

SYNTHETIC AND SYSTEMS BIOLOGY METHODS FOR APPLICATION IN GENE CIRCUITS AND MICRO- GRAVITY RELATED SPACE BIOLOGY

By
SAYAK MUKHOPADHYAY
Enrolment No: LIFE05201504008

Saha Institute of Nuclear Physics, Kolkata

A thesis submitted to the
Board of Studies in Life Sciences
In partial fulfillment of requirements
for the Degree of
DOCTOR OF PHILOSOPHY
of
HOMI BHABHA NATIONAL INSTITUTE



June, 2021

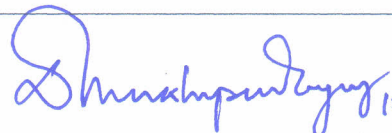
Homi Bhabha National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Sayak Mukhopadhyay entitled "Synthetic and systems biology methods for application in gene circuits and microgravity related space biology" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman - Name & Signature with date

Prof. Debashis Mukhopadhyay

 15.9.21

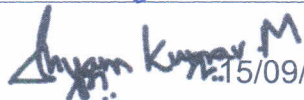
Guide / Convener - Name & Signature with date

Prof. Sangram Bagh

 15.9.21

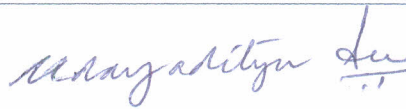
Examiner - Name & Signature with date

Prof. Shyam Kumar Masakapalli, IIT-Mandi

 15/09/2021

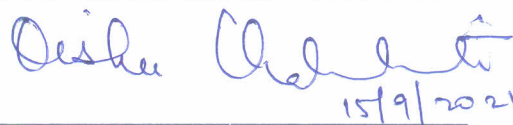
Member 1- Name & Signature with date

Prof. Udayaditya Sen

 15/9/21

Member 2- Name & Signature with date

Prof. Oishee Chakrabarti

 15/9/2021

Member 3- Name & Signature with date

Prof. Chandrima Das

 15/9/2021

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

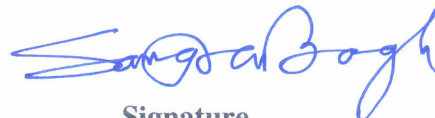
I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 15.9.21

Place: SINP, Kolkata

Signature

Co-guide (if any)



Signature

Guide

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.




Sayak Mukhopadhyay

Name & Signature
of the student

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.


Sayak Mukhopadhyay

Name & Signature
of the student

List of Publications arising from the thesis

Journal

1. A systems biology pipeline identifies new immune and disease related molecular signatures and networks in human cells during microgravity exposure., Mukhopadhyay, S., Saha, R[#], Palanisamy, A., Ghosh, M., Biswas, A., Roy, S., Pal, A., Sarkar, K., & Bagh, S., *Scientific reports*, **2016**, 6, 25975.
2. A microgravity responsive synthetic genetic device in Escherichia coli., Mukhopadhyay, S., & Bagh, S., *Biosensors & bioelectronics*, **2020**, 167, 112462.
3. Processing two environmental chemical signals with a synthetic genetic IMPLY gate, a 2-input-2-output integrated logic circuit, and a process pipeline to optimize its systems chemistry in Escherichia coli., Mukhopadhyay, S., Sarkar, K., Srivastava, R., Pal, A., & Bagh, S., *Biotechnology and bioengineering*, **2020**, 117(5), 1502–1512.

Chapters in books and lectures notes: None

Conferences

1. Oral presentation in Accelerating Biology 2016: Decoding The Deluge” At Centre For Development Of Advanced Computing (C-DAC), Pune.
2. Poster presentation at DAE-BRNS National Workshop on Particle Radiation: Effects and Application, 2016, S A Jaipuria college Kolkata.
3. Poster presentation at 19th Transcription Assembly Meeting, 2016, Bose Institute Kolkata
4. Poster presentation at International Symposium on Systems, Synthetic & Chemical Biology, SSC 2017, Bose Institute Kolkata.
5. Poster presentation at SINP International Cancer meeting 2018, SINP Kolkata.
6. Oral presentation in 19th EANA astrobiology conference, 2019, at Orleans, France.
7. Poster presentation at SSABMS 2019, SINP, Kolkata


Sayak Mukhopadhyay

Name & Signature of the student

DEDICATIONS

To My Parents

ACKNOWLEDGEMENTS

The last five and a half-year PhD journey was a transforming journey in my life. If I had pursued any other career, I might never have encountered the eureka moments. It was precisely because of these moments that I was able to persevere through the most trying times of this journey. I was extremely lucky to come across a slew of generous and kind individuals who helped make it all possible. Indeed, I can show appreciation to a much greater number of people than I can in this medium, and I apologize to anyone I missed.

First and foremost, I would like to express my gratitude to my PhD advisor Prof. Sangram Bagh, an absolute genius personality, for introducing me to Synthetic and Systems Biology and for his unwavering and selfless support during this journey. I was lucky to have met him during my Master's at Presidency University. It was in his classes, I learnt about systems thinking and emergence phenomenon, which made me appreciate the complexity of biological systems. During my PhD, he was always patient with me and guided, when in a dilemma. He inspired and motivated me to continue my intellectual and professional growth. I look up to him as a role model in academics and life as well. This thesis would not exist without him.

I want to thank my doctoral committee members, Prof. Debashis Mukhopadhyay, Prof. Udayaditya Sen, Prof. Oishee Chakrabarti and Prof. Chandrima Das and all the faculty members of the Biophysics and Structural genomics (B&SG) and Crystallography and Molecular biology (C&MB) department for their valuable advice, continuous support and encouraging words throughout my PhD. I want to thank Ex-HoD and my doctoral committee members Prof. Subrata Banerjee and Ex-Dean Prof. Abhijit Chakrabarti for their trust and support. My special thanks to Prof. Partha Saha, Dean

academic, for his support in all academic and administrative matters starting from coursework to thesis submission. I want to thank all the departmental staffs for their immense help and cooperation.

I feel fortunate to have such amazing lab members. Thank you Kathakali, Rajkamal, Deepro and Arijit Da for all the help in lab and life, technical support and mental support throughout this journey. The tea breaks, adda, memes and occasional happy hours eased the tiring long weeks at the lab. My heartfelt thanks to all my friends at B&SG dept., each one of you was very helpful and made everything about grad school enjoyable. I would miss the annual retreats, which gave fresh air and relief from grad school stress. I would like to thank my friends at the hostel, Bibhuti, Abhishek, Snehal, Satyabrata, Tulika and others, they made me feel at home away from home. I want to thank my best friend, Vivek who always pushed me academically since our college days. He always listened to my dilemmas and gave very helpful inputs. Lastly, immense love and respect for my parents, and my family members, for their love and blessings. Thank you for everything.

I would like to acknowledge the funding sources that made it possible: Department of Atomic Energy (DAE), Govt. of India; Department of science and technology, Govt. of India; Synthetic Biology Program, Biotechnology Industry Research Assistance Council (BIRAC), Govt. of India; Homi Bhabha National Institute (HBNI) International travel scheme and European Astrobiology Network Association for funding my travel to International Conference.

CONTENTS

Summary	14
List of Figures	18
List of Tables	22
Chapter 1: General Introduction.....	23
1.1 Space Biology & Microgravity: History and Objectives.	23
1.2 Facilities & tools for studying gravitational biology	30
1.2.1 Simulating microgravity on ground	31
1.3 Systems Biology.....	34
1.4 Synthetic Biology: From Bench to Space	36
Chapter 2: A systems biology pipeline identifies new immune and disease related molecular signatures and networks in human cells during microgravity exposure .	43
2.1 Introduction	43
2.1.1 Limitation of using conventional data analysis in microgravity based transcriptomic studies.....	47
2.2 Systems Biology Pipeline.....	48
2.2.1 The gene expression data	48
2.2.2 Database for predefined human molecular pathways	49
2.2.3 Identification of pathways and leading genes	51
2.2.4 Extracting gene expression patterns and networks in microgravity.....	52
2.3 Results	54

2.3.1 Molecular pathway analysis using gene set enrichment analysis (GSEA)	54
2.3.2 Novel molecular signatures related to immunity	60
2.3.3 Novel molecular signatures related to cancer	68
2.3.4 Other novel molecular signatures	69
2.3.5 Consensus non-negative matrix factorization (CNMF) clustering and protein network mapping show plausible regulatory and functional connections among genes	69
2.4 Discussion	74
2.5 Conclusion	78
2.6 Methods	79
2.6.1 Gene set enrichment analysis (GSEA)	79
2.6.2 Leading edge gene analysis	81
2.6.3 Clustering of leading-edge genes with non-negative matrix factorization consensus (NMFC)	81
2.6.4 Protein-protein associated networks of clustered genes	82
Chapter 3: A microgravity responsive synthetic genetic device in <i>Escherichia coli</i>	83
3.1 Introduction	83
3.1.1 The need of a microgravity responsive synthetic genetic device	86
3.1.2 Experimental setup	88
3.2 Results	90
3.2.1 Design and fabrication of microgravity responsive synthetic genetic device	90

3.2.2 The fabricated synthetic genetic device responds to microgravity by altering expression of a fluorescence protein.....	94
3.2.3 The basic device design is general in nature for <i>E.coli</i>	98
3.2.4 Application of the device to control native cellular properties with microgravity	102
3.3 Discussion & Conclusion.....	107
3.4 Materials and Method.....	109
3.4.1 Media, strains and growth conditions	109
3.4.2 synsrRNA and Plasmid construction.....	110
3.4.3 Working principle of Synergy HTX Multi-Mode plate reader (Biotek Instruments, USA).....	117
3.4.4 Working principle of fluorescence microscope (confocal LSM 710, Ziess)	118
3.4.5 Fluorescence Measurements and data analysis	120
3.4.6 RNA purification and cDNA synthesis	122
3.4.7 qPCR and Data analysis	122
3.4.8 Confocal microscopy.....	123
Chapter 4: Development of a process pipeline for synthetic genetic circuit fabrication and its application in creating complex genetic circuit in single cell.....	124
4.1 Introduction	124
4.1.1 Bioparts Assembly Techniques	124
4.1.2 Network brick.....	128
4.1.3 Synthetic Genetic IMPLY gate.....	129

4.2 Results and Discussion.....	131
4.2.1 Network brick base plasmid design and construction.....	131
4.2.2 Design, fabrication and characterization of NOR logic gate	140
4.2.3 Design and fabrication of IMPLY logic for sensing and processing two extracellular environmental chemicals.....	148
4.2.4 Characterization and optimization of IMPLY gate.....	149
4.2.5 Mathematical modeling, digitality, simulation and experimental validation of IMPLY device.	153
4.2.6 Fabrication and optimization of a 2-input-2output integrated circuit that integrates IMPLY and a NOT gate	159
4.2.7 Mathematical modeling, digitality, simulation and experimental validation of 2input-2output-integrated device.	162
4.2.8 Integrating synthetic circuits with native cellular process of <i>E.coli</i>	166
4.2.9 PCR based Network Brick for fabrication and characterization of Exchanger circuit.....	169
4.3 Discussion and Conclusion	172
4.4 Materials and Methods	174
4.4.1 Promoters, primers, genes, plasmids, and bacterial cell strains	174
4.4.2 Cell growth for genetic constructs characterization.....	187
4.4.3 Fluorescence and optical density (OD) measurements	187
4.4.4 Working principle of Synergy HTX Multi-Mode Reader (Biotek Instruments, USA).....	188
4.4.5 Calculating translation initiation rate in RBS Library Calculator.....	188

4.4.6 Working principle of fluorescence microscope (confocal LSM 710, Ziess)	
.....	189
4.4.7 Microscopy	190
Conclusions and future directions	191
Bibliography.....	195
APPENDIX	223

Summary

Humans must migrate past low earth orbit into deep space in order to evolve into a multi-planet civilization. The major difficulties that accompany this mission include dealing with microgravity, cosmic radiation danger, sustaining a closed life support system, and resource availability. Alternative methods such as systems and synthetic biology are being discussed to address these issues.

This thesis has three objectives: -

1. Developing a systems biology pipeline to understand the pathway level and network level picture of human cell under microgravity, which cannot be done by conventional data analysis method.

Microgravity is a prominent health hazard for astronauts, yet we understand little about its effect at the molecular systems level. We incorporated a suite of systems biology techniques and databases in this analysis and analysed over 8000 molecular pathways using previously published global gene expression datasets of human cells in microgravity. Hundreds of new pathways have been identified with statistical confidence for each dataset and despite the difference in cell types and experiments, around 100 of the new pathways are appeared common across the datasets. They are related to reduced inflammation, autoimmunity, diabetes and asthma. We have identified downregulation of NfκB pathway via Notch1 signalling as new pathway for reduced immunity in microgravity. Induction of few cancer types including liver cancer and leukaemia and increased drug response to cancer in microgravity are also found. Increase in olfactory signal transduction is also identified. Genes, based on their expression pattern, are clustered and mathematically stable clusters are identified. The network mapping of genes within a cluster indicates the plausible functional connections in microgravity. This

pipeline gives a new systems level picture of human cells under microgravity, generates testable hypothesis and may help estimating risk and developing medicine for space missions.

2. To create a microgravity responsive synthetic genetic device in *Escherichia coli*, which can integrate microgravity as a physical signal within a biochemical process in a human designed way.

Bioengineering solutions to human space travel must consider microgravity as an important component. Thus, one of the fundamental challenges of space bioengineering is to create cellular microgravity responsive device, which integrate microgravity as a signal within biochemical and cellular processes. Here, we designed, fabricated and characterized the first biochemical and cellular microgravity responsive device using an engineered genetic circuit in *E.coli*, which responded to microgravity by changing the expression of a target enhanced green fluorescent gene (EGFP). Our device design was based on the deregulation of Hfq protein in *E.coli* in microgravity, which was translated through Hfq mediated silencing of EGFP by anti-EGFP synthetic small regulatory RNAs. This resulted a reduced silencing (~28 times) of the EGFP in microgravity. We demonstrated that the basic design of the device is universal in nature for *E.coli*, by creating multiple successful devices, where target genes (EGFP, TdTomato, and FtsZ) and the promoters (inducible and constitutive) were altered. Further, we applied this device to control the cell division process by microgravity. We targeted the cell division regulator FtsZ in *E.coli*, which resulted in elongated *E.coli* cells in normal gravity. This elongated cell length got rescued to a regular length by applying microgravity.

The work showed for the first time, a way to integrate microgravity as a physical signal within biochemical processes of a living cell in a human designed way and thus, may have significance in space bioengineering and synthetic biology

3. Development of a process pipeline for synthetic genetic circuit fabrication and its application in creating complex genetic circuit in single cell.

Effective assembly of genetic elements is needed for the reliable construction of complex genetic circuits, for space synthetic biology, and for synthetic biology in general. Despite the existence of numerous DNA assembly methods, it remains a bottleneck for the majority of small laboratories such as ours. Synthetic genetic devices can perform molecular computation in living bacteria, which may sense more than one environmental chemical signal, perform complex signal processing in a human-designed way, and respond in a logical manner. IMPLY is one of the four fundamental logic functions and unlike others, it is an “IF-THEN” constraint-based logic. By adopting physical hierarchy of electronics in the realm of in-cell systems chemistry, a full-spectrum transcriptional cascaded synthetic genetic IMPLY gate, which senses and integrates two environmental chemical signals, is designed, fabricated, and optimized in a single *Escherichia coli* cell. This IMPLY gate is successfully integrated into a 2-input-2-output integrated logic circuit and showed higher signal-decoding efficiency. Further, we showed simple application of those devices by integrating them with an inherent cellular process, where we controlled the cell morphology and color in a logical manner. To fabricate and optimize the genetic devices, a new process pipeline named NETWORK Brick is developed. This pipeline allows fast parallel kinetic optimization and reduction in the unwanted kinetic influence of one DNA module over another. A mathematical model is developed and it shows that response of the genetic devices is digital-like and

are mathematically predictable. This single-cell IMPLY gate provides the fundamental constraint-based logic and completes the in-cell molecular logic processing toolbox. The work has significance in the smart biosensor, artificial in-cell molecular computation, synthetic biology, and microbiorobotics.

In a nutshell, in this thesis work we developed systems and synthetic biology tools to contribute in two specific areas i.e., space biology and gene circuits. The second chapter addresses one major limitation of using conventional differential gene expression analysis for a weak perturbation like microgravity, which hardly reveals altered pathways with statistical significance. Our analytical pipeline explained several unexplained phenomena and developed several testable hypotheses for future. The unsupervised machine learning algorithm identified some common functional and regulatory connection among different clusters of genes and hints toward an underlying subset of genes that are affected in microgravity. The third chapter showed for the first time that how microgravity can be integrated with native cellular processes in *E.coli* as a physical signal using a synthetic gene circuit and is very important towards future of space bio-engineering. Finally, the fourth chapter deals with synthetic biology in context of gene circuits and presents an efficient assembly method, which allows bidirectional assembly and the feature of parallel optimization of gene circuits. This work is important as smaller laboratories lack automation facilities; therefore, such technique can be used for gene circuit fabrication in synthetic biology research. We demonstrated the application of this technique by developing, characterizing and optimizing complex genetic gates and assembled 5 cascaded genetic cassettes in a single plasmid in *E.coli*.

List of Figures

Figure 1.1 Miss Sam, a rhesus monkey, is seen enclosed in a Mercury fiberglass contour couch model. She is being loaded into a container for the Mercury Capsule's Little Joe 1B suborbital test flight. Photo credit: NASA ⁵ (Permission for non-commercial use.)	24
Figure 1.2 Photographic documentation of <i>Arabidopsis</i> seedlings from the Petri Plants-2 experiment in the Destiny U.S. Laboratory aboard the International Space Station (ISS).	28
Figure 1.3 . Fluid shift towards upper body leading to puffy face.	30
Figure 1.4 Four different classes of ground-based microgravity simulators for biological samples.	34
Figure 1.5 The design principles of synthetic biology are borrowed from engineering ⁴¹ .	37
Figure 2.1 GSEA analysis of human cell during microgravity exposure.	52
Figure 2.2 Flowchart of the systems biology pipeline.	53
Figure 2.3 Number of statistically significant ($p < 0.001$, $q < 0.001$) altered molecular pathways identified by GSEA for each experiment.	55
Figure 2.4 GSEA identified greater number of pathways compared to DGE analysis.	56
Figure 2.5 Venn diagrams representing the number of common altered pathways in microgravity.	60
Figure 2.6 Overlapping leading edge genes between three pathways, which affect the regulation of Nf- κ B pathway.	63

Figure 2.7) Maximum number of mathematically stable clusters plausible for other experiments. Clusters with upregulated and downregulated genes are shown in black and red bars respectively.	70
Figure 2.8 Cophenetic coefficient as function of number of clusters for all the experiments.	71
Figure 2.9 Protein-protein association (PPA) networks.	72
Figure 3.1 Potential application of synthetic biology in space.	85
Figure 3.2 Projected photosynthetic bioreactors for food production.....	86
Figure 3.3 Experimental setup for <i>E.coli</i> culture in simulated microgravity condition.	88
Figure 3.4 Configuration of the HARV's	90
Figure 3.5 Relative Hfq mRNA expression level in microgravity (μ G) and 1G in <i>E.coli</i> cell strain DH5 α Z1.....	92
Figure 3.6 The design of the microgravity sensor circuit with EGFP.	93
Figure 3.7 Anti-EGFP synthetic small regulatory RNA encoding synthetic DNA sequence.	94
Figure 3.8 The experimental behavior of microgravity sensor.	95
Figure 3.9 Control experiment.	96
Figure 3.10 Anti-tdTomato synthetic small regulatory RNA encoding synthetic DNA sequence.	98
Figure 3.11 The design of the Td-tomato based microgravity sensor circuit.....	99
Figure 3.12 Experimental behavior of the TdTomato based microgravity sensor.	100
Figure 3.13 Control experiment.	101

Figure 3.14 AntiFtsz synthetic small regulatory RNA encoding synthetic DNA sequence.	103
Figure 3.15 SynsrRNA based microgravity sensor for integrating microgravity with native FtsZ protein in Dh5 α Z1 <i>E.coli</i>	103
Figure 3.16 Merged fluorescence and DIC image of <i>E. coli</i> DH5 α Z1 cells expressing SynsrRNA against FtsZ from microgravity responsive synthetic genetic device.	105
Figure 3.17 Control experiments for SynsrRNA against FtsZ from microgravity responsive synthetic genetic device.	106
Figure 3.18 DIC and Fluorescence image of the merged images a) in Figure 3.16 and b) in Figure 3.17 (control experiments).	107
Figure 3.19 Plasmid maps for the plasmids developed for this study.	115
Figure 3.20 Schematic diagram of filter-based fluorescence detection in multimode reader.	118
Figure 3.21 Schematic diagram of confocal microscope used in these experiments.	120
Figure 4.1 Biobrick assembly method.	126
Figure 4.2 Gibson isothermal assembly	128
Figure 4.3 pZ vector system for cloning and expression ¹⁵⁶	132
Figure 4.4. Network Brick, the new biochemical process pipeline.	134
Figure 4.5 Process pipeline for assembling two genetic cassettes in a single direction.	135
Figure 4.6 Schematic of RCEV and process pipeline for assembling two genetic cassettes in two opposite directions.	136
Figure 4.7. Design and testing of molecular NOR gate.	141

