
OMG srs Documentation

Release latest

Mar 27, 2019

Contents:

1	Contents	3
1.1	Introduction	3
1.2	The Overall Description	3
1.3	Specific Requirements	4
1.4	Non-Functional Requirements	4
1.5	Future Changes	4
2	Introduction	5
2.1	Purpose	5
2.2	Scope	5
2.3	Definitions, and Abbreviations.	5
2.4	References	5
2.5	Overview	6
3	The Overall Description	7
3.1	Product Perspective	7
3.2	User Characteristics	7
3.3	Constraints	7
3.4	Assumptions and Dependencies	8
4	Specific Requirements	9
4.1	Functional Requirements	9
5	Non-Functional Requirements	11
5.1	Response Time	11
5.2	Aesthetic Aspects	11
5.3	Confidentiality Policy	11
6	Future Changes	13
7	Indices and tables	15

- IBRAHIM MOHAMED IBRAHIM ISMAIL-201732120159
- Yussuf Mohamed mahdi-201732120173
- Class: SE 2017

1.1 Introduction

1.1.1 Purpose

1.1.2 Scope

1.1.3 Definitions and Abbreviations

1.1.4 References

1.1.5 Overview

1.2 The Overall Description

1.2.1 Product Perspective

System Interfaces

1.2.2 User Characteristics

1.2.3 Constraints

1.2.4 Assumptions and Dependencies

1.3 Specific Requirements

1.3.1 Functional Requirements

1.3.2 Input

1.3.3 OutPut

Table Analyzing

Plot Analyzing

1.4 Non-Functional Requirements

1.4.1 Response Time

1.4.2 Aesthetic Aspects

1.4.3 Confidentiality Policy

1.5 Future Changes

2.1 Purpose

This document was written to specify OMG's software requirements specifications for the developers, designers, maintainers, and stakeholders.

2.2 Scope

Gene analyzing web application. The application. Simply, takes a plain text file (.txt) as an input, containing three columns, gene ID , control sample, knock out sample. Analyzes the samples, and returns a table of the analyzation result, and a scatter plot demonstrating the deferentially expressed genes.

2.3 Definitions, and Abbreviations.

- **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.
- **Deferentially 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.

2.4 References

See **Software Requirements Specification for OMG Version 0.1** <https://oh-my-genes.readthedocs.io/en/latest/>

Written by: 1. 2.

2.5 Overview

In this document we specify the requirements that the application must follow, what to avoid, and (if any) future changes, as well as some predictions, and critical cases.

The Overall Description

The users can upload the data text document to the application by browsing their machine locating the directory, then hit the upload button. No registration needed and the application supports only English.

3.1 Product Perspective

This is a complement of the OMG web application project developed by ZJNU Software Engineering class 2016.

The application uses different APIs, such as flask, jinja, and werkzeug. Developers or maintainers dealing with this application must have a good background about those API libraries.

3.2 User Characteristics

- **Client User:** Biologists who devote themselves studying the expression of genes.
- **Site Maintainer:** Technicians who have the basic knowledge of flask and python to maintain a website.

3.3 Constraints

1. application must be compatible with the most used browsers around the world e.g firefox, chrome, and internet explorer.
2. application size must not be > **1GB**.
3. application memory usage must be **limited** under certain level so the machine runs properly.
4. application's budget for now **10K USD**. (**TBA**)
5. application's System downtime < **30 minutes per year**.
6. application must know how to handle different input values,for example if the gene equation happened to be 0/0, 1/0, 100/0, and so on. How the application going to handle them.

7. methods in the application must accept different values as arguments, for example if we have `search(arg1,arg2)` where `arg1` is the item searched for, and `arg2` is a key of the item name, let's say `arg1` is "avocado", `arg2` is "do" how the method going to quickly find the item and return the result.

3.4 Assumptions and Dependencies

- if the input is empty file, the application must return a human understandable error message.
- also if there is a missing header return an error message.
- as well as if there are missing values.
- or the length of columns are not the same.
- like any other web application if there happened to be a network outage, what are the options available to prevent any interrupts, some of the ideas are we can use multiple service providers, or keep a backup server.

4.1 Functional Requirements

4.1.1 Input

A valid submitted gene expression file has the following format:

- 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., Control-Sample) and in a treatment sample (e.g., KnockOutSample).

See the txt file pattern as following:

gene_ID	ControlSample	KnockOutDample
AT1G010	1.198558083	2.036161827
AT1G020	13.75736234	13.370796
AT1G030	0.833779536	0.203616183
AT1G040	9.58846466	7.126566394
AT1G050	23.81482799	21.10821094

4.1.2 Output

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

Table Results

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

gene_ID	ControlDample	KnockOutSample	Log_2[FC]
AT1G06463	68.78681169	182.3043556	3.38

Scatter Plot

The scatter plot displays differentially expressed genes. The X-axis is Control, and Y-axis is KnockOut. Replace 'Control' and 'KnockOut' with appropriate column names in the provided uploaded file. The up-regulated genes are shown in red dots, and down-regulated genes are shown in blue. When the mouse cursor points on one of the genes in the scatter plot, it must show the gene ID, for the sake of recognition and convenient.

