**Software Requirements Documentation for “Oh My Genes”**

Yeboah Martha A.

2019-03-26

**Outline**

1. **Introduction**
   1. Purpose
   2. Overview
   3. User Characteristics
   4. Terminologies and Abbreviations
2. **Functional Requirements**
   1. Input
   2. Output(Data Analyzing)
   3. User Interface
3. **Non-functional Requirements**
   1. Response time
   2. Aesthetic Aspects
   3. Confidentiality Policy
4. **Constraints**
   1. Browser Compatibility
   2. Space Complexity
   3. Budget
   4. System Downtime
5. **Change Cases**
6. **Milestone**
7. **Appendices**

**8 References**

**Introduction**

* 1. Purpose

This document is intended to specify a set of requirements for the project OMG, which means “Oh My Genes”, a web application for identifying genes’ expression differentially. Comparing two samples, it’s quite useful to analyze the different performances between the same genes automatically and quickly. The main purpose is to help biologist to analyze their experiment data.

* 1. Overview

The web application has a simple interface with a single button [Upload file]. Our scientist upload a plain text containing gene expression levels from two samples, representing two experimental conditions. Accepting the file and if an invalid expression is given, the web application returns a page informing the user to provide the correct format.

* 1. User Characteristics

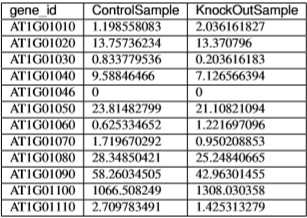
Biologists, researchers or students in related area and technicians who have the basic knowledge of flask and python to maintain a website.

* 1. Terminologies and Abbreviations Explanation
* 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.
* OMG; Oh My Genes, which is name of this project.
* 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

1. **Functional Requirements**
   1. Input

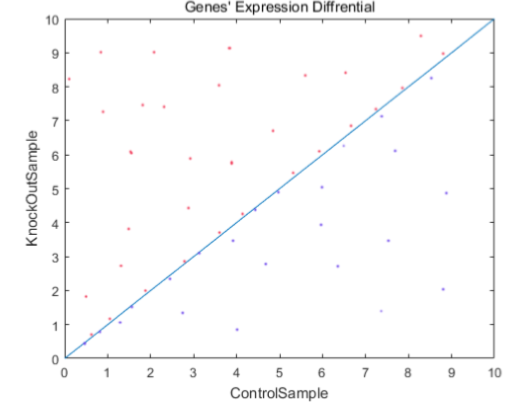
A valid submitted gene expression file has the following format:

* It is TAB-delimited, plain text file with three columns
* 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).



Accepting the file, the software will return a table of differentially expressed genes and a scatter plot of genes whose X-axis is control and Y-axis is treatment.





1. **Non-functional Requirements**
   1. Response Time

Time complexity must be taken into consideration if users have a big amount of data to analyze. In five seconds, we will give back the results.

* 1. Aesthetic Aspects

Simple is the best

* 1. Confidentiality Policy

The web application must maintain the experiment data’s reliability, integration and secrecy and the experimental data will not be divulged without the agreement of the owner.

1. **Constraints**

Before we develop the project, we have to take some main constraints into account so that we can come up any useful ideas with them in the future.

* 1. Browser Compatibility

The application must be accessible through Firefox, Chrome, and Safari.

* 1. Space Complexity

An excellent system should not need too much memory, so that it will run much faster and convenient and we will not cost many internal storages for this system. Thus, the Web space is supposed to be less than 1GB.

* 1. Budget

Budget less than 10,000 USD.

* 1. System Downtime

The downtime of the system is less than 30 minutes per year.

1. **Change Cases**

There are still some other expectation we possess as following:

In order to improve the practicability of this system and accomplish our research better, we expect that this system can give a few potential functions. For example, this system can set a threshold for log\_2[FC] by itself. Or if we upload several data, it can give the response several scatter plot, which will improve workpiece ratio of us than just upload one data table once. Next, we expect this system can have a login judge system to prevent people from outside using this system, what we mean is like a firewall.

Besides, we hope this system can judge whether the file that we upload is valid and whether it has existed or related to our research. If it is invalid, we hope it will appear some hint to guide us upload the correct file. If it has existed, the system can give corresponding information. Afterwards, we hope this system can clean up memory regularly and data recovery function in case system crash. At the same time, we hope this system can recover and display the result last time when we biologists open it next time.

And we want this system give other functions not only to identify differentially expressed genes but give a gene expression file containing two cell samples, we want more cell samples and we want to test the expression levels of genes through scientific calculations. Last but not least, our biologists may come from all over the world, so this system must have its own complicated translation function. Then, this system must be in a simple style.

1. **Milestones**
2. We want a SRS to get a rough idea by 1 May 2019.
3. SRS better before 15 May 2019.
4. We expect to get design done by yourself to using Flask and learn what you need when you design this system from internet. But it just can learn rather than copy, because we need a self-designed system which is safe than a copy system.
5. We expect this system get coding done by Python in main, Java in assistant. Because Python and Java is the main programming language in current market, and Python is simpler to use than Java or C, the operation will also be simpler according to this.
6. We expect this system can accept tests by 1 June 2019, so that we can give some suggestions of improvement and may found a few bugs to repair in a timely manner.
7. After the last adjustment and testing the system we hope this system can be released by 15 June 2019. Thus, our biologist can use this system in research as soon as possible.
8. **References**

* <https://omg.readthedocs.io/en/latest/>
* <https://www.google.com/search?q=how+to+write+a+software+requirements+document&oq=how+t&aqs=chrome.0.69i59j69i60l3j69i57j69i59.14496j0j9&sourceid=chrome&ie=UTF-8>
* <file:///F:/Documents/WeChat%20Files/Whyteasamoahgyamfi/Files/oh-my-genes.pdf>

**Writer Information**

201732120135

Yeboah Martha Asamoah

[3534763826@qq.com](mailto:3534763826@qq.com)