Figure 4.8. Design and testing of molecular IMPLY gate.....	149
Figure 4.9 Comparison of EGFP expression from P_{Lataa} and P_R promoters with identical RBS and identical plasmid vector.	150
Figure 4.10 IMPLY gate logic behavior with various RBSs (i1-i7).....	152
Figure 4.11. Dose-response curves.....	158
Figure 4.12. Simulated (A) and experimental (132 data points) (B) behavior of IMPLY gate. Plugging the parameter values in the Equation 4.4 we performed a complete simulation by varying the both IPTG and ATC concentration for total 4902 data points. We compared it with experimental behavior of IMPLY gate across different concentrations of IPTG and ATC (132 data points).	159
Figure 4.13. Design and testing of 2-input-2-output integrated logic circuit.....	160
Figure 4.14. Simulated and experimental behaviour of 2-input-2-output integrated logic circuit. Simulated behavior of A) output 1 (EGFP) and B) output 2 (TdTomato). Experimental behaviour (121 data points each) of C) output 1 (EGFP) and D) output 2 (TdTomato) as a function of 2 input chemicals in the full integrated circuit inside cell.	165
Figure 4.15. A) The IMPLY circuit with phenotypic output. B) Experimental behaviour of the engineered cell. The red arrows are pointing towards elongated cells.	167
Figure 4.16. A) The 2-input-2-output circuit with phenotypic outputs. B) Experimental behaviour of the engineered cell. The red arrows are pointing towards elongated cells.	168
Figure 4.17 Modified PCR based Network Brick vector and assembly pipeline.	170
Figure 4.18 Design and experimental characterization of Exchanger circuit constructed using PCR based Network Brick.....	171

List of Tables

Table 1.1 Several organisms which have been studied in space. Reproduced with permission from (Clément, 2006).	25
Table 1.2 Summary of opportunities and challenges for space synthetic biology.	40
Table 2.1 Brief Description of categories of gene sets used in this study	50
Table 2.2 Altered KEGG pathways identified by GSEA for dataset E-GEOD-43582.57	
Table 2.3 Microgravity induced functions are supported by multiple molecular pathways.....	65
Table 3.1 List of all the synthetic oligos used in this study	113
Table 3.2 List of all the plasmids used in this study.....	116
Table 4.1. Comparison between Network Brick and Biobrick for construction & high throughput optimization of hypothetical constructs in three cases.	137
Table 4.2. Construction steps of gene circuits using Network Brick	142
Table 4.3. RBS used in the study and their estimated translation rates from RBS calculator	151
Table 4.4. Parameter values generated by curve fitting.	164
Table 4.5. Predefined set of primers required for Network PCR Brick	171
Table 4.6 List of plasmids	176
Table 4.7 List of primers	184

Chapter 1: General Introduction

1.1 Space Biology & Microgravity: History and Objectives.

For decades, biological research has been performed in space. The first attempts with biological payloads in the United States date back to 1946, when a series of fungal spores was launched in a pioneering balloon flight from Alamogordo, New Mexico (Souza, Hogan and Ballard, 1965). The aim of biological research in the early years of the space age was to assess the ability of living organisms to withstand space flight. After it was established that animals and humans could endure microgravity, cosmic radiation, and the rigors of launch and re-entry, researchers turned their attention to the biological changes that occur during and after space travel. The outpost of humanity in space i.e. the International Space Station (ISS) is a pressurized module in orbit, is characterized by microgravity, cosmic radiation risk and absence of circadian rhythm. In order to ensure the safety of astronauts in these unique settings, it is crucial to understand the effect of space environment on living things.

Historically, both the US and Soviet/Russian space programs relied heavily on animal experiments to amass medical expertise and validate engineering design principles necessary to sustain human spaceflight. In addition, Animal replacements such as dogs, monkeys (Figure 1.1) and rats were used to assess the space environment's suitability for human presence, including launch vehicles, radiation and microgravity exposure, life support systems, and rescue procedures (*Laboratory animals in space life sciences research.*, no date). Initially, small insects and plant seed were sent into space to assess the effect of cosmic radiation. These was followed by test flights with

General Introduction

mammals and primates for evaluating the impact of acceleration and microgravity. After establishing that complex biological species could live in orbit, humans travelled into space, bringing along animals as experimental subjects. Once humans established a permanent outpost in space, biological specimens are being flown on manned and unmanned flights to further investigate the long-term physiological consequences of spaceflight, including microgravity and cosmic radiation (Souza, Hogan and Ballard, 1965; Clément, 2006). A list of some organisms which have been studied in space is highlighted in Table 1.1. Initially, several of the most important biological effects of spaceflight have been uncovered by animal studies for e.g. loss of bone density, cardiovascular deconditioning, muscle atrophy, altered metabolism and immune suppression has been studied in rat and rhesus monkey (Figure 1.1); altered development in chicken, rat and frog across several space mission ((*No Title*), no date a).

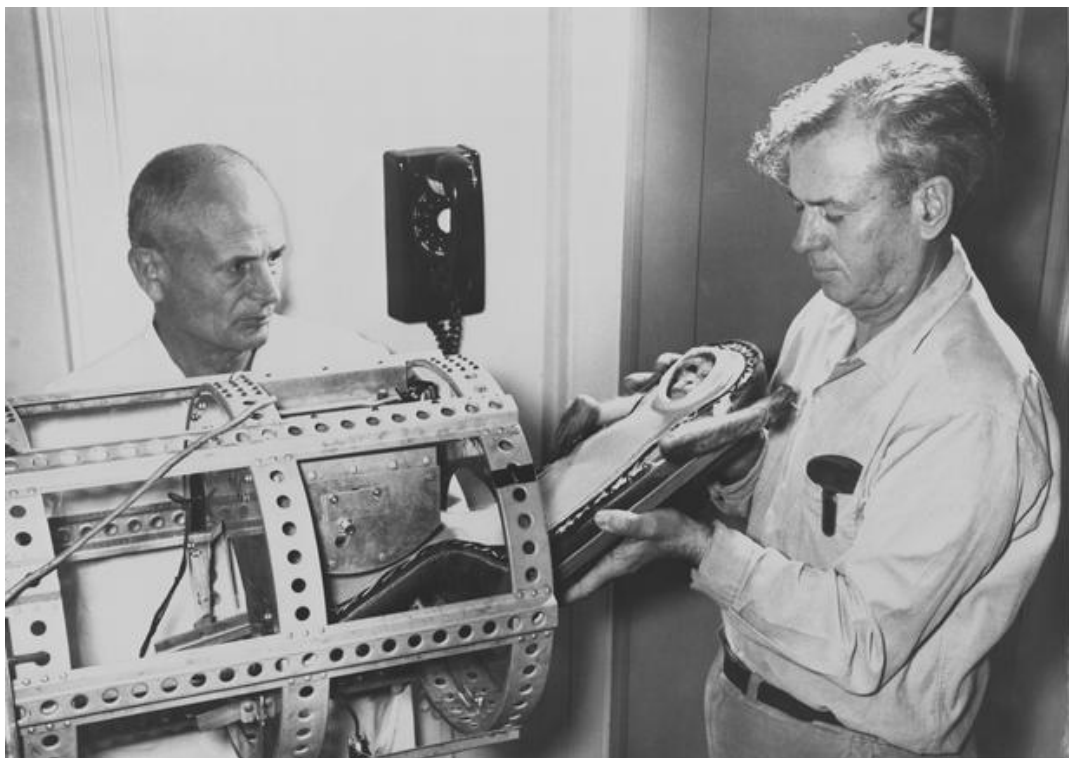


Figure 1.1 Miss Sam, a rhesus monkey, is seen enclosed in a Mercury fiberglass contour couch model. She is being loaded into a container for the Mercury Capsule's Little Joe 1B suborbital test flight. Photo credit: NASA(*Rhesus Monkey - Miss*

General Introduction

Sam - Fiberglass Couch - Little Joe (LJ)-1B Flight - Prep | NASA Image and Video Library, no date) (Permission for non-commercial use.)

There are multiple reasons which make humans unsuitable as test subjects. Firstly, studies like developmental biology and genetics cannot be performed ethically on humans. Second, utilizing tiny species is practical. In comparison to humans, the animals chosen for space research are lightweight and relatively simple to manage, sustain and multigenerational studies can be performed in single missions. Simple model organisms are now employed by investigators to study the complex effects of space on living things. Model organisms are suitable for multigenerational research due to their massive sampling sizes and rapid replication times. Additionally, the majority of often used model organisms have fully sequenced genomes.

Table 1.1 Several organisms which have been studied in space. Reproduced with permission from (Clément, 2006).

Bacteria	Invertebrates
<i>Aeromonas proteolytica</i>	<i>Acheta domesticus</i> (Cricket)
<i>Bacillus mycoides</i>	<i>Araneus diadematus</i> (Spider)
<i>Bacillus subtilis</i>	<i>Biomphalaria glabrata</i> (Snail)
<i>Bacillus thuringiensis</i>	<i>Caenorhabditis elegans</i> (Nematode)
<i>Burkholderia cepacia</i>	<i>Cynops pyrrhogaster</i> (Newt)
<i>Chaetomium globosum</i>	<i>Drosophila melanogaster</i> (Fruit fly)
<i>Deinococcus radiodurans</i>	<i>Habrobracon juglandis</i> (Wasp)
<i>Escherichia coli</i>	<i>Manduca sexta</i> (Tobacco hornworm)
<i>Nematospiroides dubius</i>	<i>Pelomyxa carolinensis</i> (Amoeba)
<i>Rhodotorula rubra</i>	<i>Poethetria dispar</i> (Gypsy moth)
<i>Salmonella typhimurium</i>	<i>Scorpio maurus</i> (Scorpion)
<i>Trichophyton Terrestre</i>	<i>Tribolium confusum</i> (Beetle)
	<i>Trigonoscelis gigas</i> (Beetle)
Vertebrates–Aquatic species	Vertebrates–Birds
<i>Arbacia punctulata</i> (Sea urchin)	<i>Coturnix coturnix</i> (Quail)

<i>Aurelia aurita</i> (Jellyfish)	<i>Gallus gallus</i> (Chicken)
<i>Fundulus heteroclitus</i> (Killifish)	Vertebrates–Mammals
<i>Lytechinus pictus</i> (Sea urchin)	<i>Canis familiaris</i> (Dog)
<i>Opsanus tau</i> (Toadfish)	<i>Felix maniculata</i> (Cat)
<i>Oreochromis mossambicus</i> (Cichlid fish)	<i>Macaca mulatta</i> (Rhesus monkey)
<i>Oryzias latipes</i> (Medaka fish)	<i>Macaca nemestrina</i> (Macaque monkey)
<i>Rana catesbeiana</i> (Bullfrog)	<i>Mus musculus</i> (Mouse)
<i>Rana pipiens</i> (Frog)	<i>Oryctolagus cuniculus</i> (Rabbit)
<i>Strongylocentrotus purpuratus</i> (Sea urchin)	<i>Pan troglodytes</i> (Chimpanzee)
<i>Xenopus laevis</i> (Frog)	<i>Perognathus longimembris</i> (Pocket mouse)
<i>Xenopus laevis</i> Daudin (South African toad)	<i>Rattus norvegicus</i> (Rat)
<i>Xiphophorus helleri</i> (Swordtail fish)	<i>Saimiri sciureus</i> (Squirrel monkey)
	<i>Testudo horsfieldi</i> (Tortoise)
Plants	
<i>Aesculus hippocastanum</i> L. (Horse chestnut)	<i>Haplopappus gracilis</i> (Haplopappus)
<i>Arabidopsis thaliana</i> (Thale cress)	<i>Helianthus annuus</i> L. (Sunflower)
<i>Avena sativa</i> (Oat)	<i>Hemerocallis</i> (Daylily)
<i>Brassica rapa</i> (Field mustard)	<i>Lepidium sativum</i> (Garden cress)
<i>Capsicum annuum</i> (Ornamental pepper)	<i>Linum usitatissimum</i> (Flax)
<i>Ceratodon</i> (Moss)	<i>Lycopersicon esculentum</i> (Tomato)
<i>Ceratopteris</i> (Fern)	<i>Neurospora crassa</i> (Fungus)
<i>Ceratophyllum demersum</i> (Hornweed)	<i>Nicotiana tabacum</i> (Tobacco)
<i>Cucumis sativus</i> (Cucumber)	<i>Oryza sativa</i> (Rice)
<i>Dactylis glomerata</i> L. (Orchard grass)	<i>Physarum polycephalum</i> (Slime mold)
<i>Daucus carota</i> (Carrot)	<i>Pseudotsuga menziesii</i> (Douglas fir)
<i>Digitalis lanata</i> (Foxglove)	<i>Pseudotsuga taeda</i> (Loblolly pine)
<i>Digitalis purpurea</i> L. (Foxglove)	<i>Saccharomyces cerevisiae</i> (Yeast)
<i>Elodea</i> (Waterweed)	<i>Tradescantia</i> (Spiderwort)
<i>Flammulina velutipes</i> , Agaricales (Fungus)	<i>Triticum aestivum</i> (Wheat)
<i>Glycine max</i> (Soybean)	<i>Triticum vulgare</i> (Wheat)
	<i>Vigna radiata</i> (Mung bean)
	<i>Zea mays</i> (Corn)

General Introduction

The focus is not limited to animals, for long-term sustained human existence in space we must be capable of consistently growing and reproducing different plant species over several generations for food and regulated environmental life support systems. In 1997, seeds collected from plants cultivated in microgravity were successfully germinated in orbit, marking a significant milestone in plant space biology. This pioneering “seed-to-seed” development of plants in space demonstrated that plants do not need gravity to reproduce (MERKIS and LAURINAVICHYUS, 1983). Multiple studies provided ample evidence in this direction (Kuang *et al.*, 2000; Sychev *et al.*, 2007; Link, Busse and Stankovic, 2014). Apart from plant reproduction experiments, vegetative propagation of plants has been successfully carried out in space (*Influence of Microgravity Environment on Root Growth, Soluble Sugars, and Starch Concentration of Sweetpotato Stem Cuttings in: Journal of the American Society for Horticultural Science Volume 133 Issue 3 (2008)*, no date; Cook and Croxdale, 2003). Some of the plant studied in space are listed in Table 1.1. Most preferred plant not only in space biology (Figure 1.2), but other research fields is *Arabidopsis thaliana*, a model organism. It has a small genome size, short life cycle (6 weeks from germination to seed maturation), limited space requirement for cultivation, produces decent number of seeds and large number of mutants are available for researchers(Ruyters and Braun, 2014)(Clément, 2006).

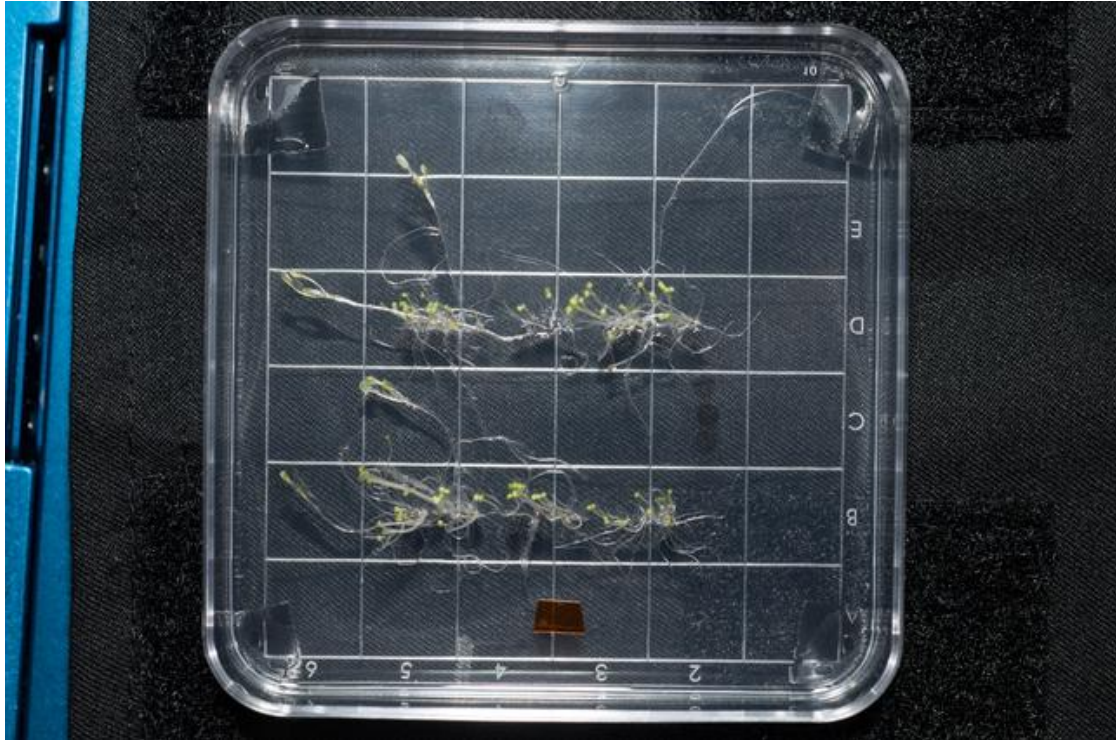


Figure 1.2 Photographic documentation of *Arabidopsis* seedlings from the Petri Plants-2 experiment in the Destiny U.S. Laboratory aboard the International Space Station (ISS).

Photo credit: NASA(*Petri Plants-2 Experiment Plate Final Survey* | *NASA Image and Video Library*, no date) (Permission for non-commercial use.)

The human body has long been known to undergo adaptive or maladaptive changes when transitioning from a terrestrial to these unique space environments. As they adapt to life in space, astronauts experience nausea, disorientation, upward shift in body fluids (Figure 1.3), disturbance of their sleep schedule, immune response depression, and other symptoms. The majority of these conditions disappear once they return to Earth, some immediately, some gradually. Sometimes conditions, such as bone mass loss (Lang *et al.*, 2006), may take years to heal, and for some astronauts, the damage may be permanent. These changes appear to occur at the cellular level, evident from changes in calcium balance. These conditions are reviewed in the introduction section of the second chapter.

The interest in space biology research stems from two grand objectives. First, we must consider and develop steps to counteract the most adverse effects of spaceflight on biological systems in order for humans to participate in long-term space travel (Figure 1.3). Second, understanding how organisms respond to microgravity allows one better to understand fundamental biological processes (Figure 1.3). Life on Earth has evolved in a 1-g environment. The influence of gravity is not well understood, except that there is clearly a biological response to gravity in the structure and functioning of living organisms. The effects of gravity are relatively obvious at total organism or system levels, as observed in astronauts, but they are not immediately apparent at the cellular level. The information gained from this research can be used to improve human health and the quality of life (Clément, 2006).

Finally, understanding how different plants, animals, and microorganisms respond and interact in closed environments is critical for designing advanced life support systems for long-term missions. Spacecraft can be replenished in regular intervals in Earth's orbit. On the other hand, long-term expeditions to other planets and celestial bodies would require the development of self-sustaining ecosystems. Plants and microbes with bio-regenerative properties would be critical components in such systems, performing functions such as food production, CO₂ absorption, and waste recycling. In addition to aiding space exploration, research into these properties can have a wide range of applications on Earth. Thus, as in any other research discipline space biology has both basic and applied aspects. Further details of the field can be found in (Clément, 2006).

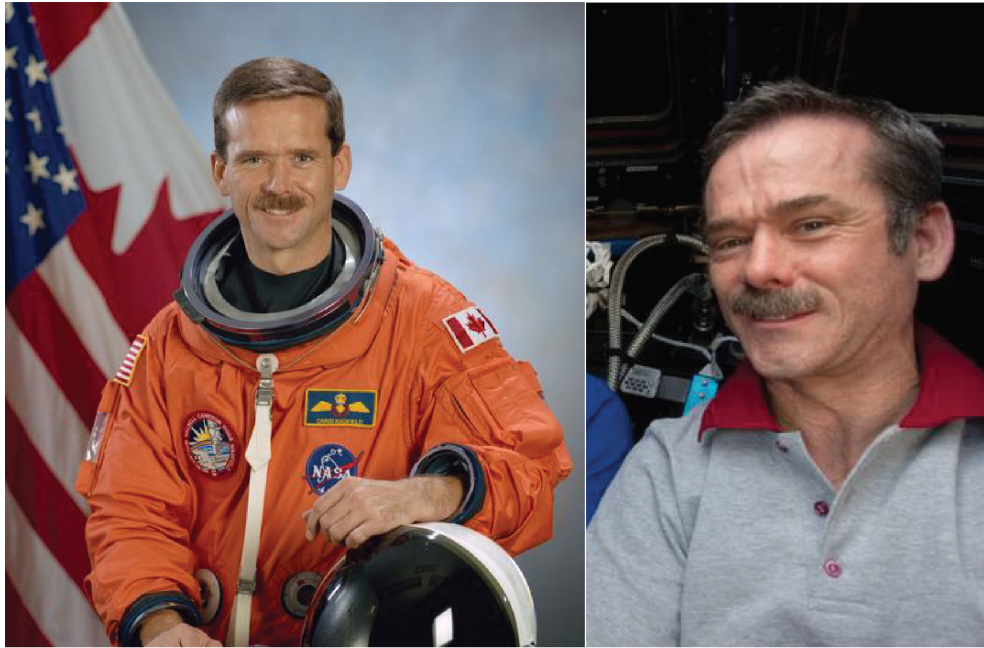


Figure 1.3 . Fluid shift towards upper body leading to puffy face.

(Left) Before flight, (Right) in-flight. Photo credit: NASA(*Petri Plants-2 Experiment Plate Final Survey* | *NASA Image and Video Library*, no date) (Permission for non-commercial use.)

