

3.2 Hardware interfaces

3.3 Software interfaces

3.4 Communication protocols and interfaces

4. System Features

4.1 System feature A

4.1.1 Description and priority

4.1.2 Action/result

4.1.3 Functional requirements

5. Other Non-functional Requirements

5.1 Performance requirements

5.2 Safety requirements

5.3 Security requirements

5.4 Software quality attributes

6. Other Requirements

Appendix A: Terminology/Glossary/Definitions list

Appendix B: To be determined

**Preface**

This document describes the Little Hill Lab's initial requirements for an online application (Oh My Genes) which allows our scientists to upload gene expression files and quickly get differentially expressed genes.

**1. Introduction**

**1.1 Purpose**

This text descripts the application’s (Oh My Genes) request. The text will be used by the members of the project developers who develop the system function and who test the correctness of the system function.

**1.2 Document conventions**

In this text, it will use font small 2 and overstriking for primary title, font small 3 for secondary title and font 4 for the content. And it will use the italic when mentions the name of the application Oh My Genes.

**1.3 Intended audience**

This SRS about Oh My Genes is for developers, project managers, salesmen, users and testers. The article mainly introduces the overall description, external interface requirements, system features and other non-functional requirements. I suppose the developers and project managers to read the whole article carefully and salesmen pay attention to overall description especially. Users and testers read the system features carefully.

**1.4 Project Scope**

Oh My Genes is used to identify differentially expressed genes given a gene expression file containing two cell samples. And to decrease the workload of a number of repeated calculations, then draw a visual picture to show the result. You can see the detailed content the article about the application GEO2R GEO online analysis tool GEO2R analysis of differential expression genes.

**1.5 Contact information/SRS team members**

SRS team members: Zhang Tingting (id: 201636900118)

Zhao Xuan (id: 201636900119)

**1.6 References**

【1】Biological longed for class write the GEO online analysis tool GEO2R analysis of differential expression genes and the address is http://agetouch.blog.163.com/blog/static/2285350902016101083320329

**2. Overall Description**

**2.1 Product perspective <- context, overview**

A web application preferably using the Flask framework (<http://flask.pocoo.org/>).

**2.2 Product functions \*\*\***

Help the users to use this application to grab the information of genes, such as different expressions of the two cells.

**2.2.1 Overview**

The web application has a simple interface with a single button [Upload and GO].

Our scientists upload a plain text file containing gene expression levels from two samples, representing two experimental conditions. Accepting the file, the software will return a table of differentially expressed genes and a scatter plot of these genes whose X-axis is control and Y-axis is treatment. If an invalid gene expression is given, the web application returns a page informing the user to provide the correct format.

**2.2.2 Input**

A valid submitted gene expression file has the following format. It is a TAB-delimited, plain text file with three columns (see the attached file for a full example). The file contains an optional head line, followed by each gene's expression in a control sample (e.g., ControlSample) and in a treatment sample (e.g., KnockOutSample).

gene\_id ControlSample KnockOutSample

AT1G01010 1.198558083 2.036161827

AT1G01020 13.75736234 13.370796

AT1G01030 0.833779536 0.203616183

AT1G01040 9.58846466 7.126566394

AT1G01046 0 0

AT1G01050 23.81482799 21.10821094

AT1G01060 0.625334652 1.221697096

AT1G01070 1.719670292 0.950208853

AT1G01080 28.34850421 25.24840665

AT1G01090 58.26034505 42.96301455

AT1G01100 1066.508249 1308.030358

AT1G01110 2.709783491 1.425313279

**2.2.3 Output**

The web application displays a table and a scatter plot given a gene expression file.

The table contains a list of differentially expressed genes with the following format:

gene\_id control\_sample treat\_sample log\_2[FC]

AT1G01010 1.198558083 2.036161827 0.76

The scatter plot displays differentially expressed genes. The X-axis is Control, and Y-axis is Treatment.

Replace 'Control' and 'Treatment' with appropriated column names if provided in the uploaded file. The up-regulated genes are shown in red dots, and down-regulated genes are shown in blue.

**2.2.4 Response time**

<5 seconds.

**2.3 User classes and characteristics**

Our application mainly serves the people who study on the genee. Mostly they are biologists and scientists, but it doesn’t rule out some other people want to use this application to get some useful information.

**2.4 Operating environment**

Our software is a multi-functional software system based on the windows platform. It is compatible and can run on 64 - bit laptop or ordinary desktop.

**2.5Design/implementation constraints \*\*\***

Our application must upload the gene file and then get the result, it seems too fussy. We must consider about the arrangement and beautification of the interface, Prioritization of processing operations and it deepens the difficulty of coding and testing. This application needs users to upload the gene file, but the result is given by the application after it analyses. Every one uploads different files, so the application must analyses a variety of documents. It will raise the error rate.

