

---

# OMG Documentation

*Release 1.0*

GavinGu

Jun 27, 2018



---

## Contents

---

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Purpose . . . . .	3
1.2	Overview . . . . .	3
1.3	User Characteristics . . . . .	3
1.4	Conventions . . . . .	3
<b>2</b>	<b>Functional Requirements</b>	<b>5</b>
2.1	Function 1: I/O format checking . . . . .	5
2.2	Function 2: User Interface . . . . .	6
<b>3</b>	<b>Nonfunctional Requirements</b>	<b>7</b>
3.1	response time . . . . .	7
3.2	Aesthetic aspects . . . . .	7
3.3	Confidentiality policy . . . . .	7
<b>4</b>	<b>Constraints</b>	<b>9</b>
<b>5</b>	<b>Change cases</b>	<b>11</b>
<b>6</b>	<b>Milestones</b>	<b>13</b>



**Group** whatever we like

**member**

- Gavin Gu, Architecture Analyst / AA, e-mail: [369713635@qq.com](mailto:369713635@qq.com)
- Idyllic, Product Design / PD, e-mail: [850092018@qq.com](mailto:850092018@qq.com)

**date** 2018/05/03 14:18



### 1.1 Purpose

Comparing two samples, it's quite useful to analyse the different performances between the same genes automatically and quickly.

### 1.2 Overview

a web applicaiton based on Flask framework for anlysing the difference between genes.

### 1.3 User Characteristics

biologists and some researchers or students in related area.

### 1.4 Conventions

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.





---

## Functional Requirements

---

### 2.1 Function 1: I/O format checking

A valid submitted gene expression file has the following format. It is a 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).

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

If an invalid gene expression is given, the web application returns a page informing the user to provide the correct format.

## 2.2 Function 2: User Interface

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

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.

---

### Nonfunctional Requirements

---

#### **3.1 response time**

In 5sec, we will give back the results.

#### **3.2 Aesthetic aspects**

Simple is best.

#### **3.3 Confidentiality policy**

We will NEVER use the sotred results to seek any benefits. And it's still ok to get the result without sign in.



---

### Constraints

---

There are some constraints in this system as following:

1. According to the market shares of all kinds of browser and the compatibility of Internet Explorer, we decide that this software is accessible through Chrome, Firefox and Safari.
2. 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 less than 1GB.
3. The most important is the Budget. The less funds we use, the better we hope. But less budget does not mean just complete what we want this software can do, that is to say we expect using the least funds to use the best technology and realize the most functions. At last, we expect that the Budget ought to be limited in 10,000 USD.
4. As we need use this software to help our research the whole year, and we wish that the software makes errors or shut down accidentally much less, our team want the downtime of the system is less than 30 minutes per year



---

### 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 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, maybe 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 cash. 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 can give other functions not only to identify differentially expressed genes 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 we want this system can with its own complicated translation function. Then, we expect this system is in simple style, as we are doing research we do not need pretty beautiful interface just in simple and clear is OK.

Besides, we want this system possess privacy and security. Because our research is very confidential and authoritative, the result is all biologist





---

### Milestones

---

For submit this system, we have some requirements as following:

1. We want a SRS to get a rough idea this system by 1 May 2018.
2. We want to get SRS approved, much earlier much better, it is better before 15 May 2018.
3. 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 relative safe than a copy system.
4. 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.
5. We expect this system can accept tests by 1 June 2018, so that we can give some suggestions of improvement and may found a few bugs to repair in a timely manner.
6. After the last adjustment and testing the system, we hope this system can be released by 15 June 2018. Thus, our biologist can use this system in research as soon as possible.