1.2 Facilities & tools for studying gravitational biology

Microgravity, or more precisely micro-weightlessness, occurs in an orbital spacecraft in free fall, that is, without any external force acting on it other than gravitational forces. Thus, near-weightlessness on the range of 10^{-4} to 10^{-6} g is encountered in today's spacecraft and is often referred to as microgravity. The facilities for attaining microgravity are as follows: -

Drop towers: - The most cost-effective and readily accessible method of providing microgravity is to use drop towers or shafts up to 100 meters in height, which may use vacuum pumps and allow an experiment capsule to fall freely for a brief period before decelerating, giving 2-5 seconds of microgravity (*ZARM: General Information*, no date). These are good for material sciences and physics research (Clément, 2006).

Parabolic flights: - An aircraft is maneuvered so that it enters a free-fall ballistic trajectory, i.e., a parabola, which lasts about 20 seconds and achieves weightlessness. Each mission typically consists of between 30 and 40 parabolas, allowing for the replication of studies. On these trips, the experimenter may carry and operate rather substantial pieces of equipment. Parabolic flights have been widely utilized to study human and animal physiology and gravitational biology under low gravity conditions (Clément, 2006).

Sounding rockets: - These are suborbital rockets transporting cargo above the Earth's atmosphere for up to 15 minutes without putting it into orbit around the Earth. Typically, such rockets reach an altitude of 250–350 kilometers, at which point the payload is detached and stabilized free-falls to a parachute landing (Clément, 2006). Such facilities have been used to investigate the processes behind gravity sensing in a variety of plants and animals (Clément, 2006).

International space station (ISS): - The ISS is the world's largest and longest continuously inhabited space station, occupied by humans since November 2000. It is the most expensive microgravity platform for conducting experiments but because it provides months or even years of experimental time, crew operation, and download capabilities (Böhmer and Schleiff, 2019).

However not all experiments are possible in space, therefore ground-based microgravity simulations are used for routine microgravity research.

1.2.1 Simulating microgravity on ground

Since living organisms are adapted to the ubiquitous force of gravity on earth, it became essential to study how life responds to microgravity with the start of crewed space missions. Using spaceflights for conducting experiment is not feasible, due to high cost and

General Introduction

limited flights. This highlights the critical importance of ground-based facilities (GBFs) for establishing baselines and conducting rigorous research of the biological system in order to address gravity-related concerns prior to conducting space experiments. They also allow stand-alone studies, providing additional and cost-effective platforms for gravitational study. The numerous microgravity simulators that gravitational biologists commonly use are as follows: -

Rotating Wall Vessel: Rotating wall vessels (RWVs) or rotating bioreactors (Rotating Cell Culture System, originally created by NASA) comprises of a horizontally rotating cylindrical culture vessel (Figure 1.4 A), perpendicular to the gravitational vector, with a central co-axial oxygenator. When the vessel is filled completely with culture media, devoid of any bubbles and rotated, the fluid spins along a horizontal axis like a rigid body resulting in laminar flow and reduced shear force. Thus, the biological sample is maintained in a state of suspension. The vessel's rotation generates an upward hydrodynamic drag force in opposition to the downward gravity force. As gravitational and centrifugal forces are balanced, a microgravity-like culture environment is formed inside the cylinder.

Clinostat: A clinostat is a mechanism that rotates samples to avoid the perception of the vector of gravity by the biological system. Clinostats may have a single horizontal rotation axis (2D-clinostat) (Figure 1.4 A) or two rotation axes (3D-clinostat). There are several configurations depending on the number of rotation axes, the rotation speed, and the rotation direction (Briegleb, 1992; Klaus, Todd and Schatz, 1998). Classical Clinostats were initially used in plant science to rotate objects steadily (1–10 revolutions per minute). Gravitropic responses were not observed in seedlings or small plants rotated slowly in the 2-D clinostat axis. However, morphological experiments later

General Introduction

demonstrated that slow clinorotation (1–2 rpm) causes ultrastructural disruptions not observed during spaceflight (Hensel and Sievers, 1980). These findings indicate that the slow rotation not only blocked a gravity-induced growth response, but also triggered omnilateral mechanical stress in some sensitive plant tissues. Briegleb (Briegleb, 1992) pioneered the idea of a fast spinning clinostat for the purpose of achieving "functional weightlessness" for small objects, mostly single cells. It is assumed that fast and constant rotation mechanically prevents sedimentation by changing the direction of the gravity vector in a continuous and constant manner.

Random Positioning Machine: The Random Positioning Machine (RPM) works on the principle of random rotation of biological samples along two axes (Mesland *et al.*, 1996). This rotation continuously reorients the gravity vector acting on the samples. A microgravity-like condition is then generated by rotating the samples in such a way that the gravity vector's trajectory points in all directions (from the samples' perspective). As the gravity vector is averaged over time, it mathematically converges to zero. To allow rotation of the samples in either direction, RPM systems typically employ two independently rotating gimbal-mounted frames (Figure 1.4 C).

Magnetic Levitation: Magnetic levitation is now well developed as a viable means of simulating microgravity environments and therefore one of the Earth-based solutions to space-based experiments. The fundamental concept is that a magnetic force is used to compensate for the gravitational force. Magnetic levitation systems suspend cells in fluid by applying a strong magnetic field to cells conjugated to ferromagnetic particles (Souza *et al.*, 2010) (Figure 1.4 D).

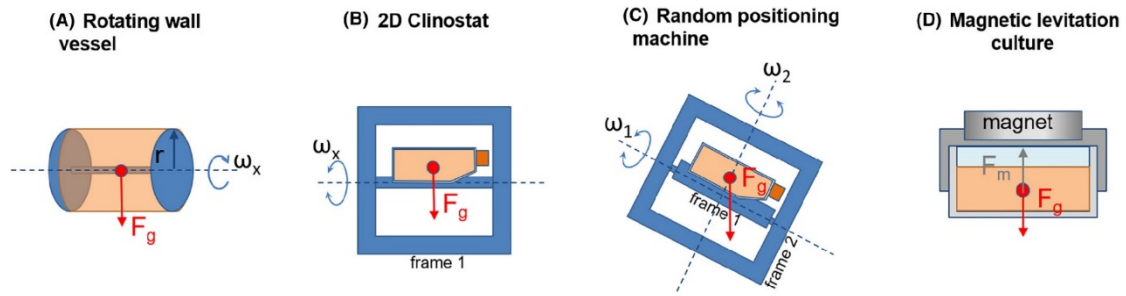


Figure 1.4 Four different classes of ground-based microgravity simulators for biological samples.

RWVs rotate a complete cylindrical vessel horizontally, clinostats rotate samples around a single rotational axis, RPMs and 3D clinostats rotate samples on two perpendicular axes, and magnetic levitation systems add heavy magnetic fields to suspend cells that are normally conjugated to ferromagnetic particles. ω represents angular velocity, F_g represents gravitational force, F_m represents magnetic force, and r represents radius. (The image has been reproduced from (Poon, 2020) under the terms of the Creative Commons Attribution License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, provided the original author and source are credited..)

1.3 Systems Biology

One of the broader definition of systems biology is “To understand complex biological systems requires the integration of experimental and computational research — in other words a systems biology” by Kitano (Kitano, 2002) .Biological systems exhibit complexity at different spatial layers. Several functionally diverse components in biological systems interact highly selectively and nonlinearly to achieve coherent behaviors (Kitano, 2002). These components can be mRNA, gene network, protein network, cells in a tissue, organs and even whole organism in a community. The common features at each level are the overwhelming number of components and the selective and nonlinear interactions between them, which make the behaviors of these systems difficult to understand intuitively. The Human genome project (Lander *et al.*, 2001) pushed and popularized a new approach in biology, which we know as a systems approach. In systems biology, specific genes or proteins are not studied one at a time, otherwise known as the

General Introduction

reductionist approach, a successful mode of biology for the past 30 years (Ideker, Galitski and Hood, 2001). Instead, it focuses on the regulation and interactions of all of the elements in a biological system. These data can then be combined, visualized graphically, and computationally modelled to understand the system's structure and response to any perturbations. A system-level interpretation of a biological system can be obtained by examining the system structure (Kitano, 2002). This involves the network of gene interactions and biochemical processes and the molecular mechanisms by which these interactions affect the physical properties of intracellular and multicellular structures. The valuable source for such information is a large-scale, comprehensive database on gene-regulatory and biochemical networks, e.g. Molecular signature database (Subramanian, Subramanian, *et al.*, 2005). Systems biology publications has grown over the last two decades, covering wide range of topics such as gene expression analysis (Perou and Borresen-Dale, 2011), evolutionary biology (Koonin and Wolf, 2006), stem cells (Kinney *et al.*, 2019), network biology (Censi *et al.*, 2018), molecular diagnostics (Subramanian, Tamayo, Vamsi K Mootha, *et al.*, 2005), cancer (Fu, 2014), etc..

Microgravity-exposed cells (both in outer space and in simulated microgravity) experience significant alterations in cytoskeleton architecture, cell shape, molecular pathways, and gene expression, among other things (Hammond *et al.*, 1999; VASSY *et al.*, 2001). These impacts on cells and tissues, in the long run, may result in serious physiological and medical disorders, including osteoporosis, muscular atrophy, heart failure, metabolic and immune system abnormalities, and others (*Space Physiology - Jay C. Buckey - Oxford University Press*, no date). Indeed, too many activities and cell properties change in concert as a result of microgravity exposure, therefore it has been

suggested that the effect of microgravity can only be understood by looking at the system as a whole (Bizzarri, 2014).

We have used this approach in the second chapter to understand the global changes in human genetic pathways and networks in response to microgravity.

1.4 Synthetic Biology: From Bench to Space

“Synthetic biology is the design and construction of new biological entities such as enzymes, genetic circuits, and cells or the redesign of existing biological systems. Synthetic biology builds on the advances in molecular, cell, and systems biology and seeks to transform biology in the same way that synthesis transformed chemistry and integrated circuit design transformed computing” as defined by Engineering Biology Research Consortium (EBRC) (*What is Synthetic/Engineering Biology?* | EBRC, no date). The focus on design-based approach and control makes it different from conventional genetic engineering. In synthetic biology the emphasis is given in design and construction of genetic circuit or metabolic pathways, whose behavior can be modelled, understood and optimized to meet certain performance criteria, as well as the assembly of these smaller components and devices into larger integrated systems to solve specific problems.

Its objective is to develop novel artificial living organisms designed in the same way as electronic devices to perform complex user-defined tasks (Ro *et al.*, 2006; Fernandez-Rodriguez *et al.*, 2017). The fundamental concept is to view the basic cell components as distinct modules and use them to design entirely new species or allow current organisms to perform functions that they cannot perform in nature (Serrano, 2007). The design principles are derived from engineering; similarly to how an electrical circuit is built by linking resistors, capacitors, and diodes, genetic circuits are

General Introduction

constructed by ligating DNA sequences such as chromosomes, promoters, or other regulatory sequences (Andrianantoandro, Basu, David K Karig, *et al.*, 2006). The modularity concept in biological systems suggests that biological systems can be engineered through the development and connection of independent biochemical modules to produce higher-order functional systems. Mathematically, the functions of these basic modules can also be modelled by solving ordinary differential equations (ODEs), which characterize biochemical reactions in quantitative terms (Szallasi, Stelling and Periwal, 2010).

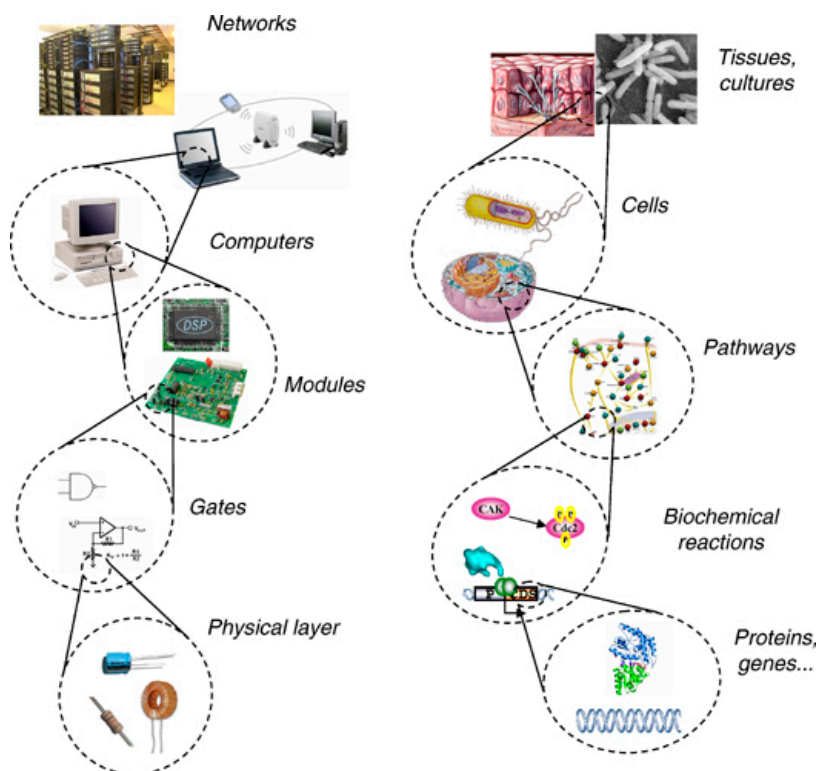


Figure 1.5 The design principles of synthetic biology are borrowed from engineering (Andrianantoandro, Basu, David K Karig, *et al.*, 2006).

The genes or proteins, biochemical reactions, pathways, cellular system, and tissue are analogous to the physical layer, the electronic logic circuits, electronic modules, an entire computer, and a network of computers. This picture also highlights the modular nature of biology and the hierarchical nature (Reproduced with permission from (Andrianantoandro, Basu, David K Karig, *et al.*, 2006)).

Over the last decade, we have seen exponential progress in commercial DNA synthesis in terms of cost, time and maximum length (Carlson, 2009); this has enabled

General Introduction

the possibility of creating existing organisms from scratch. In 2010, Craig Venter Institute (JCVI) created the first bacterial cell *Mycoplasma mycoides* JCVI-syn1.0, with complete synthetic genome (Gibson *et al.*, 2008), it was able to replicate and showed the expected phenotype. Although the general approach in this field is to engineer pre-existing organisms to perform a novel task, the entire genome need not be artificially constructed, instead, the novel function is implemented by introducing an exogenous genetic program into the organism of interest (also called host or chassis).

The exogenous genetic program, which imparts novel function to a cell, comprises synthetic gene circuits. Genetic modules, like promoter, regulatory and effector genes, ribosome binding site, etc., are assembled based on a regulatory and quantitative design to achieve a function in a synthetic gene circuit. In traditional engineering, a circuit is composed of well-characterized standard parts. A Similar fashion of standardization has been adopted for synthetic biology. One such popular strategy was Biobrick standard (Shetty, Endy and Knight, 2008), where each genetic parts were flanked by four specified restriction sites, which helped in the assembly of these parts on a plasmid using standard cloning technique; this also helped in building an initial repository of parts (*Registry - igem.org*, no date), which facilitated the exchange of parts among labs in a popular competition iGEM (*igem.org*, no date). The details of such standards and the assembly techniques have been discussed in chapter four, where we introduced a new method for assembly of complex gene circuits and demonstrated its application by developing integrated logic circuit (IC) and controlled cellular property in *E.coli*. We have also developed synthetic gene circuits with application in microgravity; see chapter three.

General Introduction

Synthetic genetic circuits have been used to engineer bacterial (Cameron, Bashor and Collins, 2014) and mammalian cells (Kis *et al.*, 2015) for plethora of applications like blood glucose homeostasis (Ye *et al.*, 2011) biosensors (Pardee *et al.*, 2014; Saltepe *et al.*, 2018), bioremediation (Lorenzo *et al.*, 2018) and bioproduction (Guo *et al.*, 2018), etc.. The Keasling lab at the University of Berkeley obtained the first high-impact breakthrough of synthetic biology when they developed a microorganism to create the antimalarial drug precursor artemisinic acid, which can then be converted to the active ingredient artemisinin (Ro *et al.*, 2006).

In the year 2010, realizing the potential of synthetic biology in NASA's future missions, a workshop was organized on the potential roles for Synthetic Biology in NASA's Mission (Langhoff, Cumbers and Worden, 2011) to identify and discuss opportunities in space synthetic biology. A comprehensive account of such opportunities and challenges is summarized in Table 1.2, referred from (Menezes, Montague, *et al.*, 2015). These challenges can be addressed by conventional Physico-chemical means, but that would require a lot of machinery, increased power demand and launch volume in terms of propellants, food and raw material. In long-distance space missions, it will increase dependency on resupplies, thus make the mission unfeasible. Biological technologies are versatile, and it can reduce power demand and launch volume as well as dependency on resupplies, by harnessing solar energy and available destination nutrients for example from Mars, to biomanufacture food, fuel, pharmaceuticals and biopolymers (Menezes *et al.*, 2014; Menezes, Montague, *et al.*, 2015). Microorganisms have the potential to be highly beneficial. Humans have consumed and used microorganism-derived resources on Earth throughout human existence: oxygen produced by

cyanobacteria and eukaryotic microalgae, food and beverages produced by edible microorganisms and fermented items.

Additionally, we depend on them for a variety of essential procedures, such as waste recycling. All species known to date originated on Earth and are not suited to the most habitats occurring elsewhere (Verseux *et al.*, 2015). In extreme environment of space, the majority of the substrates on which they typically depend are absent. If we must transport all starting compounds required for microbial processes to occur from Earth, we can merely reduce the launch volume, not solve it. For example, the mass of metabolic consumables required to maintain a six-person crew on Mars for 1000 days using the European life support system MELiSSA (Gòdia *et al.*, 2002) has been estimated to be approximately 30 metric tonnes, not including hygiene water (Langhoff, Cumbers and Worden, 2011). Then again, the conditions found outside of Earth are usually extremely hostile to all known microorganisms, and replicating Earth-like conditions within a vast volume and surface area will be expensive.

Thus, synthetic biology can be a complimentary solution to improve the fitness of and confer new functions to these microbes, as well as to employ them to solve pertinent concerns. A detailed account of space synthetic biology can be found in these articles (Langhoff, Cumbers and Worden, 2011; Montague, George H McArthur, *et al.*, 2012; Menezes *et al.*, 2014; Menezes, Montague, *et al.*, 2015) and book chapter (Verseux *et al.*, 2015).

Table 1.2 Summary of opportunities and challenges for space synthetic biology.

A part of this table is reproduced from (Menezes, Montague, *et al.*, 2015) under the terms of the Creative Commons Attribution License

<http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, provided the original author and source are credited.

1. *In-Situ* Resource utilization: Relying on ‘on’-site available raw materials, to generate valuable resources.

- (a) Ensuring functionality in extreme environments.
- (b) Providing the capacity to harness three kinds of resources: wastes, volatiles and minerals.
- (c) Producing feedstocks for manufacturing processes and cell-based biomaterials for construction processes.

Recent developments:- Engineered *E.coli* was able to bind silicon from lunar regolith (Lehner *et al.*, 2019).

2. Manufacturing

- (a) Satisfying construction-related desires with adhesives to bind regolith, biocement and biopolymers.
- (b) Generating fuel for power and propulsion.
- (c) Revisiting abiotic manufacturing and construction technologies to leverage existing or synthetic biology capabilities.

Recent developments:- Metabolically Engineered *Bacillus subtilis* to produce a polymer precursor (Averesch and Rothschild, 2019).

3. Life support (Gòdia *et al.*, 2002)

- (a) Improving the biological management of waste, especially wastewater.
- (b) Treating, conditioning and recycling air, water and solid wastes through incorporating biology into traditionally inanimate structures, e.g., creating a ‘living’ habitat.
- (c) Producing flavourful, texture-rich and nutritious food.
- (d) Providing nutrients, and assisting with the recycling of nutrients.

Recent developments:-Microbial fuel cell, to generate electricity from wastewater (‘NASA - Synthetic Biology and Microbial Fuel Cells: Towards Self-Sustaining Life Support Systems’, no date), Potential candidates for Microbial electrosynthesis and CO₂ fixation(Abel *et al.*, 2020).

4. Space medicine and human health

- (a) Preventing disease and maintaining the human microbiome.
- (b) Manufacturing synthetic drugs to combat disease, radiation damage and the effects of reduced gravity.
- (c) Developing radiation-resistant, self-healing protective clothing and personal shielding.

Recent developments:- Strategies on molecular pharmaceutical development for space application (McNulty *et al.*, 2020).

5. Space cybernetics

(a) Developing device-level biological control systems: biological sensors, actuators and controllers.

(b) Designing biological control systems that are either completely composed of biological parts or partially integrate biological controllers and systems with abiotic sensors and actuators as a form of artificial life.

Recent developments: - We developed a microgravity responsive synthetic genetic device in *Escherichia coli*, refer to chapter three, synthetic biological circuit in plant tested in spaceflight (Kitto *et al.*, 2021).

6. Terraforming

(a) Paraterraforming with few multi-functional species that complete the carbon and nitrogen cycle

Recent developments:- Design hypothesis for synthetic genetic circuits for earth terraforming (Solé, Montañez and Duran-Nebreda, 2015).

Chapter 2: A systems biology pipeline identifies new immune and disease related molecular signatures and networks in human cells during microgravity exposure

2.1 Introduction

The future plan of manned mission to Mars and asteroids (Binzel, 2014) requires astronauts to spend years in space. Spaceflight poses multiple challenges that can compromise the health of the astronauts. Human spaceflight since 1960's provided much information on impact of environmental stressors like weightlessness (also called microgravity), space radiation on human physiology and disrupted circadian rhythm (Souza, Hogan and Ballard, 1965). Human research program has classified five categories of hazards which is encountered by astronauts: -

- A) **Space Radiation:** Space radiation is different from common terrestrial forms of background radiation. Our magnetosphere protects us from dangerous radiation from the sun and space. The sun's radiation is made up of varying amounts of high-energy protons. Low levels of heavy charged particles make up space radiation. Both shielding materials and biological structures may be affected by high-energy protons and charged particles. Radiation raises cancer risk, affects the central nervous system, can affect the cognitive capacity, reduce motor functions and induce behavioural changes ((*No Title*), no date b; Mars, 2018; Löbrich and Jeggo, 2019).