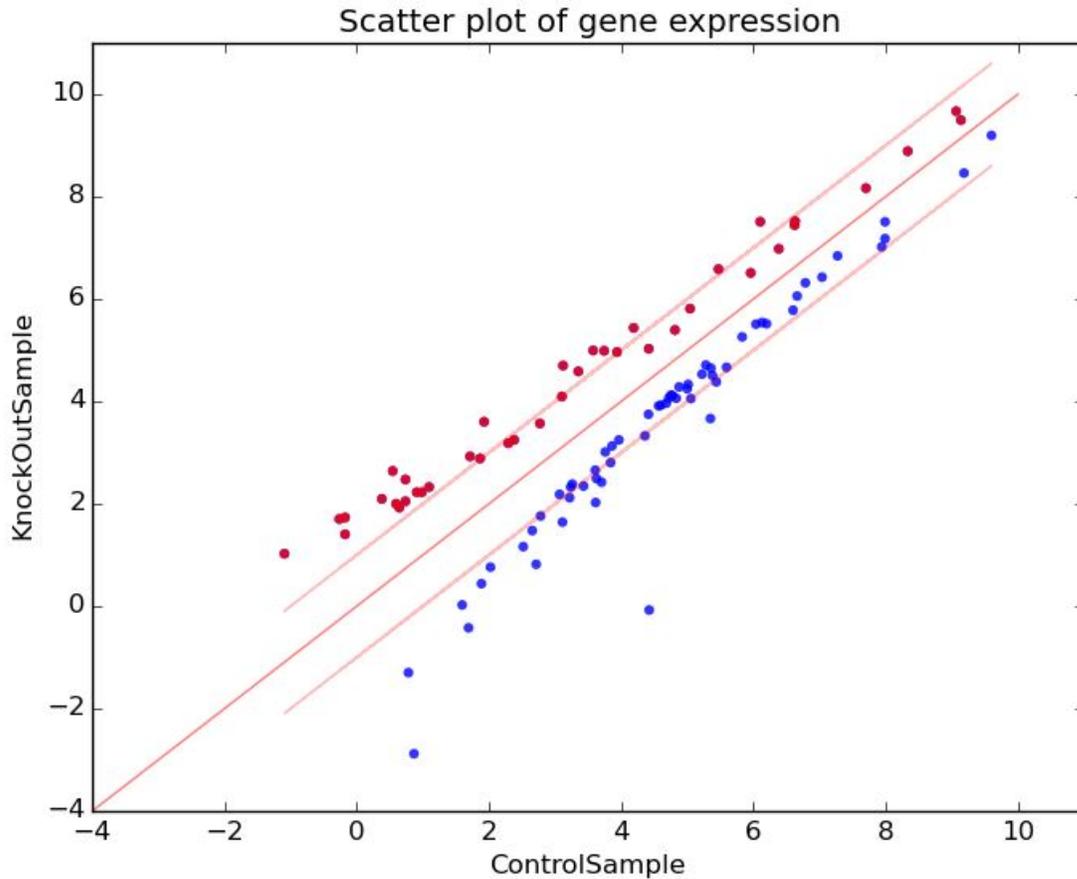


Fig. 1: figure 1.0: Scatter Plot expressing the gene levels.

Non-Functional Requirements

5.1 Response Time

Since this web application is used for biologists all around the world, the server may have to stand a heavy load of requests. Thus may cause some problems such as the system will go down for several times or lose response to its clients. On the other hand, if users have a big amount of data to analyze, the time complexity must be taken into consideration. For the optimization aspects, the response time of the application depends on the amount of data uploaded, for most cases if the data is between 10K-30K the response time must be between 10-15 seconds, and less than that if the data < 10K.

5.2 Aesthetic Aspects

As a web application to solve biological analyzing. The interface should be designed as simple as possible, and the scatter plot must show only the interesting genes levels.

5.3 Confidentiality Policy

As a public web application, the security of data must be a significant part to be considered especially for biologists, every experiment result is derived uneasily. Therefore, when users are using the application, their data must be used and stored properly. In other words, the web application must maintain the experiment data's reliability, integration and secrecy and the experimental data will not be **published** without the agreement of the owner.

CHAPTER 6

Future Changes

1. The application may support more file extensions (.txt,.doc).
2. Improvement in the home page design.
3. Download the analyzing results.
4. Feedback section for more improvement ideas.

CHAPTER 7

Indices and tables

- `genindex`
- `modindex`
- `search`