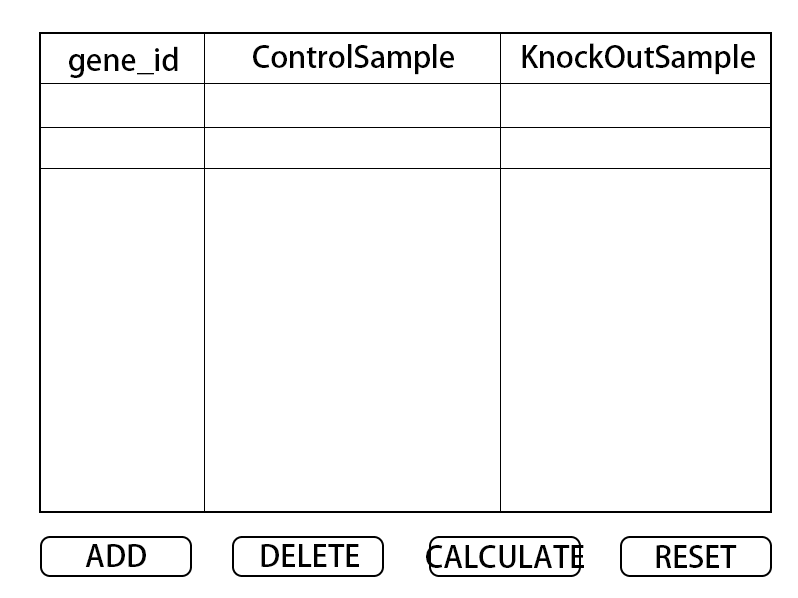
**2.6 Assumptions and dependencies**

This application needs users to upload the file to get the information what the users want. It means it will have lots of new genes needing our updating. In other words, the people who manage the application should know about the knowledge of the genes.

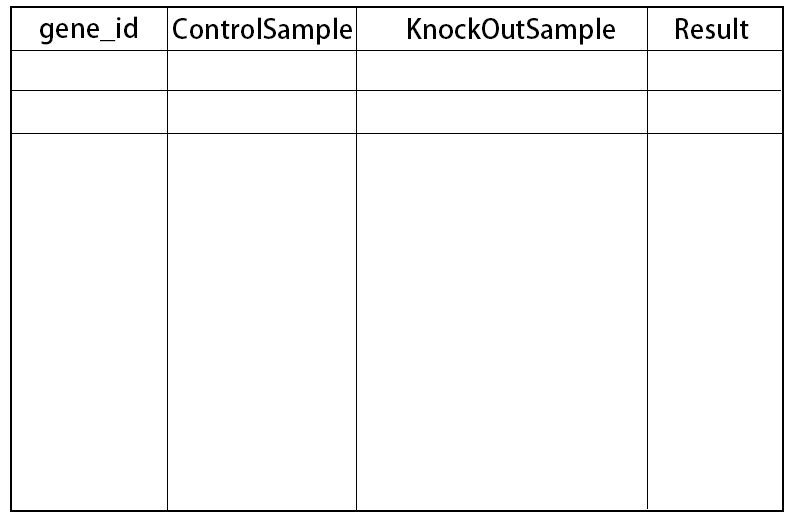
**3. External Interface Requirements**

**3.1 User interfaces**

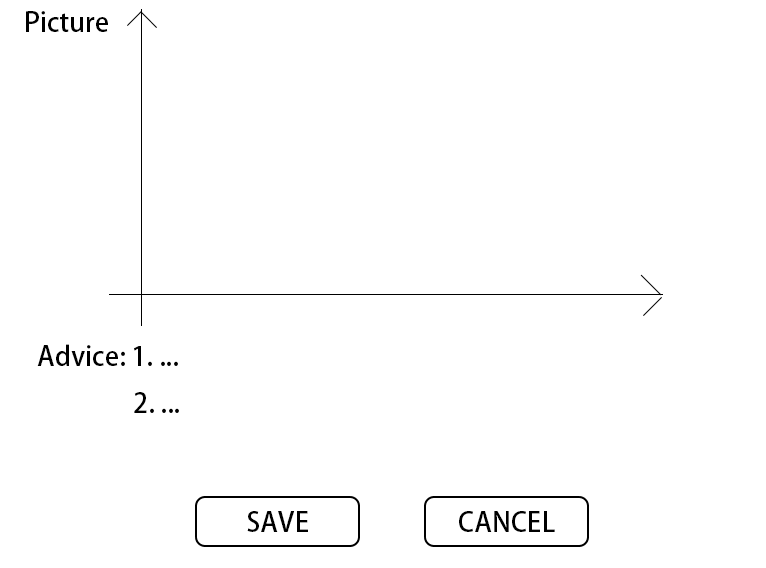
UI-1: It will show you a table, the first column you can input the *gene\_id,* the second column you can input the ControlSample and the last column you can input KnockOutSample. It will show you four rows first if you want to add rows click the add key and vise versa. Or you can input the data by database. And at the end of the table it has two buttons: calculate and reset.