- B) Isolation and confinement:** Isolation and confinement can cause mood, cognition, morale, and interpersonal relationships to deteriorate. Isolation and confinement for an extended period of time can lead to fatigue and depression. It's important to consider group dynamics and suitability for spaceflight because astronauts must work together for long periods of time in confined spaces. Devices like an actigraphy can be used to monitor and enhance sleep in order to improve mood and alertness. LED lights that sync to circadian rhythms can also help you sleep better and feel better (*Human Health on the Space Station*, no date; Mars, 2018).
- C) Distance from earth:** The crew's self-sufficiency is vital due to the distance between the ISS and Earth. The crew must be able to complete their task without assistance. In space, astronauts are responsible for their own wellbeing (*Human Health on the Space Station*, no date; Mars, 2018).
- D) Hostile/closed environments:** It's essential to know the microbes that reside in the ISS's sealed climate (Checinska Sielaff *et al.*, 2019). The spread of disease-causing microbes is determined by living in close quarters. This, coupled with reduced immunity and higher stress levels, suggests that humans in space have a higher risk of being sick. Researchers can keep track of possible contaminants by monitoring air quality and swabbing the ISS. In addition, keeping a close eye on improvements in astronauts' immune systems might help them avoid getting sick (*Human Health on the Space Station*, no date; Mars, 2018).
- E) Microgravity:** Microgravity is one of the most prominent health hazards for astronauts (Anon, 2015; Setlow, 2003). Physiological changes in many organ systems due to a short to moderate microgravity exposure (days to months)

Chapter two

have been recorded since the 1960s and 1970s space exploration. Bone and muscle loss, puffiness in the face, change in cardiovascular physiology, catecholamine cardiomyopathy, insufficient blood flow in the brain, genitourinary issues and disturbance in neurovestibular system are common among space travel crewmembers (Anon, 2015; Jones et al., 2005; Otsuka et al., 2015; Setlow, 2003; Sun et al., 2015; White and Averner, 2001). Microgravity is largely a unknown risk and one of the factor which is currently limiting attempts to visit Mars, necessitating extensive research on its effects on astronauts (*Human Health on the Space Station*, no date; Mars, 2018).

Among these major hazards, microgravity is the most challenging, as it is ubiquitous in space. Apart from the physiological changes mentioned before, microgravity is known to induce deregulation of human immune systems (Verhaar *et al.*, 2014a; Crucian, Raymond P Stowe, *et al.*, 2015). The earliest report came from Soviet immunologists in the early 1970s (IV Konstantinova, 1973). Multiple gene expression studies showed microgravity-induced signature of early inhibition in T cell activation (Hauschild *et al.*, 2014), impaired endothelial cell function (Jeanette A M Maier *et al.*, 2015), cellular senescence (Liu and Wang, 2008), alteration of genes related to cell cycle (Versari *et al.*, 2013a; Vidyasekar *et al.*, 2015), cell adhesion (Jeanette A.M. Maier *et al.*, 2015), oxidative phosphorylation (Versari *et al.*, 2013b) and apoptosis (Versari *et al.*, 2013a). It has been showed that the reduced immunity may result from inhibition of NF- κ B/Rel pathway, downregulation of early T cell activation genes, IFN- γ and EL-2R α genes (Chang *et al.*, 2012) and impairment of Jun-N-terminal kinase activity (Verhaar *et al.*, 2014b). The compromised immune system increases the risk of infection by pathogen like *Salmonella*, virulence of which is increased in microgravity (Weinstein

Chapter two

and Mermel, 2013). *Salmonella* and *Pseudomonas aeruginosa* PAO1 infection among astronauts is a well-known health hazard documented starting from Apollo and Skylab missions (Hawkins and Ziegelschmid, 1975; Weinstein and Mermel, 2013). Further, microgravity alters level of micro RNAs (miRNAs), many of which are related with inflammation (Girardi, 2014) and multiple cancer types (Mangala *et al.*, 2011; Girardi *et al.*, 2014; Vidyasekar *et al.*, 2015). However, the studies showed controversial inference based on the expression of different microRNAs. For example, expression of hsa-miR-423-5p and hsa-miR-222 in microgravity suggest the induction of breast cancer, whereas expression of hsa-miR-141 suggests the decrease in the same (Mangala *et al.*, 2011). Similar controversial miRNA expression pattern was observed for leukaemia and lung cancer (Girardi *et al.*, 2014). Further, as a single miRNA is related with several cancer types and opposite results in miRNA alternation are observed among studies (Mangala *et al.*, 2011; Girardi *et al.*, 2014; Vidyasekar *et al.*, 2015), there is uncertainty to identify specific cancer signatures, if any, associated with microgravity. No cancer related signatures and inflammation signature were identified in normal human cells through gene expression data alone. Thus, the connection of cancer induction with microgravity is undefined and no assessment reports included microgravity-associated cancer as a risk factor. However, the ambitious plan for sending humans to Mars and asteroids requires a thorough understanding about the effect of microgravity at the cellular level to estimate the risk for all potential diseases and health conditions and develop protocols against any adverse effect of space on the astronauts. In spite of its prevalence, a detailed molecular systems level picture on how various molecular pathways in human cells get affected by microgravity is largely unknown.

2.1.1 Limitation of using conventional data analysis in microgravity based transcriptomic studies.

Several gene expression experiments were performed in the 1990s to determine the impact of microgravity on cells, but these studies appeared to focus on just a few genes at a time (Souza, Hogan and Ballard, 1965; Souza, Etheridge and Callahan, 1991). Large-scale genome-wide studies were only possible with the advent of high-throughput genomic technologies such as microarrays (Hammond *et al.*, 1999). In the previous microgravity studies, the transcriptomics data of human cells were analysed by differential gene expression analysis, followed by passive pathway mapping (Ward *et al.*, 2006; Chang *et al.*, 2012; Versari *et al.*, 2013a; Girardi *et al.*, 2014). Differential gene expression analysis relies on arbitrary cut-off value (>1.5 – 2 folds) in fold change of individual genes. Limitations with such approach is that it may overlook the pathway level picture due to absence of genes with lower expression values. For example, this gene centric method cannot identify the downregulation of oxidative phosphorylation pathway in diabetes, where the mean decrease in member genes' expression is about 1.2 folds (Patti *et al.*, 2003; Subramanian, Tamayo, Vamsi K. Mootha, *et al.*, 2005). This is specifically critical for the situation like microgravity, which results an overall low fold change in the global gene expression compare to other perturbation like cancers (Hibbs *et al.*, 2004; Fu *et al.*, 2014). Further, previous studies relied on KEGG, GO databases and a few manually curated ontologies for pathway analysis, missing a huge number of disease and immunity related pathways. The assignments of the pathways were also arbitrary. Pathways were assigned even when the fraction of mapped genes were as low as 2% of the whole pathway (Ward *et al.*, 2006; Versari *et al.*, 2013a).

Chapter two

In this work, by integrating a set of statistical and machine learning tools systems and databases, we have established a systems biology pipeline (Figure 2.2) and analyzed the effect of microgravity on more than 8000 molecular pathways on normal human cells from published global gene expression data (See next section). We have identified new pathways, mechanism and plausible regulatory and functional connections across the gene networks in microgravity, which cannot be identified by conventional analysis.

2.2 Systems Biology Pipeline

2.2.1 The gene expression data

The global gene expression datasets of human cell under microgravity from 5 important works were mined from ArrayExpress (Parkinson *et al.*, 2007), a repository of functional genomics data, URL: <https://www.ebi.ac.uk/arrayexpress/>. Three out of those five experiments were performed in space-flight conditions. Dataset with accession number E-GEOD-38836 (Chang *et al.*, 2012) represents International Space Station (ISS) study of human T-cells, E-GEOD-43582 (Versari *et al.*, 2013a) represents Progress 40 P space flight mission (ESA-SPHINX) with human umbilical vein endothelial cells (HUVEC) and E-GEOD-54213 [unpublished] represents space shuttle STS-135 study of human endothelial cells. Datasets E-GEOD-4209 (Ward *et al.*, 2006) and E-GEOD-57418 (Girardi *et al.*, 2014) represent ground based simulated microgravity studies with activated human T-cells and human peripheral blood lymphocyte (PBMC) respectively. To our knowledge, no other normal human cell gene expression study under microgravity condition were available in the public domain at the time of this study.

2.2.2 Database for predefined human molecular pathways

Those 8000 predefined molecular pathways or gene sets, which covers almost all known aspects of human cell including positional (chromosomal) gene set, chemical and genetic perturbation, canonical pathways, cancer gene neighbourhoods, cancer modules, oncogenic signatures, immunogenic signature and hallmark gene sets, were extracted from Molecular Signature Database v 5.0 (MSigDB), which brought together pathways from all standard databases on human. A Gene set houses genes that encode products involved in a particular metabolic pathway, are found in the same cytogenetic band, or have the same GO classification, or are known to function. Brief overview of the categories of gene sets are tabulated in Table 2.1, detailed description is available in this URL <http://www.gsea-msigdb.org/gsea/msigdb/index.jsp>. However, we have excluded Gene Ontology (GO) classification, as it appears too broad and less specific.

Table 2.1 Brief Description of categories of gene sets used in this study

Serial No.	Major Collection	Name	No of gene sets	Brief Description
1	Curated gene sets c2	chemical and genetic perturbation	3395	Gene sets represent expression signatures in response to genetic and chemical perturbations, assorted from biomedical literature.
2		Canonical Pathways	1330	Gene sets derived from online pathway databases and knowledge of domain experts. It comprises of cellular, metabolic and disease pathways from KEGG, Biocarta and Reactome.
3	Computational gene sets c4	Cancer Module	431	Gene sets defined by Segal et al. 2004. Briefly, the authors compiled gene sets ('modules') from a variety of resources such as KEGG, GO, and others. They identified 456 such modules as significantly changed in a variety of cancer conditions by mining a large compendium of cancer-related microarray data.
4		Cancer Gene Neighbourhood	427	Gene sets defined by expression neighborhoods centered on 380 cancer-associated genes. This collection is described in Subramanian, Tamayo et al. 2005
5	c6	Oncogenic Signature	189	Gene sets that represent signatures of cellular pathways which are often dis-regulated in cancer.
6	c7	Immunologic Signature	1910	Gene sets in this collection represent cell states and perturbations within the immune system
7	c1	Positional Gene sets	326	Gene sets corresponding to each human chromosome and each cytogenetic band
8	H	Hallmark Gene sets	50	Hallmark gene sets are coherently expressed signatures derived from the aggregation of many MSigDB gene sets to represent well-defined biological states or processes.
		Total	8058	

2.2.3 Identification of pathways and leading genes

We have applied Gene Set Enrichment Analysis (GSEA) (Subramanian, Tamayo, Vamsi K. Mootha, *et al.*, 2005) (URL: <https://www.gsea-msigdb.org/gsea/index.jsp>), a computational method which determines the likeliness of any specific predefined molecular pathways (gene set) to be involved with microgravity directly from the global gene expression data, in a statistically significant way. The p value and False Discovery Rate (FDR) q value of entire gene set are calculated in GSEA. Thus, unlike passive pathway mapping in differential gene expression analysis, each altered pathway in GSEA is chosen or discarded based on statistical parameters. In GSEA, from the global gene expression data, a ranked list of genes is prepared according to the difference in gene expression values between microgravity and normal gravity divided by the standard deviation between biological replicates. Figure 2.1 (a) shows the heat map of a ranked list of an experiment for ease of visualization. A running sum statistic estimates the likeliness of genes from a specific molecular pathway to be appeared in top or bottom of the list in a statistically significant way (Figure 2.1 b). Details of the method can be found in methods section (2.6). Next, we have identified leading edge genes (Table 2.6) before enrichment score (ES) peak values (Figure 2.1 b). Leading edge genes contributed the most for a pathway to be altered in microgravity.

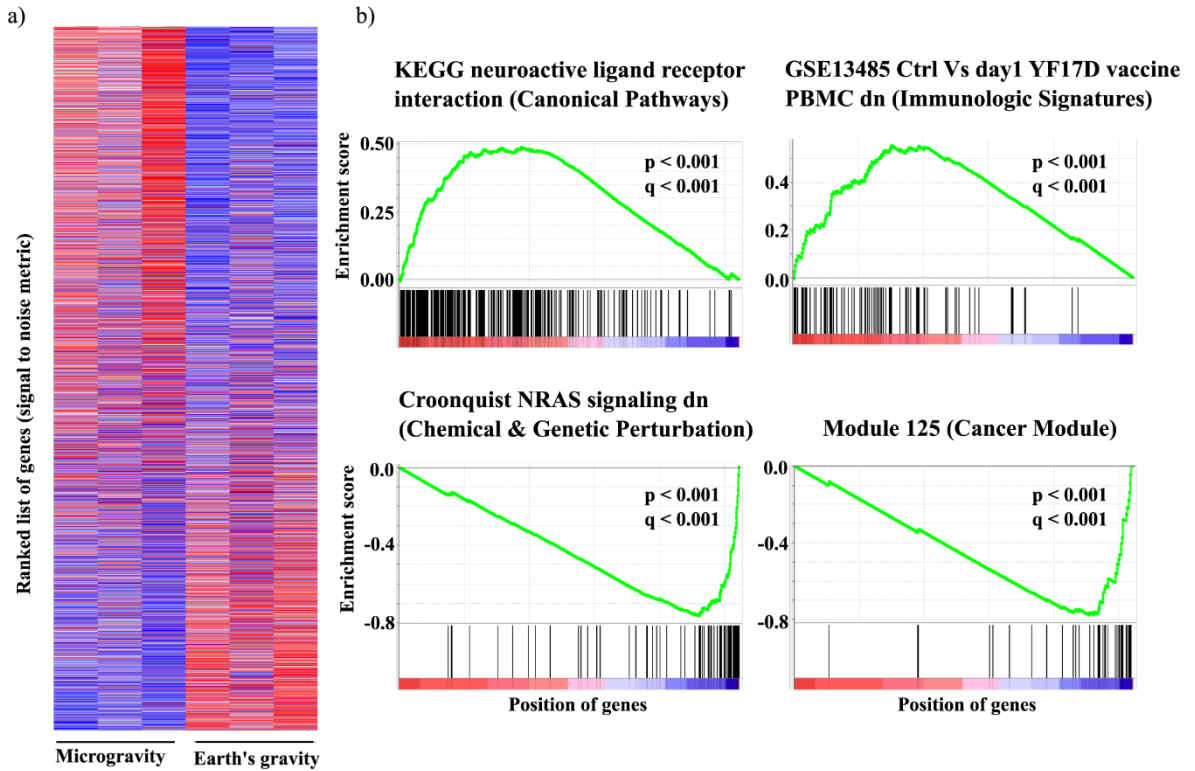


Figure 2.1 GSEA analysis of human cell during microgravity exposure.

(a) Genome wide expression profile of E-GEOD-4209 for three biological replicates in microgravity and normal gravity. The heatmap represents ranked genes, which is created by signal-to-noise matrix (fold change in expression between microgravity and normal gravity divided by the standard deviation in gene expression among the replicates). (b) Enrichment score of four representative, statistically significant gene sets/pathways from four different gene set modules. The top and bottom figures represent upregulated and downregulated pathways respectively. The heatmap of the ranked gene list is shown at the bottom of each pathway. The black lines within the heatmap represent the position of the pathway genes in that ranked list.

2.2.4 Extracting gene expression patterns and networks in microgravity

We have applied an unsupervised non-negative matrix factorization coupled with a consensus-clustering algorithm (NMFC) (Lee and Seung, 1999; Brunet *et al.*, 2004) to extract the genes (leading edge genes) with similar expression patterns in microgravity. NMFC was found superior to extract the functional meaningful subset of genes from genome scale transcriptomics data compare to other clustering algorithm including

Chapter two

principal component analysis, K means, self-organizing maps, singular value decomposition and hierarchical clustering (Kim and Tidor, 2003; Brunet *et al.*, 2004; Devarajan, 2008). The genes with similar expression patterns within a mathematically stable cluster suggest plausible regulatory or functional connections in microgravity. Those genes within a single cluster were further mapped on STRING network databases (Franceschini *et al.*, 2013) to evaluate the functional networks. Further details about the NMFC and network mapping can be found in results and method sections.

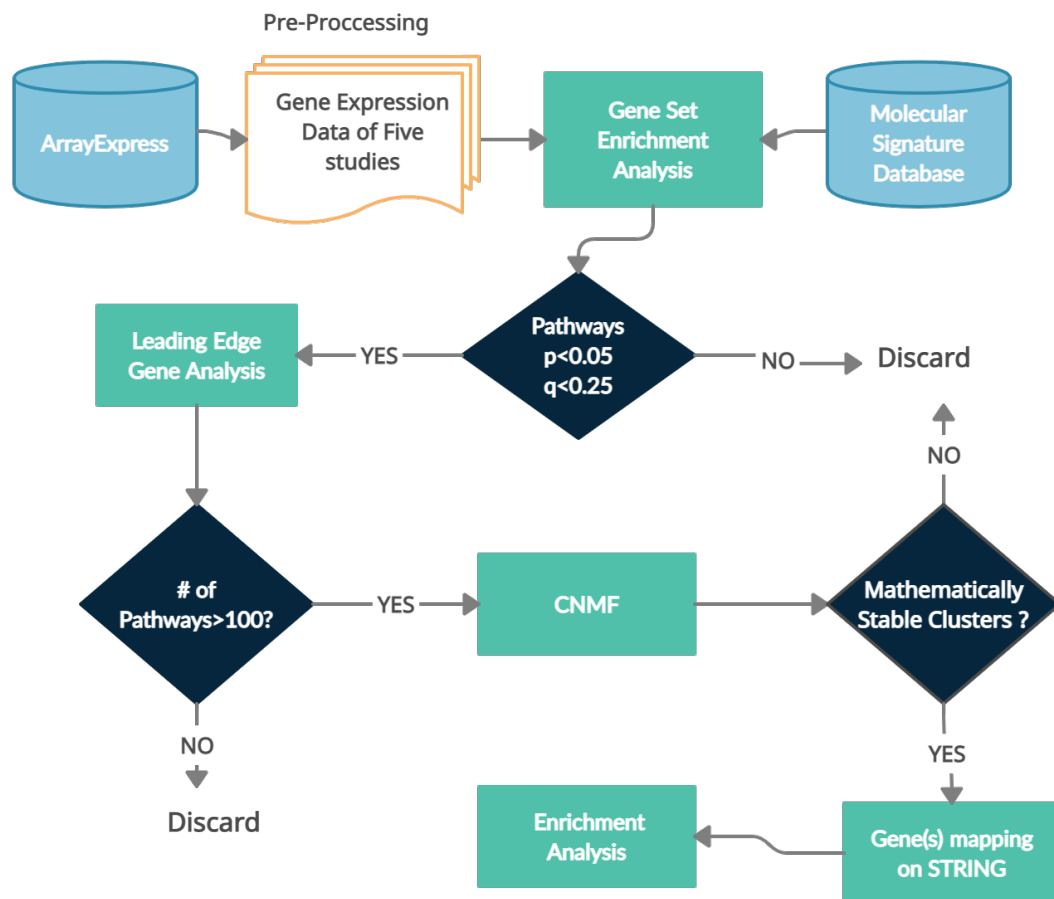


Figure 2.2 Flowchart of the systems biology pipeline.

The pipeline consists of four significant steps (in green rectangular box) with intermediate steps. The gene expression datasets for five microarray experiments on microgravity exposed human cells are mined from Arrayexpress. These input files are pre-processed before gene set enrichment analysis (GSEA) step. The GSEA is integrated with Molecular signature database, a repository of various categories of pre-defined annotated gene sets or molecular pathways of human cell. In the next step the gene set

enrichment analysis identifies which molecular pathways are altered in a given condition, with statistical confidence. Next, we identified the common altered pathways with high statistical confidence (high normalized enrichment score (NES) values and $p < 0.05$), overlapping at least three experiments. Next, we identified the leading-edge genes from these common pathways and applied non-negative matrix factorization coupled with consensus-clustering algorithm (CNMF) to cluster genes with similar expression pattern. Finally, these mathematically stable clusters were mapped on STRING network database to evaluate function networks.

2.3 Results

2.3.1 Molecular pathway analysis using gene set enrichment analysis (GSEA)

During microgravity exposure, a large number of molecular pathways from each study are found to be upregulated or downregulated in a statistically significant way ($p < 0.01$, $q < 0.25$) (Figure 2.3). A total number of 1976, 560, 1224, 0 and 836 pathways were altered spanning all the modules for E-GEOD-4209, E-GEOD-38836, E-GEOD-43582, E-GEOD-54213 and E-GEOD-57418 respectively. We have excluded E-GEOD-54213 (unpublished data) from main analysis, as it shows no statistically significant altered pathway.

Chapter two

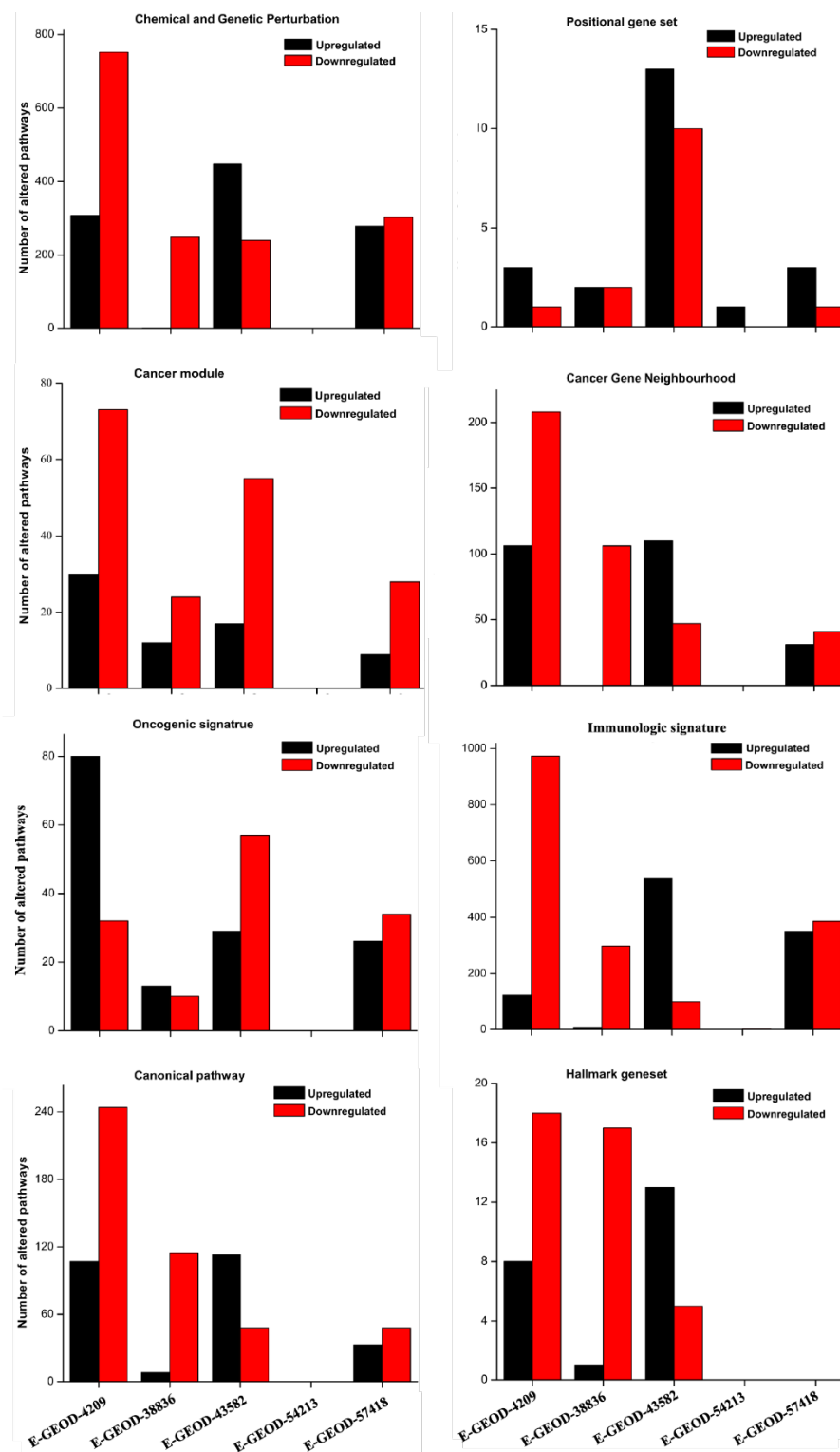


Figure 2.3 Number of statistically significant ($p < 0.001$, $q < 0.001$) altered molecular pathways identified by GSEA for each experiment.

Gene set enrichment analysis on microarray data of each experiment identified statistically significant altered gene sets or molecular pathways. The gene sets were extracted from Molecular Signature Data-base v 5.0 and they belonged to various categories of almost all known aspects of human cell including positional (chromosomal) gene set,

Chapter two

chemical and genetic perturbation, canonical path-ways, cancer gene neighborhoods, cancer modules, oncogenic signatures, immuno-genic signature and hallmark gene sets. Each bar chart represents one of the eight gene set modules from these categories.

GSEA shows ability to extract more pathways with statistical confidence than differential gene expression analysis (DGE). For example, reference (Versari *et al.*, 2013a) (E-GEOD-43582) identifies only 8 KEGG pathways without any quantitative confidence. Using GSEA we have identified total 45 altered KEGG pathways (26 upregulated and 19 downregulated) with $p < 0.05$ and high normalized enrichment score (NES) value (Table 2.2 & Figure 2.4).

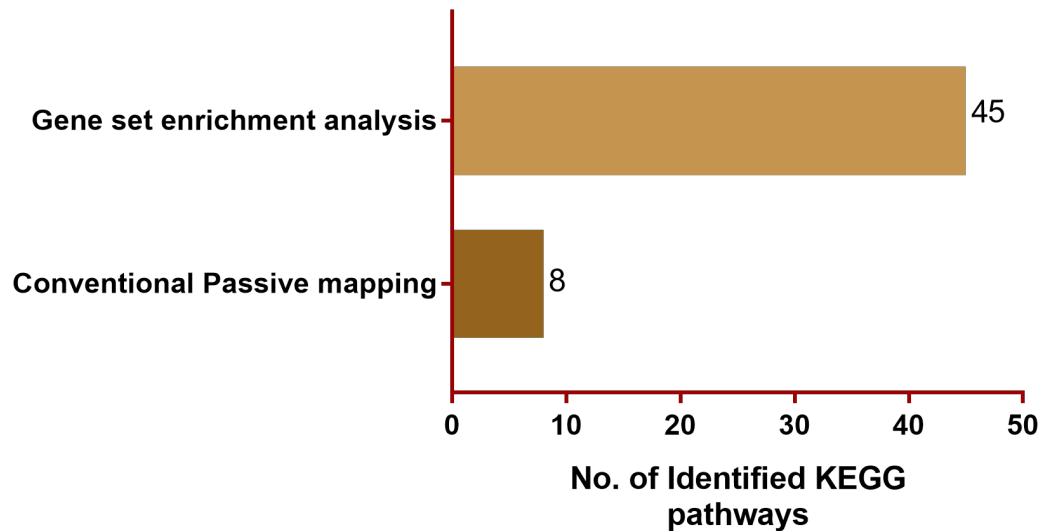


Figure 2.4 GSEA identified greater number of pathways compared to DGE analysis.

We compared the performance of GSEA and conventional passive mapping approach used in the experiment (E-GEOD-43582) in identifying altered pathways. The original study E-GEOD-43582 identified only 8 KEGG pathways without any statistical confidence, whereas GSEA identified 45 KEGG pathways with $p\text{-value} < 0.05$ and high normalized enrichment score (NES) for the same experimental dataset. Each bar represents the number of identified KEGG pathways in each case for experiment E-GEOD-43582.

Table 2.2 Altered KEGG pathways identified by GSEA for dataset E-GEOD-43582.

No. of path- ways	Upregulated KEGG pathways in E-GEOD-43582 by GSEA	NES	p value	q value
1	KEGG_ABC_TRANSPORTERS	1.537154	0.028419	0.159726
2	KEGG_AMINOACYL_TRNA_BIOSYNTHESIS	1.59197	0.010676	0.118971
3	KEGG_ASCORBATE_AND_ALDARATE_METABOLISM	1.590584	0.032727	0.118964
4	KEGG_BASAL_TRANSCRIPTION_FACTORS	1.645014	0.014572	0.084383
5	KEGG_BASE_EXCISION_REPAIR	1.688345	0.012302	0.065025
6	KEGG_CELL_CYCLE	2.052853	<0.001	0.002014
7	KEGG_DNA_REPLICATION	2.406814	<0.001	<0.0001
8	KEGG_GLYCEROLIPID_METABOLISM	1.447228	0.041451	0.2281
9	KEGG_GLYCOSYLPHOSPHATIDYLINOSITOL_GPI_ANCHOR_BIO- SYNTHESIS	1.990355	0.001773	0.004016
10	KEGG_HOMOLOGOUS_RECOMBINATION	2.111701	<0.001	7.17E-04
11	KEGG_LYSOSOME	1.395146	0.032761	0.279942
12	KEGG_MISMATCH_REPAIR	2.1679	<0.001	2.13E-04
13	KEGG_N_GLYCAN_BIOSYNTHESIS	1.796955	<0.001	0.027784
14	KEGG_NUCLEOTIDE_EXCISION_REPAIR	2.179221	<0.001	1.92E-04
15	KEGG_OLFACTORY_TRANSDUCTION	1.361321	0.009943	0.307224
16	KEGG_ONE_CARBON_POOL_BY_FOLATE	1.48253	0.049632	0.198121
17	KEGG_OOCYTE_MEIOSIS	1.656462	0.001689	0.076775
18	KEGG_P53_SIGNALING_PATHWAY	1.409219	0.046763	0.26277
19	KEGG_PENTOSE_AND_GLUCURONATE_INTERCONVERSIONS	1.779207	0.005618	0.030507
20	KEGG_PEROXISOME	1.643054	0.003442	0.084623
21	KEGG_PROGESTERONE_MEDIATED_OOCYTE_MATURATION	1.545982	0.009917	0.154961
22	KEGG_PYRIMIDINE_METABOLISM	1.410386	0.024712	0.264212

Chapter two

23	KEGG_SELENOAMINO_ACID_METABOLISM	1.669546	0.009488	0.071862
24	KEGG_STARCH_AND_SUCROSE_METABOLISM	1.742877	0.003571	0.042692
25	KEGG_TASTE_TRANSDUCTION	2.350056	<0.001	<0.0001
26	KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS	1.705043	0.001626	0.058719
Downregulated KEGG pathways in E-GEOD-43582 by GSEA				
27	KEGG_ALLOGRAFT_REJECTION	-1.91765	<0.001	0.050728
28	KEGG_ARACHIDONIC_ACID_METABOLISM	-1.6956	0.002304	0.128111
29	KEGG_ASTHMA	-1.93317	0.002262	0.049885
30	KEGG_AUTOIMMUNE_THYROID_DISEASE	-1.8415	<0.001	0.063863
31	KEGG_CALCIUM_SIGNALING_PATHWAY	-1.55394	<0.001	0.19537
32	KEGG_CARDIAC_MUSCLE_CONTRACTION	-1.67928	<0.001	0.129146
33	KEGG_CHEMOKINE_SIGNALING_PATHWAY	-1.44179	0.005634	0.310231
34	KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	-1.81668	<0.001	0.063237
35	KEGG_DILATED_CARDIOMYOPATHY	-1.53684	0.005013	0.215787
36	KEGG_FRUCTOSE_AND_MANNOSE_METABOLISM	-1.48836	0.030093	0.262869
37	KEGG_GRAFT_VERSUS_HOST_DISEASE	-1.43127	0.046083	0.321264
38	KEGG_HEMATOPOIETIC_CELL_LINEAGE	-1.63429	0.002604	0.145686
39	KEGG_HYPERTROPHIC_CARDIOMYOPATHY_HCM	-1.62532	0.007246	0.15097
40	KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION	-1.84496	0.00232	0.068358
41	KEGG_LINOLEIC_ACID_METABOLISM	-1.4954	0.039627	0.254681
42	KEGG_MAPK_SIGNALING_PATHWAY	-1.38432	<0.001	0.383581
43	KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION	-1.90394	<0.001	0.050412
44	KEGG_RETINOL_METABOLISM	-1.47001	0.034325	0.27182
45	KEGG_TYPE_I_DIABETES_MELLITUS	-1.68162	0.011792	0.131513

Chapter two

Further, we have analysed thousands of pathways from various modules (cancer modules, oncogenic signature, chemical and genetic perturbation etc.), which were not analysed before. More than 98% of those identified pathways are found to be new for each dataset compared to the corresponding published results. Thousands of pathways are altered (Figure 2.3) in our analysis whereas tens of pathways are reported in previous publications (Ward *et al.*, 2006; Chang *et al.*, 2012; Versari *et al.*, 2013a; Girardi *et al.*, 2014).

Next, we have compared the GSEA identified pathways from all 4 datasets and identified more than 100 altered pathways (high NES values and $p < 0.05$), which overlapped among at least three data sets (Figure 2.5 & Table A.2). Around 100 of those overlapped pathways were identified as novel molecular signatures, which were not reported in any of those source publications (see Table A.2), the novel pathways found were marked as ‘This Study Only’. The short description of all the overlapping pathways can be found in Table A.1 . We further classified those overlapped pathways based on specific immunity and disease related functions. All those functions and their molecular signatures (pathways) are tabulated in Table 2.5.

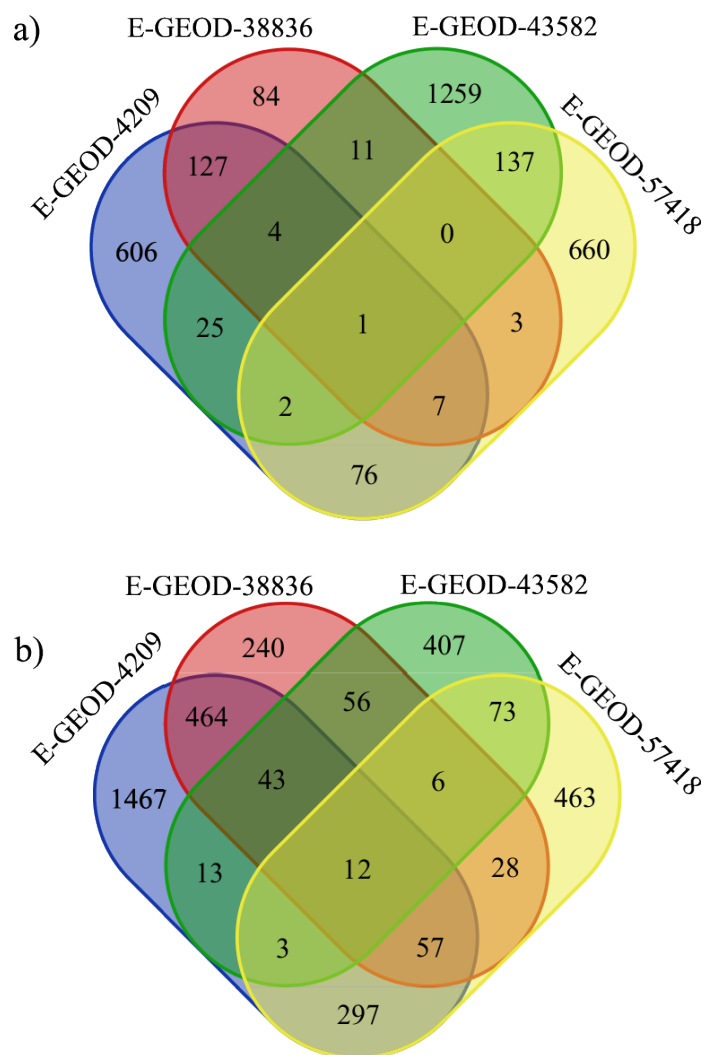


Figure 2.5 Venn diagrams representing the number of common altered pathways in microgravity.

We compared the GSEA identified pathways from four experiments for (a) upregulated pathways and (b) downregulated pathways and represented them using venn diagram. More than 100 altered pathways (high NES values and $p < 0.05$), overlapped among at least three data sets. All altered pathways from 8 gene set categories are combined in this diagram.

2.3.2 Novel molecular signatures related to immunity

Those novel, overlapping pathways can be divided in two categories. In first category, the earlier reported pathways in one study are found in other datasets, where the pathways were never reported. Reference (Chang *et al.*, 2012) showed that one of the reasons for impaired immunity was the downregulation of tumour necrosis factor (TNF) mediated NF- κ B /Rel pathways. We have identified the downregulation of the same

Chapter two

pathway (appeared as Hallmark_TNFA_signaling_via_NF- κ B in our analysis, Table 2.5 and Table 2.3 by analysing data of reference (Chang *et al.*, 2012). In addition, we identified the same pathway in two other studies E-GEOD-43582 (Versari *et al.*, 2013a) and E-GEOD-4209 (Ward *et al.*, 2006), where this pathway was not reported. Similarly, the pathway responsible for generation of second messenger molecules in T-cell regulation (TCR) signalling (Reactome second messenger molecules signalling pathway, Table 2.5 and Table A.2) is repressed in microgravity for all four studies. However, only one study (Girardi *et al.*, 2014), (E-GEOD-57418) suggested the same without any statistical confidence. Those messenger molecules are associated with the activation of both NF- κ B and PKC pathways along with calcium mobilization.

In the second category, we found completely new pathways and class of functions, which were not reported in any of the studies with source datasets. We identified the downregulation NF- κ B pathway via Notch1 signalling (Vilimas Notch1 targets up, Table 2.5 and Table A.2) as a new route for NF- κ B deregulation. In addition, we have found another new pathway, the downregulated genes of LPS induced stimulation of DC macrophage (Zhou inflammatory response live up, Table 2.5 and Table A.2), as an associated signature. It was shown that Notch1 mediated NF- κ B pathway induced LPS stimulated macrophage activation (Monsalve *et al.*, 2009). We found 2 completely new pathways, which could be seen as associated molecular signatures in connection with the reduced second messenger molecules. The, reduced allograft rejection (KEGG allograft rejection, Table 2.5 and Table A.2) and reduced NFAT (nuclear factor for activated T cell) pathway (PID_NFAT_TF pathway, Table 2.5 and Table A.2) in our analysis are most likely connected with the downregulation of the second messenger molecules,

Chapter two

as blocking of this pathway prevents graft rejection in transplanted patients (Fric *et al.*, 2012) and calcium mobilization is required for NFAT activation (Kliem *et al.*, 2012).

To exclude the possibility that this new route (Vilimas Notch1 targets up) for NF- κ B pathway deregulation appeared in our analysis is due to the presence of common genes from Hallmark_TNFA_signaling_via_NF- κ B and Reactome generation of second messenger molecules, we compared the leading edge genes of three pathways for each of the experiments (Figure 2.6 a, b, c). The results showed a negligible overlap (Figure 2.6).

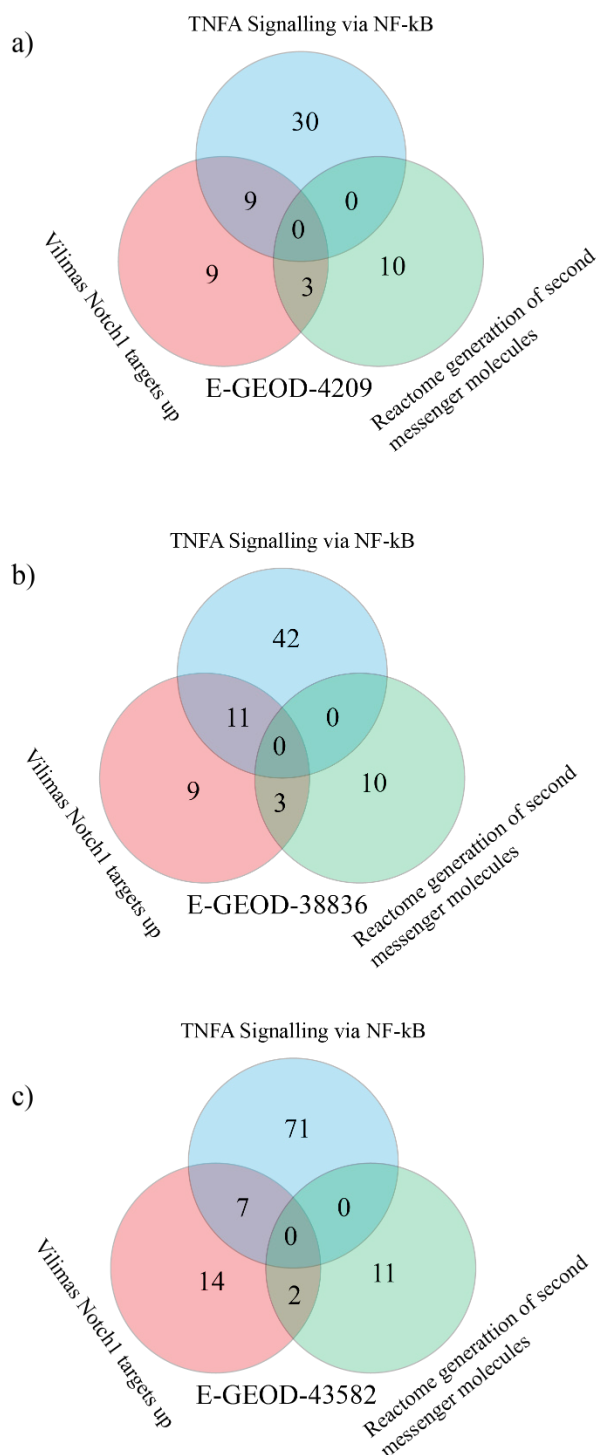


Figure 2.6 Overlapping leading edge genes between three pathways, which affect the regulation of Nf- κB pathway.

In our analysis we found three pathways known to regulate Nf- κB pathway, Vilimas Notch1 targets up, Hallmark_TNFA_signaling_via_NF- κB and Reactome generation of second messenger molecules. To exclude the possibility that these three pathways appeared due presence of common genes, we compared the leading-edge genes of each pathway.

Chapter two

We have identified reduced autoimmunity and LPS induced gene expression as two new classes of signature in microgravity. Those functional classes are supported by multiple novel pathways (Table 2.3). Further we have found multiple novel and specific signatures, which suggests the reduced inflammatory response as a functional class (Table 2.5). We have also identified repressed intestinal immune network for immunoglobulin A (IGA) production and cytokine-cytokine receptor interactions, which were not reported before in the context of microgravity. Several other signatures for reduced immunity were also observed. A comprehensive list is tabulated in Table 2.5.