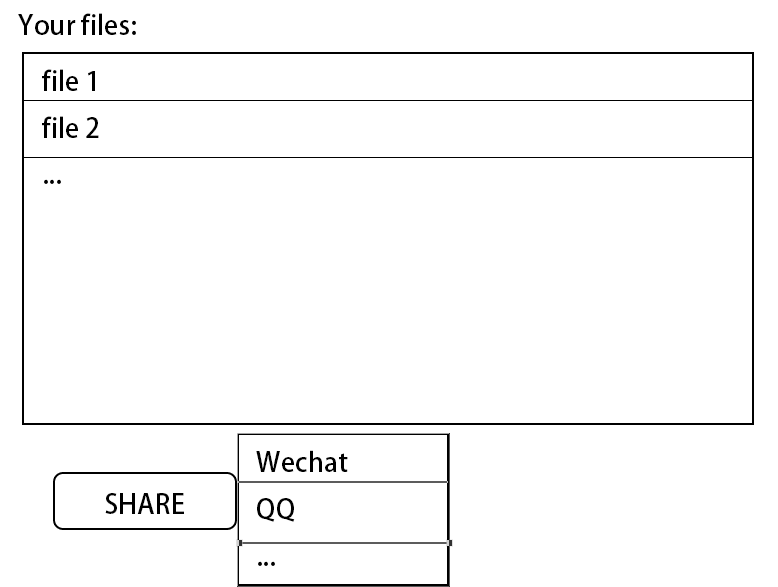
UI-2: It will show you what you input before and the result of the calculation at the forth column.



UI-3: It will show you the picture about the calculating result and give advice about the genes. And it has two buttons at the bottom of the page: save and cancel.



UI-4: It will show you the list of file you saved before. And you can see the completed result in this page. And you can transfer the file to others through this page.



**3.2 Hardware interfaces**

The requests of the hardware for the web application are as followed:

1. 64 bits laptop or desktop.

2. TCP protocol.

**3.3 Software interfaces**

The web application will create links between these as followed:

1. IE browser or Microsoft-edge browser.

2. SQL server.

3. Windows operating system.

**3.4 Communication protocols and interfaces**

1. TCP protocol.

2. Secure transmission and encryption techniques.

3. Microsoft-edge browser.

4. File transfer rate.

**4. System Features**

**4.1 System feature A**

**4.1.1 Description and priority**

Add and update are the highest priorities in this web application. Every time the users upload a new file, the application should save and update the gene list.

Regular log records and updates are designed to better improve our software system, which is lower in comparison, so the priority is low in the software system.

The last main user group to manage and maintain the data directly in the background is the database administrator. Although the group is not very large, its importance can not be ignored. After all, the highlight of the software system is the database, so its priority is the highest. I believe this is also the center of the application database system software.

**4.1.2 Action/result**

**Action:** There are two buttons on the web application (upload and go). Users upload a plain text file containing gene expression levels from two samples, representing two experimental conditions.

**Result:** Accepting the file, the software will return a table of differentially expressed genes and a scatter plot of these genes whose X-axis is control and Y-axis is treatment. If an invalid gene expression is given, the web application returns a page informing the user to provide the correct format.

**4.1.3 Functional requirements**

The software needs to draw the uploaded data to graph. The points on the graph will show the users some information, such as gene id and name.

This web application can update the data intelligently, and save the new genes into the data base.

**5. Other Non-functional Requirements**

This part will show you other non-functional requirements. It contains four parts about performance requirements, security requirements, software quality attributes and user documentation.

**5.1 Performance requirements**

It can support large amount of calculations and the response time should in 3 seconds whether the file is large or not.

Users can add or delete the data and the web application can save them timely. And the application can’t occupy large space.

And the picture of the result should be clear either the points or the lines and what the coordinates present for and the correct units.

**5.2 Safety requirements**

The data users input or from database should be safety. And the whole file transfer to others should be security as we can encrypt it.

And it is important to create backups because there always exit careless users. And if the application has errors, backups can decrease the loss.

And the application should proceed the user authentication.

**5.3 Security requirements**

First, the web application should be easy to understand and operate.

Second, it’s easy to modify the file and check the file.

Finally, it can be used in many different machines.

**5.4 Software quality attributes**

This part will contain two requests of the user documentation:

1. User’s manual: electronic document.

2. Tutorial: electronic document.

**6. Other Requirements**

**Appendix A: Terminology/Glossary/Definitions list**

**Control sample** - a cell sample prepared in its normal condition.

**Treatment sample** - a cell sample treated by special chemicals, or in which some genes are altered.

**Differentially expressed genes** - the genes which have significantly different expression levels between two samples.

**Up-regulation** - a gene is said to be up-regulated if it has higher expression in treatment than in control.

**logFC** - log fold change of gene expression. log\_2 [T/C], where T is the gene expression level from a treatment sample, while C is the gene expression level from a control sample.

**the Flask framework -** Flask is a lightweight Web application framework written in Python. The WSGI toolbox uses Werkzeug, and the template engine uses Jinja2. Flask uses BSD authorization.

Flask is also known as "micro-framework" because it uses simple core and uses extension to add other functions. Flask has no default database and form validation tools.

**Transmission Control Protocol -** TCP provides reliable data transmission in IP environment. Its services include data flow transmission, reliability, effective flow control, full duplex operation and multiplexing. Sending through connection oriented, end-to-end and reliable packets

**Appendix B: To be determined**

Due to the limitations of personal level, we know little about the technology of information security. If there is time to finish in the later stage, the message encryption technology will be added.