Table 2.3 Microgravity induced functions are supported by multiple molecular pathways.

Function	Supporting up-regulated gene-sets*	Supporting downregulated gene-sets*
Immunity		
Repressed immunity		KEGG intestinal immune network for IGA production; Biocarta inflame pathway; Biocarta cytokine pathway; KEGG cytokine cytokine receptor interaction; PID NFAT TF pathway; Zhou inflammatory response live up; Vilimas Notch1 targets up; Goldrath antigen response; Hallmark TNFA signaling via NFkB; GSE17721 0.5h Vs 4h CPG BMDM up; GSE3982 B cell Vs EFF memory CD4 T-cell up; Biocarta NKT pathway; Reactome second messenger molecules signaling pathway
Reduced inflammatory response		Biocarta inflame pathway; Galindo immune response to enterotoxin; Tian TNF signaling not via NFkB; Seki inflammatory response LPS up; GSE9988 LPS Vs LPS and anti TREM1 monocyte dn; GSE14000 unstim Vs 16h LPS DC up
Repressed autoimmunity signature		KEGG allograft rejection; KEGG autoimmune thyroid disease; KEGG graft Vs host disease; KEGG TypeI diabetes mellitus; Reactome PD1 signaling
Blood cell differentiation signature (inhibition)		Mori mature B lymphocyte up; Oswald hematopoietic stem cell in collagen gel up; Basso CD40 signaling up
Cancer		
Induced liver cancer		Module 75; Module 46
Induced B Lymphoma, diffuse large B cell lymphoma, leukomia	Module 47	Module 6; Module 123; Verhaak AML with NPM1 mutated dn
Induced Lung carcinoid and reduced pro-survival		Phong TNF targets up

Chapter two

ESR positive breast cancer	Doane breast cancer ESR1 dn	
Induction of Head and Neck cancer	Rickman head and neck cancer C	
Oncogenic signature	KRAS300_UP.V 1_UP; KRAS600_UP.V 1_UP	Amit delayed early genes
Ovarian cancer growth		Lu EZH2 targets up
Increased responsiveness to cancer treatment	Heller HDAC Targets silenced by methylation up	Kobayashi EGFR signaling 6hr dn; Becker Tamoxifen resistance up; Peng Rapamycin response dn, Lee liver cancer survival dn
Induction of apoptosis		GSE37416 CTRL Vs 12 h F Tularessis L Vs neutrophil dn; Brocke apoptosis reversed by IL6; Dairkee TERT targets up
Other disease signatures		
Reduced asthma signature	Bosco epithelial differentiation module	KEGG asthma
Reduced diabetes signature	Servitja islet HNF1A targets dn	KEGG Type I diabetes mellitus; GSE9006 healthy Vs type 2 diabetes PBMC at DX up
Psychiatric		Kim all disorders duration corr dn; Stark prefrontal cortex 22Q11 deletion dn
Cellular and metabolic pathways		

Chapter two

Reduced oxidative phosphorylation	Reactome TCA cycle; Reactome respiratory electron transport; ATP synthesis by chemiosmotic coupling and heat produced by uncoupling protein; Module152; Mootha VOXPPOS; Hallmark oxidative phosphorylation	
Reduced post transcription	Module 114; Reactome mRNA processing; LI DCP2 bound mRNA	
Increased SLC-mediated transmembrane transport	Reactome mediated membrane transport	SLC trans-

2.3.3 Novel molecular signatures related to cancer

Radiation induced cancer risks (and death risk) for astronauts in international space station were estimated for various cancers including breast cancer, leukaemia and lung cancer (Cucinotta *et al.*, 2001). Though experiments showed alternation of microRNAs related to cancers in microgravity as discussed in the introduction, till now, no microgravity study on healthy human cells showed induction of cancer signature in gene expression study. We were curious whether there is any cancer signature is embedded in the gene expression data, which were obscured in conventional analysis. We found strong signature on induction of liver cancer and leukaemia as evident by several molecular pathways (under cancer in Table 2.5). In addition, we have identified signatures related to lung cancer, breast cancer, ovarian cancer and head and neck cancer (under cancer in Table 2.5). We have also found induction of oncogenic KRAS signalling pathway (KRAS.300_UP.V1_UP and KRAS.600_UP.V1_UP, Table 2.5 and Table A.2). In space, there is always a possibility of radiation exposure even within ISS or in-flight experiments. Therefore, those cancers related signatures might not be unlikely, where the cells were studied in space (E-GEOD-38836 and E-GEOD-43582). However, it is intriguing that the cell cultured in earth-based microgravity simulator (E-GEOD-4209 and E-GEOD-57418) without any radiation exposure also showed those cancer signature in our analysis (Table A.2). Thus, our study suggests that microgravity alone may induce several cancer related signatures. Further, we have found the signatures about increased drug response on B lymphoma, liver cancer, breast cancer and non-small lung cancer in microgravity (Table 2.5).

2.3.4 Other novel molecular signatures

Our results showed reduction of diabetic signature through multiple gene sets, which includes the upregulation of the repressed genes in pancreatic islet upon HNF1A knock-out (Servitja islet HNF1A targets dn, Table 2.5) and downregulation of type I and type II diabetic genes (KEGG Type I diabetes mellitus, GSE9006 healthy Vs type 2 diabetes, Table 2.5). Similarly, reduced asthma, blood differentiation (under other diseases, Table 2.5) and post-transcription processes are also identified (please see under cellular and metabolic pathways in Table 2.5) as new signatures.

2.3.5 Consensus non-negative matrix factorization (CNMF) clustering and protein network mapping show plausible regulatory and functional connections among genes

To understand the regulatory or functional networks among the leading edge genes, we have clustered them based on their expression patterns using an unsupervised non-negative matrix factorization (NMF)(Lee and Seung, 1999) , coupled with a consensus clustering (NMFC) (Brunet *et al.*, 2004). NMFC is applied on the combined leading edge genes from upregulated and downregulated gene sets separately, from canonical pathway modules. The maximum number of mathematically stable clusters was estimated from the cophenetic coefficient as a function of number of clusters. Figure 2.8 shows such plots for leading edge genes from each experiment. In almost all figure, a sharp and continuous decline is observed after a certain number of clusters (shown by red arrow). This number can be taken as the maximum possible number of mathematically stable clusters. The stable numbers of clusters for each experiment are shown in

Chapter two

Figure 2.7 and cophenetic plots for each experimental condition are shown in Figure 2.8.

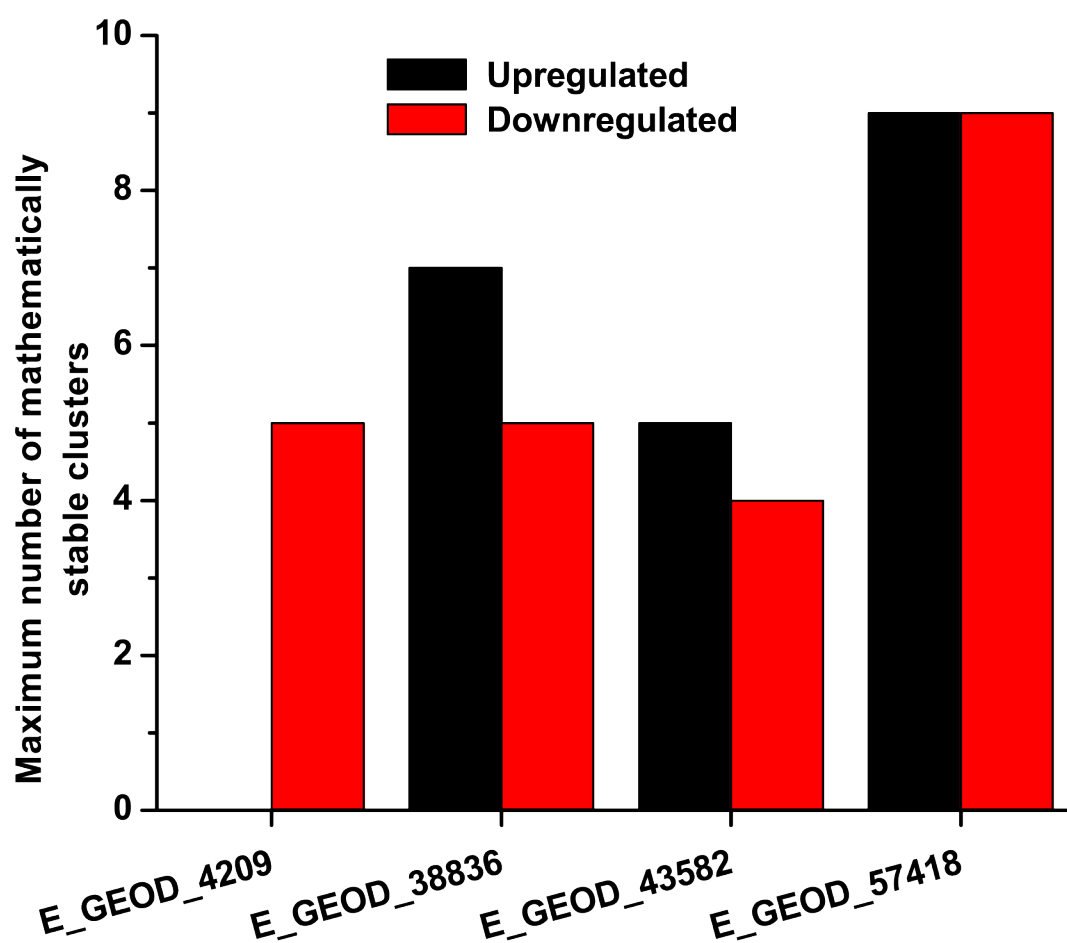


Figure 2.7) Maximum number of mathematically stable clusters plausible for other experiments. Clusters with upregulated and downregulated genes are shown in black and red bars respectively.

Chapter two

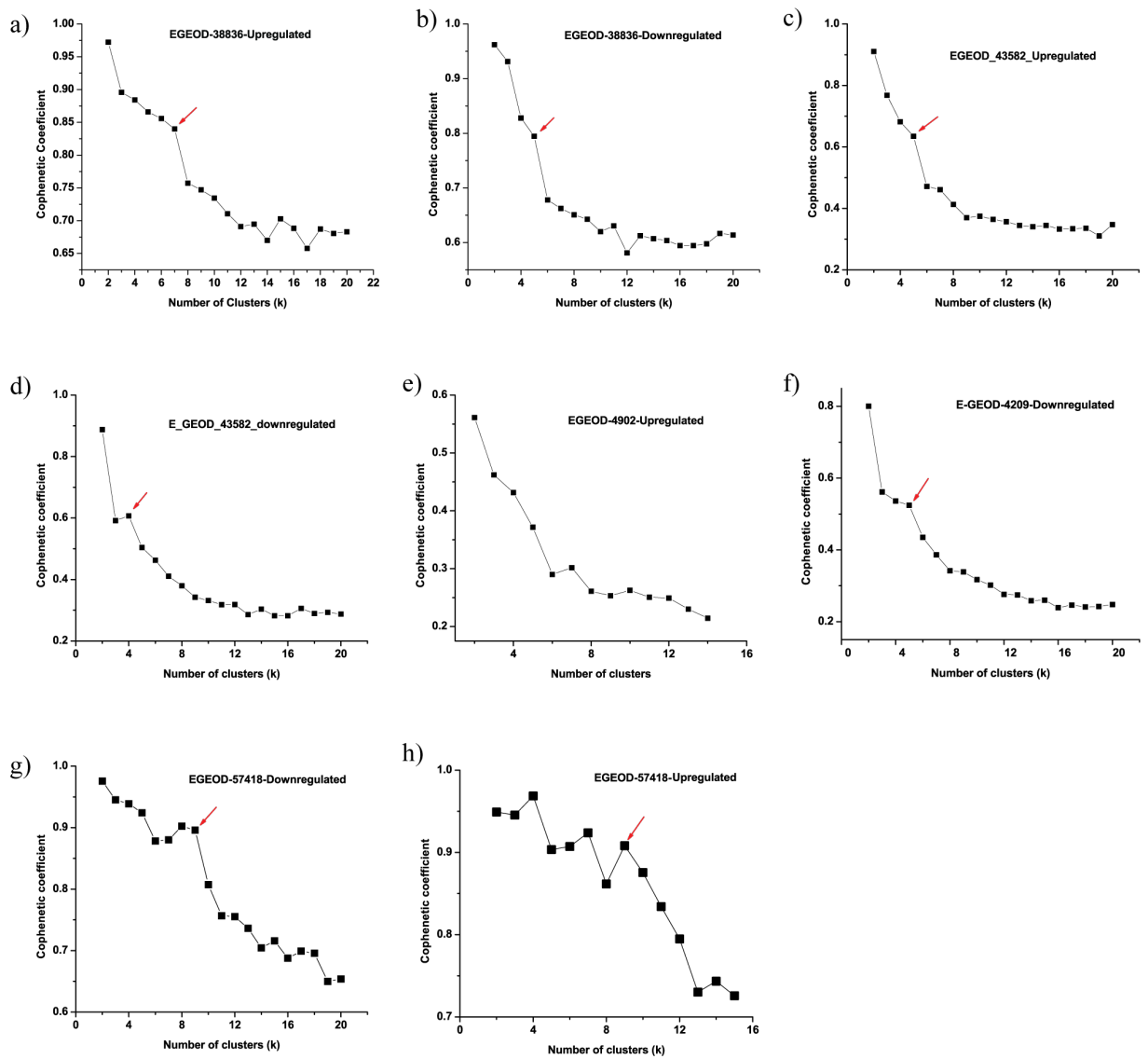


Figure 2.8 Cophenetic coefficient as function of number of clusters for all the experiments.

To gain a better understanding of the regulatory or functional networks among the leading edge genes, we used an unsupervised non-negative matrix factorization (NMF) in conjunction with a consensus clustering (CNMF) to cluster them based on their expression patterns. CNMF is applied independently to the leading edge genes of upregulated and downregulated gene sets within canonical pathway modules. The maximum number of mathematically stable clusters (indicated with an arrow) was determined using the cophenetic coefficient as a function of the cluster number.

The genes from each stable cluster are further mapped on the STRING protein network database. The resultant networks showed densely interactive and localized set of protein-protein interactions. Those interactions were enriched with KEGG and

Chapter two

Reactome pathways (the two constitutive member databases in Canonical Pathways in MSigDB) with a built-in tool in the STRING. Enrichment with p value $< 10^{-5}$ were considered for further analysis. Figure 2.9 represents a set of such PPA networks. The functional annotations with statistical parameters of all the networks are tabulated in Table .

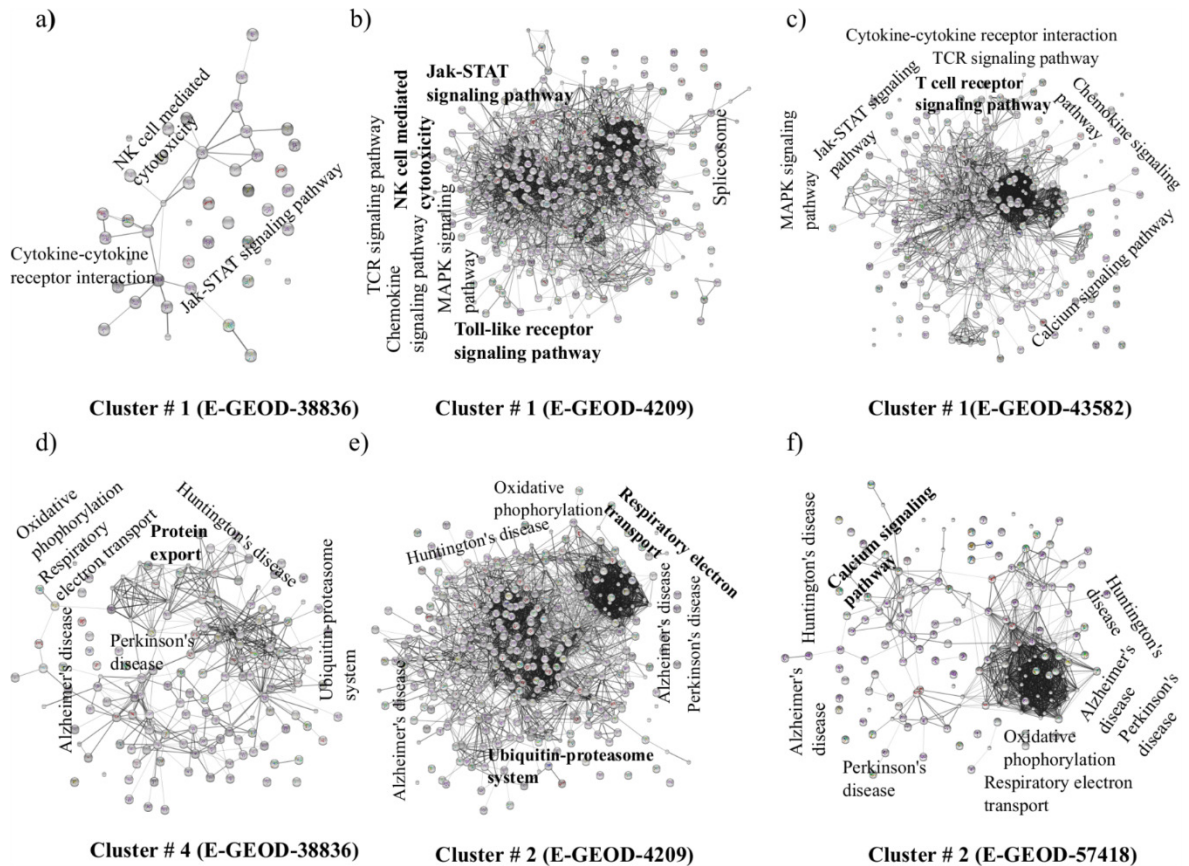


Figure 2.9 Protein-protein association (PPA) networks.

Networks (a–f) represent CNMF clusters from various datasets of downregulated genes from canonical pathways in microgravity. Each node represents a protein and the line connecting the nodes (edge) represents the functional association. The relative thickness of each line signifies the confidence level of such association. The functional annotations of different parts of the network are marked on the figure.

The small but tightly connected PPA networks within a cluster suggest regulatory and/or functional connections among genes and provide more specific and detailed functionality. In PPA networks, we mainly observed i) distributed functions, where same functions are distributed among multiple clusters for an experiment (dataset) and

Chapter two

ii) unique functions, which are confined only within a specific cluster for a dataset. Example of unique functions include Neurotrophin signalling pathway, respiratory electron transport and activation of caspase, which are only found in cluster 1, 2 and 4 respectively, among all the 5 downregulated clusters from E-GEOD-4209 (Table A.4). All unique and distributed functions can be found in Table A.4. The appearance of unique functions among clusters, which resulted from unsupervised clustering, suggests a set of genetic networks respond to microgravity with distinct gene expression patterns. As cell types and experimental conditions are different for the datasets, the functional enrichment of clusters differs among experiments. However, the association of the similar functions, among multiple experiments is of particular interest as they indicate towards plausible regulatory connections in response to microgravity. There are few instances in our results, which demands further attention.

We have already mentioned that GSEA identifies upregulated olfactory transduction pathway in microgravity. However, PPA networks pinpoint it as olfactory receptor G protein trimer complex, which appeared in upregulated clusters of multiple experiments (cluster#6 E-GEOD-38836, cluster#1 and 4 E-GEOD-43582, Upregulated Canonical Pathways, Table 2.7). Next, a network level association between Jak-STAT, cytokine-cytokine receptor interaction, TCR signalling pathway, NK cell mediated cytotoxicity and chemokine signalling is appeared in multiple datasets (Figure 2.9 a–c).

Immunology studies, not related with microgravity, showed that the cytokine-cytokine receptor interaction and Jak-STAT signalling pathway are directly connected through a cascade of biochemical reaction (O’Sullivan *et al.*, 2007). Hyper-activation of this cascade results inflammation and asthma (O’Sullivan *et al.*, 2007). It is interesting that this proportional relation is observed in our analysis and those disease pathways

Chapter two

are found to be downregulated. This functional network interface was never reported in the context of microgravity. Further, a close association among oxidative phosphorylation/respiratory electron transport, neurodegenerative conditions (Alzheimer's disease, Huntington's disease and Parkinson's disease) and ubiquitin-proteasome system are appeared in clusters of three experiments (Figure 2.9 d–f). Oxidative phosphorylation/respiratory electron transports are shown related with neurodegenerative condition (Hroudová, Singh and Fišar, 2014) and with ubiquitin-proteasome system (Ross, Olson and Coppotelli, 2015) but togetherness of those three functions with similar gene expression patterns never reported in response to microgravity. Thus, the interface among various functions (Figure 2.9 a–f) within a genetic network may guide us towards better understanding how microgravity influence several of the cellular functions through a small set of genes and warrant further investigations.

2.4 Discussion

An important aspect of this pipeline is to identify microgravity induced pathways with statistical confidence and presence of multiple signatures for a functional class. In this way the reduction of generation of second messenger molecules, which was merely suggested by one study (Girardi *et al.*, 2014), appeared as a statistical significant pathway in all the 4 datasets in our analysis along with multiple new pathways as associated molecular signatures (shown in the result section). Microgravity generates low shear stress around the cells in liquid media (Tsao *et al.*, 1994; Nickerson *et al.*, 2004). It was particularly found that low shear stress (<1 dynes/cm²) reduces the amount of second messenger molecule IP₃ by 20% (Prasad *et al.*, 1993) and also influence the Ca²⁺ mobilization (Wiesner, Berk and Nerem, 1997) in mammalian cells. Therefore, the inhibition of immune systems may directly originate from microgravity induced low sheer

Chapter two

stress on the cells through the reduced amount of second messenger molecules (all 4 studies showed this new pathway in our analysis), which negatively influence the NF- κ B pathway. In addition, the PPA network shows the reduction of Jak-STAT signalling pathway (Fig. 5a–c), which has a proportional relationship with fluid shear stress (Licorresponding, 2013). This reduction may directly be related with the reduced inflammation and asthma through cytokine-cytokine receptor interaction (O’Sullivan *et al.*, 2007) and all of those pathways were found downregulated in our results. Based on this, we hypothesize that the low shear stress in microgravity is responsible for partially impaired immunity.

Our new results support functional observations from other studies, where no gene expression analysis was performed and shed some light on couple of microgravity-induced issues, where the genetic basis is still unknown. This analysis showed reduction of the several LPS induced gene expression signature (See Table 2.5 under Reduction in LPS induced gene expression and immunogenic signature in Table A.2). Several of the cytokines production was shown reduced after LPS stimulation in microgravity (Crucian, Raymond P. Stowe, *et al.*, 2015). This analysis showed the presence of several signatures of autoimmune repression (Table 2.5), which were never reported. A recent study speculated microgravity could be a measure for fighting autoimmune disease (Verhaar *et al.*, 2014b). Further, tissue engineering model in microgravity showed the enhanced survival, better secretory profile and better insulin normalization of beta islet cells compare to ground control (Barzegari and Saei, 2012). Several of the reduced diabetic signatures appeared in our results (Table 2.3). The change in the smelling behaviors among astronauts in space was documented and attenuation of olfactory components as a result of microgravity induced upward shift of body fluid was postulated as

Chapter two

plausible reason (Olabi *et al.*, 2002). Our results identify the upregulation of olfactory signal transduction pathways, suggesting molecular genetic origin of changed smelling behaviour in space.

Apart from the new and overlapping pathways, we have showed that how gene expression patterns are possibly connected with functional behaviour at microgravity. An emphasis was given to extract similar types of gene expression patterns across the dataset. To our knowledge, this is the first study in this direction. We were able to show i) association between Jak-STAT, cytokine-cytokine receptor interaction, TCR signalling pathway, NK cell mediated cytotoxicity and chemokine signalling and ii) a close association among oxidative phosphorylation/respiratory electron transport, neurodegenerative conditions (Alzheimer's disease, Huntington's disease and Parkinson's disease) and ubiquitin-proteasome system across multiple datasets (see result section for details). We have shown from the literature that few of those pathways are connected in human cells but never reported in context of microgravity. Those patterns (cluster/gene network) in gene expression are derived from unsupervised mathematical analysis without imposing any prior functional knowledge. Thus, the identification of common functional interfaces from those gene networks across multiple data-sources suggests the plausible functional and regulatory connections among genes in microgravity and how microgravity may influence several functions through a small set of genes.

In this analysis the data sets were taken for 3 different cell types. Microgravity may work differently on different cell types and experimental conditions. This is evident in our results as major fraction of the pathways altered for each dataset are non-overlapping (Figure 2.5). However, the question remains that what could be the minimum

Chapter two

set of cellular pathways, if any, influenced by the microgravity irrespective of the cell types. Our results shed some light in this aspect. Though the four studies were performed in different experimental conditions with different cell types, microgravity conditions (simulated and spaceflight), media, fluid shear around cells, hydrostatic pressure, temperature, aeration, reactor and different experimental groups, around 100 pathways overlap among at least 3 (out of 4) experiments. Apart from the overlapping pathways, we further identified, across the datasets, common interface of multiple functions in gene networks. Those gene networks were derived from mathematical pattern recognition (unsupervised clustering) without any prior functional knowledge. Thus, the common functional interface within a gene network (cluster) across multiple datasets, suggest that the overlapping pathways identified from various cell types are not arbitrary. Those examples at least suggest that multiple pathways are perturbed in microgravity by similar way, irrespective of cell types. However, similar pathway may not lead to the similar functions in all cell types and should be validated by functional experiments. The previous studies mostly suggested the microgravity-induced pathways without statistical confidence or a set of broad functional categories missing the specific details. This approach prevents pinpointing a pathway from several others and gathering helpful details for future experiment design. Our analytical pipeline has generated a set of specific and testable hypotheses, which are subjects of near future experimentation. The analysis predicts the i) reduction of second messenger molecules as a fundamental reason for T cell regulation dysfunction and ii) the increased effectiveness of the drugs (specifically Tamoxifen and Rapamycin, see Becker Tamoxifen resistance up; Peng Rapamycin response dn under cancer in Table 2.3 on cancers in microgravity. Two important future experiments may include i) measuring T cell regulation in T cells as a

Chapter two

function of amount of second messenger molecules and the associated gene expression in microgravity and ii) the effect of drugs on various cancer cell types in microgravity. Further our results suggest that in microgravity, the mRNA processing gets partially impaired (Table 2.3), which can be tested on varieties of cell types to understand the effect of microgravity on fundamental cellular processes.

2.5 Conclusion

Here we represent an analytical pipeline, which gives an integrated molecular systems level picture of healthy human cells under microgravity. Adequate cellular and human level data for assessing the health risks for interplanetary travel are not available till date (Setlow, 2003). Enough human level data may not be available in near future as the number of astronauts flown to space is low. Therefore, assessment of health risks from molecular signatures may serve as a key criterion for future manned space mission. In addition, experiments in space are costly business. Therefore, getting maximum level of insight from a space experiment is of a crucial importance. Our approach identifies statistically significant pathways directly from the gene expression data, even in the low fold change regime, which cannot be possible in conventional differential gene expression analysis followed by passive pathway mapping. The pan molecular pathways analysis, encompassing almost all known aspects of human disease and immunity, followed by comparative analysis identifies a set of new altered molecular signatures, which are appeared across studies. Most functions are supported by multiple pathways and associated (downstream) molecular pathways. Thus, it gives a high degree of confidence. Further, unsupervised clustering suggests the plausible regulatory connections among altered genes in microgravity. Our results suggest a set of specific hypotheses, which can be tested directly in an earth-based microgravity simulator or in space flight

Chapter two

condition and may help assessing risks and developing new medicine for microgravity induced health hazards.

2.6 Methods

2.6.1 Gene set enrichment analysis (GSEA)

GenePattern, a platform for multiple genomic analyses was used for performing Gene Set Enrichment Analysis (GSEA) (Reich *et al.*, 2006). The module was run with global gene expression datasets in an appropriate file (.GCT) format. A separate file (.CLS) was developed to differentiate the gene expression data between 1g and microgravity. The molecular pathways from MSigDB are linked with GenePattern server. First, we generated a ranked list of genes (Fig. 1a) by applying ‘signal-to-noise’ metric, which can be obtained by $(A_{\text{microg}} - A_{1g}) / (SD_{\text{microg}} + SD_{1g})$. Here, A_{microg} , A_{1g} represent the average expression of genes in microgravity and normal gravity condition, respectively. SD_{microg} and SD_{1g} represent standard deviation associated with microgravity and normal gravity gene expression data. A running sum statistics (GSEA algorithm) with a weighted Kolmogorov-Smirnov like scoring scheme (Subramanian, Tamayo, Vamsi K. Mootha, *et al.*, 2005) was applied on the ranked list to determine if the genes from a gene set were distributed either top or the bottom of it. This statistic estimates a score (enrichment score or ES) for each gene set, which increases when a gene from a gene set hits the same gene in the ranked list. The score is evaluated by running down the ranked gene list (RGL) and if genes from a particular gene set (GSet) gives positive hit in the RGL, the score would increase, otherwise decrease. The probabilities of positive hit and negative hit (miss) can be estimated from the following equations.

$$P_{+hit}(GSet, i) = \sum_{\substack{gj \in GSet \\ j \leq i}} \frac{|rj|^p}{N_R}$$

Equation 2.1

$$P_{-hit}(GSet, i) = \sum_{\substack{gj \notin GSet \\ j \leq i}} \frac{1}{N - N_H}$$

Equation 2.2

where, gj denotes the j th gene in $RGL = \{g_1, \dots, g_N\}$, $r(gj) = r_j$ $p = 1$ and

$$N_R = \sum_{j \in GSet} |rj|^p$$

The highest deviation of $P_{+hit} - P_{-hit}$ from '0' gives the estimation of enrichment score (ES). The GSet permutation type was set to a value of 1000, i.e. the GSets would be permuted randomly 1000 times among all gene sets (molecular pathways) to estimate the statistical significance (null hypothesis testing). The false discovery rate (FDR-q value) was estimated from the distribution of the p values. GSEA analysis was run for atleast three times to check the consensus in ES, NES, p value and q value among various runs. The molecular pathways in MSigDB are manually checked in detail for its interpretation in the context. One category in MSigDB is cancer modules. Every module is related with its clinical annotation through its module map in (<http://robotics.stanford.edu/~erans/cancer/index.html>). For each of the module, we have assigned the clinical annotation, when it is highly significant ($p < 10^{-10}$) and when the minimum hit of the module genes on cancer pathways is atleast 50%.

2.6.2 Leading edge gene analysis

Each common statistically significant pathway for individual experiments also informs about the genes, which are most significant (high NES value) in a given pathway. The genes before the peak (highest NES value) are the leading edge genes. Those genes were extracted from the GSEA results for further analyses.

2.6.3 Clustering of leading-edge genes with non-negative matrix factorization consensus (NMFC)

The consensus non-negative matrix factorization (NMFC) clustering is run in GenePattern. It requires only positive entry in gene expression data table. Therefore, we checked for negative elements on each expression sets for NMF and we found none. The expression data of leading-edge genes combining all statistically significant gene sets from canonical pathways, which is a collections of 1330 gene sets in MSigDB, were chosen for NMFC run. The upregulated and downregulated leading-edge genes have been run separately. Based on the gene expression patterns of the multiple cell samples between microgravity and normal gravity, NMF clusters the genes. NMF consider the gene expression dataset as a positive matrix M of the size $R \times C$, with N number of clusters. Then its iteratively computes for matrices Y and Z so that $M = YZ$, with $Y \times N$ and $N \times Z$ sizes. In each step, iteration was updated to find a minimum in an appropriate function. The details of the NMF can be found in refs 25,26. The iteration number per clustering was set as 2000 times. On top of the NMF, a consensus clustering was run to evaluate the mathematical stability of the NMF clusters. We sequentially ran the consensus clustering by setting the number of clusters from 2 to 20. The cophenetic coefficient associated with each distribution of the clusters was estimated and the rate of the

change of this coefficient estimated the highest number of plausible stable clusters. The cluster, after which the cophenetic coefficient drops sharply, denotes the maximum possible number of mathematically stable clusters. The associated genes for each stable cluster were tabulated (Table A.3) for further analysis.

2.6.4 Protein-protein associated networks of clustered genes

The genes from each stable cluster were mapped on STRING network database (version 9.1) to study protein-protein association networks. The list of genes from each cluster were uploaded and assigned *Homo sapiens* as target organism on STRING. Further the enrichment score and associated p values (Table A.4) for functional annotation from KEGG and Reactome pathway database were determined from the in-built functional tools. As KEGG and Reactome databases are common in STRING and canonical pathways in MSigDB, we have chosen canonical pathways for NMFC.

Chapter 3: A microgravity responsive synthetic genetic device in *Escherichia coli*

3.1 Introduction

Historically, space exploration is driven by cultural, scientific, and political motivations (Huntress *et al.*, 2006). Adding to this reason is the projected human population in the near future by UN (*No Title*, no date c); it is estimated that in five billion years, our Sun will enter its red giant phase and will engulf nearby planets, the increased temperature will be enough to boil oceans and kill life (*The sun won't die for 5 billion years, so why do humans have only 1 billion years left on Earth?*, no date; Schröder, Smith and Apps, 2001; Schröder and Connon Smith, 2008). Theoretical physicist, using their understanding of the quantum theory of gravity, has predicted that sometime in the distant future our universe will undergo a phase transition that will destroy us and everything else around us instantaneously, they recommended spreading of civilization on different worlds in different parts of the universe to escape complete annihilation (Sen, 2015). Agencies like NASA, ESA as well as private companies like SpaceX are working towards a common goal of becoming a multi-planetary species (Vernikos *et al.*, 2016; Cichan *et al.*, 2017; Musk, 2017). Deep space travel is resource intensive, thus enormous logistical costs of launching and resupplying resources from Earth is a major limitation (Huntress *et al.*, 2006; Montague, George H McArthur, *et al.*, 2012). Therefore, the focus is on research and development of newer and alternative technologies for sustaining long duration manned missions. Synthetic biology is being considered as the feasible technology for augmenting human space

Chapter three

travel. Although biology offers plethora of resources, none of these resources are adapted to the requirement of deep space or other planets.

In a workshop held on 2010, NASA discussed the potential role of synthetic biology in its future missions (Langhoff, Cumbers and Worden, 2011). It highlighted various problems which can be addressed using synthetic biology, a) biological in situ resource utilization (ISRU) for ISS, lunar and Mars surface, b) biosensing capabilities, c) biomaterials, d) human health, e) life support. It is projected that using synthetic biology for in-situ resource utilization in space, bioproduction of food, fuel and other essentials, waste recycling and other applications (Figure 3.1) (Menezes, Montague, *et al.*, 2015), will significantly reduce payload mass (Nickerson *et al.*, 2004; Montague, George H. McArthur, *et al.*, 2012; Menezes, Cumbers, *et al.*, 2015). Various ways in which biological in situ resource can be implemented on lunar regolith and Martian soil have been discussed in details in multiple studies (Montague, George H. McArthur, *et al.*, 2012; Menezes, Cumbers, *et al.*, 2015).

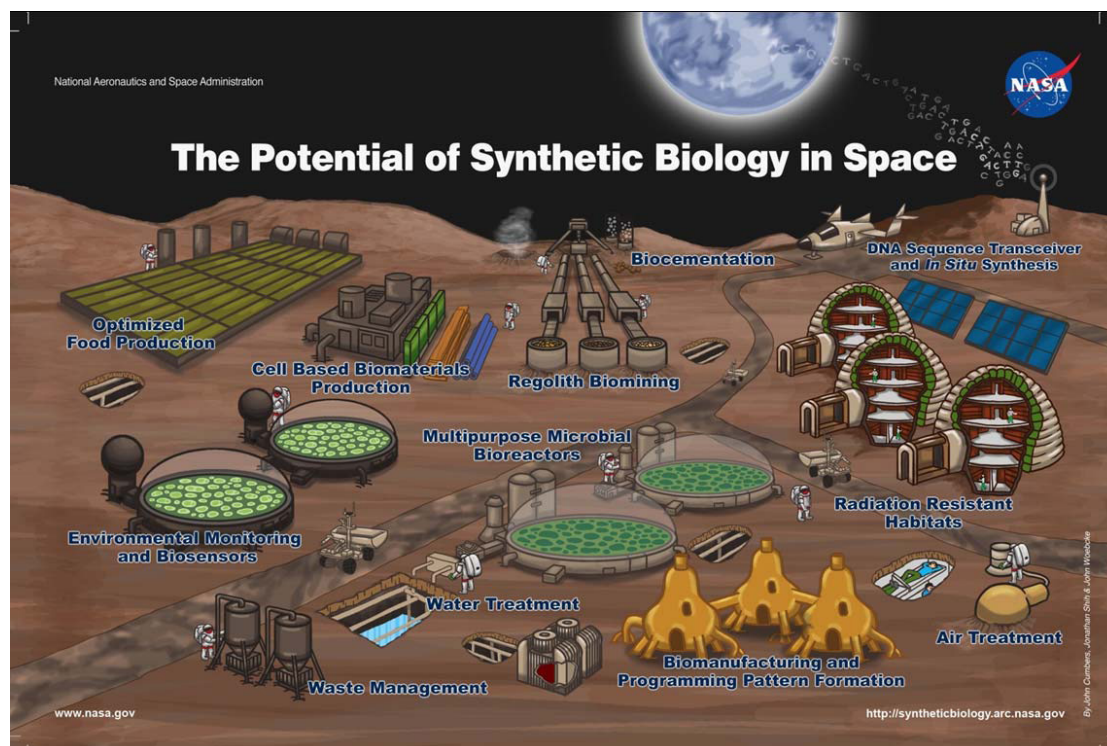


Figure 3.1 Potential application of synthetic biology in space.

Synthetic biology can be used for sustaining extra-terrestrial colonies on Moon, Asteroids or Mars by generating essential resources using local raw materials and maintaining life support systems. (reproduced from (Robertson, no date) with permission from Dr. John Cumbers.)

Apart from that several synthetic biological applications relevant to spaceflight have been suggested, which is mainly using engineered organisms for Environmental Control and Life-Support Systems (ECLSS), which focuses on urine and solid waste processing (Langhoff, Cumbers and Worden, 2011; Montague, George H McArthur, *et al.*, 2012); using biological fuel cell to generate electricity while purifying air and water ('NASA - Synthetic Biology and Microbial Fuel Cells: Towards Self-Sustaining Life Support Systems', no date); food production using photosynthetic bacteria (Figure 3.2), which can be improved by synthetic biology not only in terms of nutrition but taste as well (Flynn, no date; Langhoff, Cumbers and Worden, 2011; Menezes, Cumbers, *et al.*, 2015).

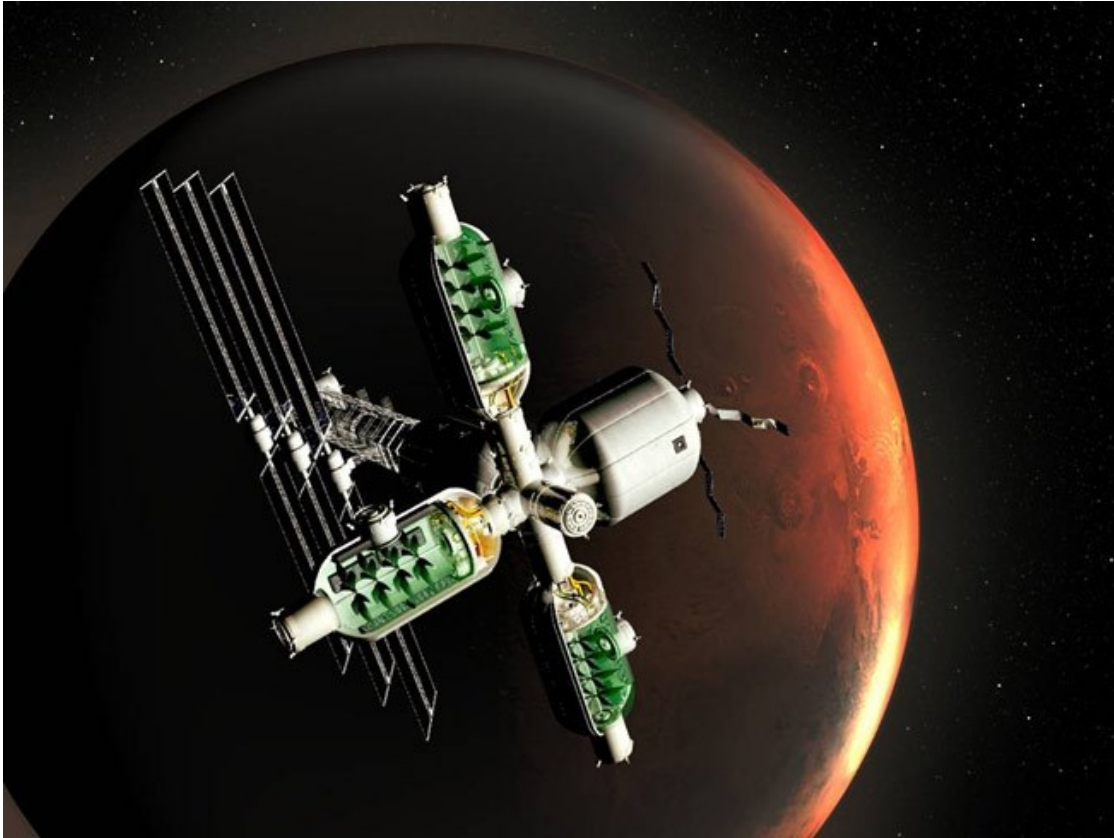


Figure 3.2 Projected photosynthetic bioreactors for food production.

Engineered photosynthetic bacteria (e.g., *Synechococcus*, spirulina) can be used as source of food for long term manned spaceflight. (Image reproduced from (*Synthetic life could make trips to Mars more comfortable*, no date) with permission from Dr. John Cumbers.)

3.1.1 The need of a microgravity responsive synthetic genetic device.

It is estimated that, in a 916 day return trip to Mars, around 410 days will be spent on the spaceflight itself (Drake, 2009). Therefore, it is important to consider the effects of microgravity on astronauts as well as microorganisms, which are a popular candidate for synthetic biology applications. Microgravity influence biological processes at the molecular, cellular and organism level, including astronauts' health and immunity (in previous chapter) (Planel, 2004; Wilson *et al.*, 2007; Mukhopadhyay *et al.*, 2016; Çelen *et al.*, 2019) Studies have been performed on multiple organisms to understand how they sense microgravity at the cellular and molecular level (INGBER,

Chapter three

1999; Wilson *et al.*, 2007; Zheng *et al.*, 2015; Mukhopadhyay *et al.*, 2016; Roy, Shilpa and Bagh, 2016; Häder *et al.*, 2017; Kohn, Hauslage and Hanke, 2017; Çelen *et al.*, 2019) Several mechanisms have been proposed including the presence of gravireceptor (Häder *et al.*, 2017), the effect of low shear stress (Nickerson *et al.*, 2004; Wilson *et al.*, 2007), altered transport phenomena due to microgravity induced low fluid shear (Nickerson *et al.*, 2004), alteration of membrane fluidity (Kohn, Hauslage and Hanke, 2017) and the change in second messenger molecules (Chapter 2). However, no definite mechanistic details were known.

One of the challenging requirements for space synthetic biology is to create engineered biochemical systems to integrate microgravity as a physical signal within molecular and cellular processes. The ability to integrate microgravity in cellular process will not only help in engineering biology for diverse application related to space-flight but may also help in understanding bacterial response to microgravity. Engineered bacteria with synthetic genetic circuits have been shown to sense and integrate various environmental as well as intracellular chemical signals and compute complex human-designed functions through designed molecular interactions and reactions (Saltepe *et al.*, 2018; Gupta *et al.*, 2019). Apart from the chemical signals, cellular devices have been created to sense and integrate physical signals like temperature (Bagh *et al.*, 2011a) and light (Olson and Tabor, 2014; Liu *et al.*, 2018). However, no molecular or biological microgravity responsive device has been created. The aim of this study is to create a synthetic genetic device in *Escherichia coli* (*E. coli*), which integrates microgravity as a physical signal with biochemical processes in a human designed way and responds by changing the expression of a target protein.

3.1.2 Experimental setup

To create simulated microgravity in the lab, we used rotary cell culture system (RCCS). It is an important microgravity laboratory apparatus, designed by NASA (Nickerson *et al.*, 2000). It consists of High Aspect Ratio Vessels (HARV), which are 50 ml bioreactors, four such bioreactors can be mounted on a rotator base, the rotation speed is adjustable from an external power supply unit. The HARV's along with the rotator base is kept in an incubator at required temperature (Figure 3.3). Oxygen supply and carbon dioxide removal is achieved through a gas permeable silicone rubber membrane.

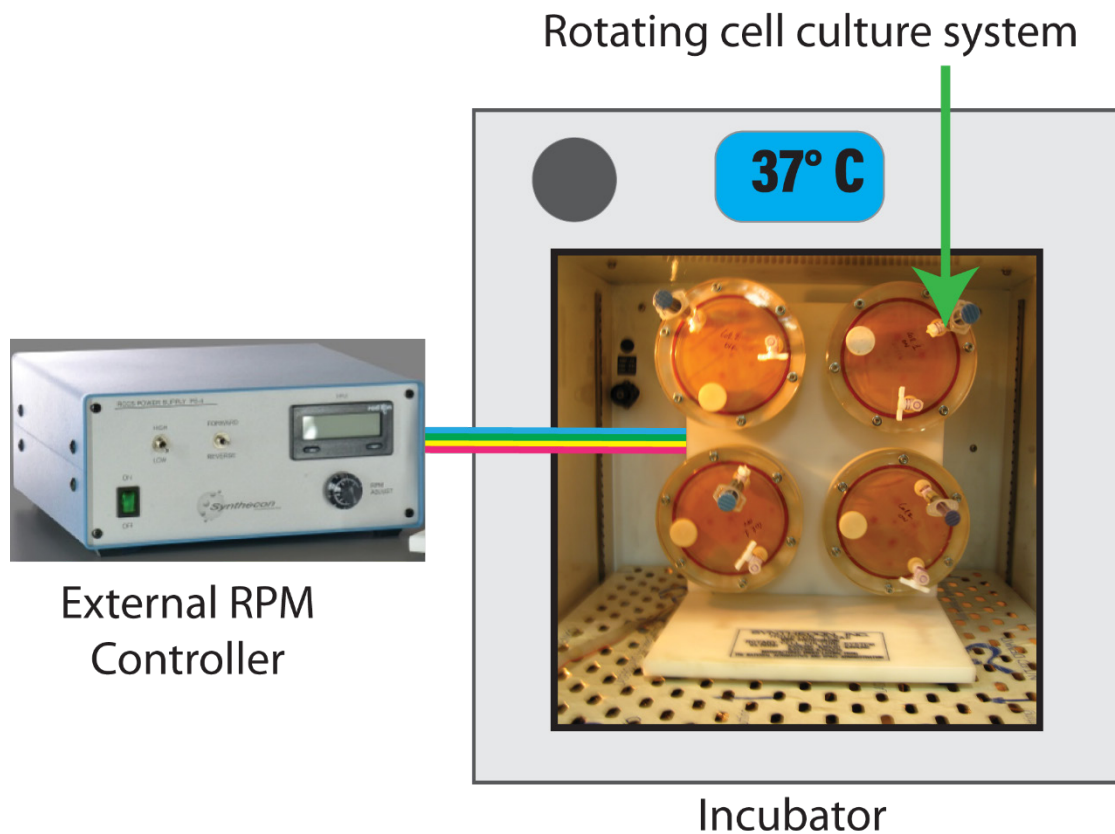


Figure 3.3 Experimental setup for *E. coli* culture in simulated microgravity condition.

The Rotating Cell Culture System is placed in incubator at 37°C, on the rotator base, four HARV's completely filled with inoculated media devoid of air bubbles can be rotated at a time, the rotation speed of HARV's is controllable using external power unit.

Chapter three

When the high aspect ratio rotating wall vessels are positioned in such a way that the axis of rotation was perpendicular to the gravity vector (Figure 3.4 a), the rotation offsets the effect of sedimentation of the cells of the gravity force. Thus, the cells are in continual suspension and in free-fall condition. This mimics the microgravity (Wilson *et al.*, 2002; Gilbert *et al.*, 2020). On the other hand, when the axis of the rotation was parallel to the gravity vector (Figure 3.4 b), this did not offset the gravity force and allowed the cells to settle down. This position provided a normal gravity environment and was used earlier in multiple studies (Wilson *et al.*, 2002; Gilbert *et al.*, 2020). Therefore, we used the same protocol to simulate the microgravity and the normal gravity condition in our experiment (Figure 3.4 a and b). The Rotating Cell Culture System is placed in incubator at 37°C (Figure 3.3). The HARV's are rotated at 25 rpm, this was chosen as multiple earlier studies (Wilson *et al.*, 2002; Gilbert *et al.*, 2020) used this.

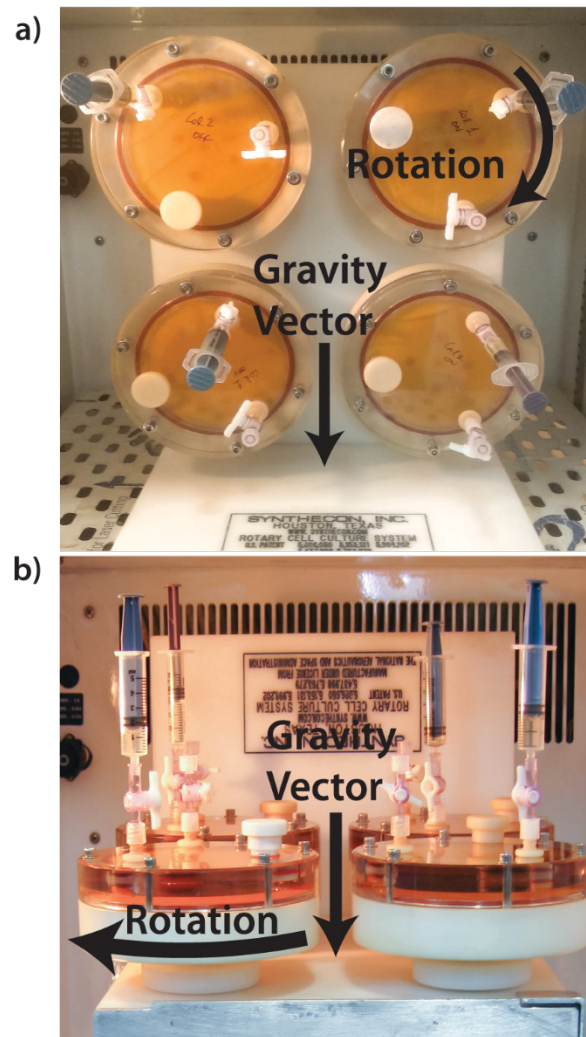


Figure 3.4 Configuration of the HARV's

a) The volume of each of the RWV reactor is 50 mL. In vertical condition, they rotate in the shown direction. Thus, the cells within the reactors are always in free fall condition mimicking microgravity (μG) b) For the Earth's gravity control (1G) the reactors are placed in horizontal position.

3.2 Results

3.2.1 Design and fabrication of microgravity responsive synthetic genetic device

Though the microgravity brings prominent phenotypic changes in organisms including bacteria, the molecular genetic studies showed that the microgravity-induced alterations

Chapter three

of expression of individual genes at both the transcript and the protein level were low (Mukhopadhyay *et al.*, 2016; Roy, Shilpa and Bagh, 2016). Here we targeted a small change in specific biomolecules in response to microgravity and processed that signal through a molecular signal-processing device to get a bigger response and better control.

In bacteria, there exists a family of RNA regulators which are generally short transcripts (~50 – 300 nucleotides) that interact and regulate their targets mRNAs, these are known as small RNA or sRNA. The most studied type of sRNA is trans-encoded sRNA, which base pairs at or near the ribosome binding site (RBS) of their target mRNA and hinders translation by occluding ribosomes (Storz, Vogel and Wassarman, 2011). In Gram-negative bacteria, the RNA binding protein Hfq is usually required for the function and/or stability of this family of sRNAs (Storz, Vogel and Wassarman, 2011) and its role in bacteria has been recently reviewed (Gottesman and Storz, 2011; Vogel and Luisi, 2011). Protein Hfq binds with a specific scaffold made with small regulatory RNAs (srRNA) and the resultant complex binds with target mRNAs to degrade it in bacteria including *E.coli* (Na *et al.*, 2013). Using this principle, synthetic srRNA (SynsrRNA) based systems were created to repress gene expression (Na *et al.*, 2013). The design of Synthetic sRNAs has two parts: a scaffold sequence and a target-binding sequence. In *E. coli*, naturally occurring sRNAs have a consensus secondary structure that serves as a scaffold for binding the Hfq protein, which aids in sRNA-target mRNA hybridization and mRNA degradation. MicC, MicF, DicF, Sgrs etc. are some of the examples of such scaffold (Na *et al.*, 2013). In this study we used MicC scaffold because of its repression capability was superior than most of the scaffolds in *E.coli* (Na *et al.*, 2013).

Chapter three

Interestingly, in recent studies it was found that the intracellular quantity of Hfq within bacterial cells including *Salmonella*, *Pseudomonas*, and *E.coli* was reduced in microgravity (Wilson *et al.*, 2007; Roy, Shilpa and Bagh, 2016; Aunins *et al.*, 2018).

To test if the Hfq was down regulated in *E.coli* cell strain DH5 α Z1 (Lutz and Bujard, 1997a), chassis of our synthetic genetic device, we compared the expression of Hfq mRNA in *E.coli* DH5 α Z1 between microgravity and Earth gravity. It showed 50% less Hfq mRNA expression in microgravity compared to the earth's gravity control (Figure 3.5).

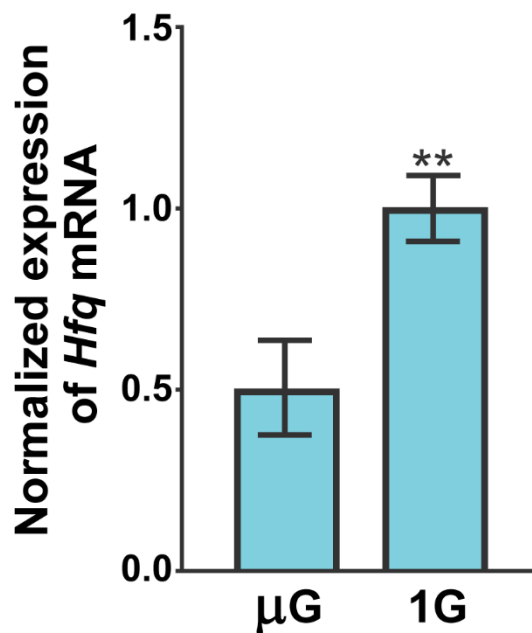


Figure 3.5 Relative Hfq mRNA expression level in microgravity (μ G) and 1G in *E.coli* cell strain DH5 α Z1.

E.coli cell strain DH5 α Z1 was grown in microgravity and earth gravity at 37°C, 25 RPM for 24 hrs in HARVs according to setup shown in Figure 3.4. Three biological replicates were used for each of the gravity conditions. Total mRNA was extracted and qPCR was performed for Hfq and 16sRNA (Housekeeping). At least three biological and technical replicates were used for qPCR. Error bars represents standard deviation from the mean. Unpaired t-test was used to measure the extent of statistical significance where (**) means P value lies between 0.001 to 0.01).

We connected those two facts, as our design hypothesis. The decrease in Hfq protein in *E.coli* DH5 α Z1 in microgravity would alter the repression capability of a

Chapter three

designed SynsrRNA against a target protein expression compare to the earth's normal gravity. In that direction, we designed a SynsrRNA based molecular network (Figure 3.6), which might work as a microgravity sensing and responsive device in *E. coli*.

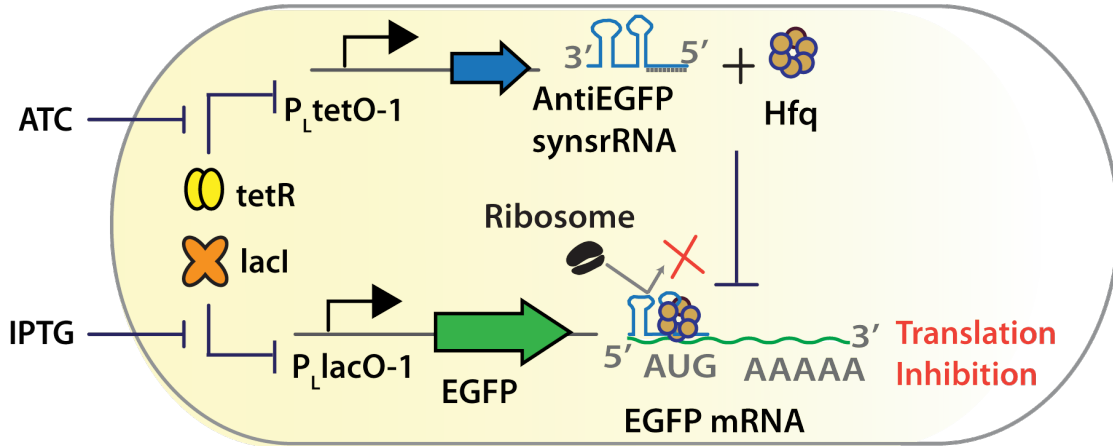


Figure 3.6 The design of the microgravity sensor circuit with EGFP.

The synsrRNA is expressed under ATC inducible promoter $P_{LtetO-1}$, the target protein is expressed under IPTG inducible promoter $P_{LlacO-1}$, translation inhibition of target mRNA by synsrRNA is mediated by Hfq protein constitutively expressed in *E. coli* cell strain DH5 α Z1. Tet Repressor(tetR) and Lac inhibitor(lacI) is constitutively expressed in DH5 α Z1. This device design was incorporated in two plasmid L_EGFP_AC and T_antiEGFP_synsrRNA_UA.

Our designed SynsrRNA against enhanced green fluorescence protein (EGFP) mRNA was around 130 bp long (Figure 3.7) and its expression was under the control of an anhydrotetracycline (ATC) induced promoter pLtetO-1 (Lutz and Bujard, 1997a). The designed SynsrRNA consisted of an 87 bp MicC scaffold for Hfq binding (Na *et al.*, 2013) and a 40 bp sequence to bind the mRNA of the target enhanced green fluorescent protein (EGFP). The DNA sequence of the full SynsrRNA has been shown in Figure 3.7. The EGFP is our machine-readable reporter protein, which was under the control of an Isopropyl β -D-1-thiogalactopyranoside (IPTG) regulated pLacO-1 promoter. The lacI and tetR proteins, which repress the pLacO-1 and pLtetO-1 promoter

Chapter three

respectively, were produced constitutively in the *E.coli* cell strain DH5 α Z1, the chassis of our device. Therefore, in presence of the IPTG only, the EGFP will produce and would give a high fluorescence signal. However, in the simultaneous presence of ATC and IPTG, the anti-SynsrRNA against EGFP mRNA would produce and reduce the EGFP expression and fluorescence signal. The device was fabricated according to the design and incorporated in two different plasmid vectors, where plasmid L_EGFP_AC (Figure 3.19 a) carries EGFP under IPTG inducible pLlacO-1 promoter in low copy (p15Ori) and plasmid T_antiEGFP_synsrRNA_UA (Figure 3.19 b) carries SynsrRNA against EGFP mRNA under ATC inducible pLtetO-1 promoter in a very high copy (PUC Ori). Both the plasmids of our system were transformed into DH5 α Z1 cells.

5'**gcaccaccccggtgaacagctcctcgcccttgctcaccatTTTCTGTTGGGCCATTGCATT-GCCACTGATTTTCCAACATATAAAAAGACAAGCCCGAACAGTCGTCCGGGCTTTTTTTCTCGAGCT**3'

Figure 3.7 Anti-EGFP synthetic small regulatory RNA encoding synthetic DNA sequence.

The sequence highlighted in green is the target (EGFP mRNA) binding region (40bp) and the sequence highlighted in yellow is the MicC scaffold region (87bp), which helps to bind Hfq protein.

3.2.2 The fabricated synthetic genetic device responds to microgravity by altering expression of a fluorescence protein

Next, the individual colonies of the engineered cells were subject to overnight growth with appropriate antibiotics at 37°C in a shaker incubator. Overnight cell culture from a single colony was divided into two groups. One group of cells was treated with both the inducers ATC and IPTG. The other group of cells was treated with IPTG only. The overnight culture was diluted 100 times and re-suspended in microgravity simulator

Chapter three

reactors filled with fresh media with appropriate antibiotics and inducers. The cells in the simulators were grown for 24 h at 37°C, collected, washed, re-suspended in PBS (pH:7.4) and the expression of EGFP was measured in a multimode micro-plate reader. We calculated (see section 3.4) the fold changes in EGFP expression between IPTG only (express EGFP but not anti-EGFP SynsrRNA) and IPTG + ATC condition (express both EGFP mRNA and anti-EGFP SynsrRNA). A substantial difference in fold change (28.2) values between microgravity and normal gravity was observed (Figure 3.8).

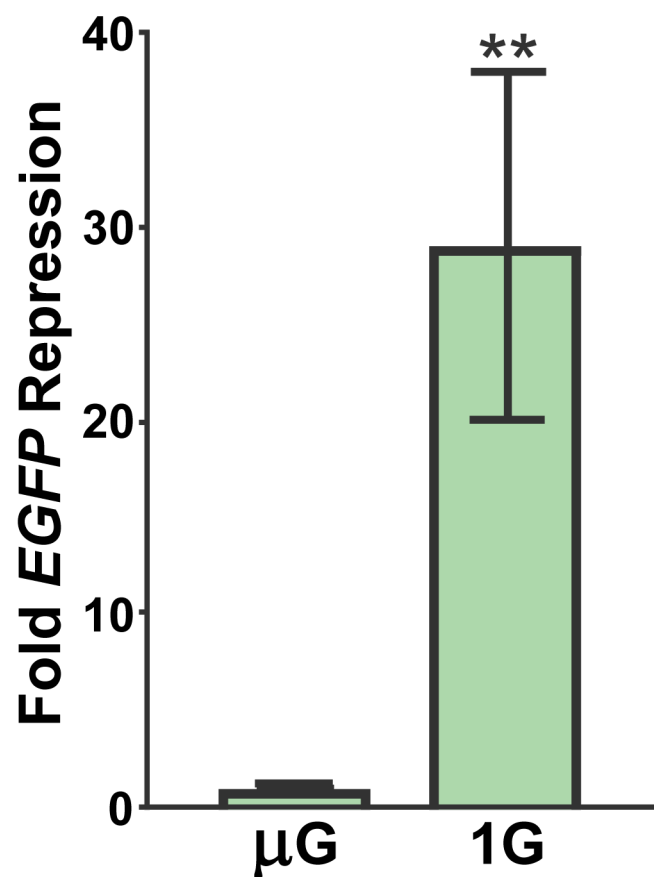


Figure 3.8 The experimental behavior of microgravity sensor.

The device shown in Figure 3.6 was introduced and characterized by transforming two plasmid L_EGFP_AC and T_antiEGFP_synsrRNA_UA in *E.coli* cell strain DH5αZ1. The EGFP repression capability of the antiEGFP synsrRNA was compared between microgravity and earth gravity. Here, each bar represents the mean EGFP expression folds change between IPTG only (absence of anti-EGFP SynsrRNA) and IPTG + ATC (presence of anti-EGFP SynsrRNA) conditions. At least three biological replicates were used. Error bar represent standard deviation from the mean. Unpaired t-test was used to

Chapter three

measure the extent of statistical significance where (** means P value lies between 0.001 to 0.01)

Next, to test, if the difference was stemmed from the microgravity induced changes in inherent expression of EGFP, instead of information processing through our device, we ran similar experiments with cells having EGFP under $P_{L_{lacO-1}}$ (Figure 3.19 a) but without the plasmid carrying anti-EGFP SynsrRNA (Figure 3.19 b), in two conditions (IPTG only and IPTG + ATC). We observed no significant differences (0.94 times, p value = 0.4) between microgravity and normal gravity (Figure 3.9).

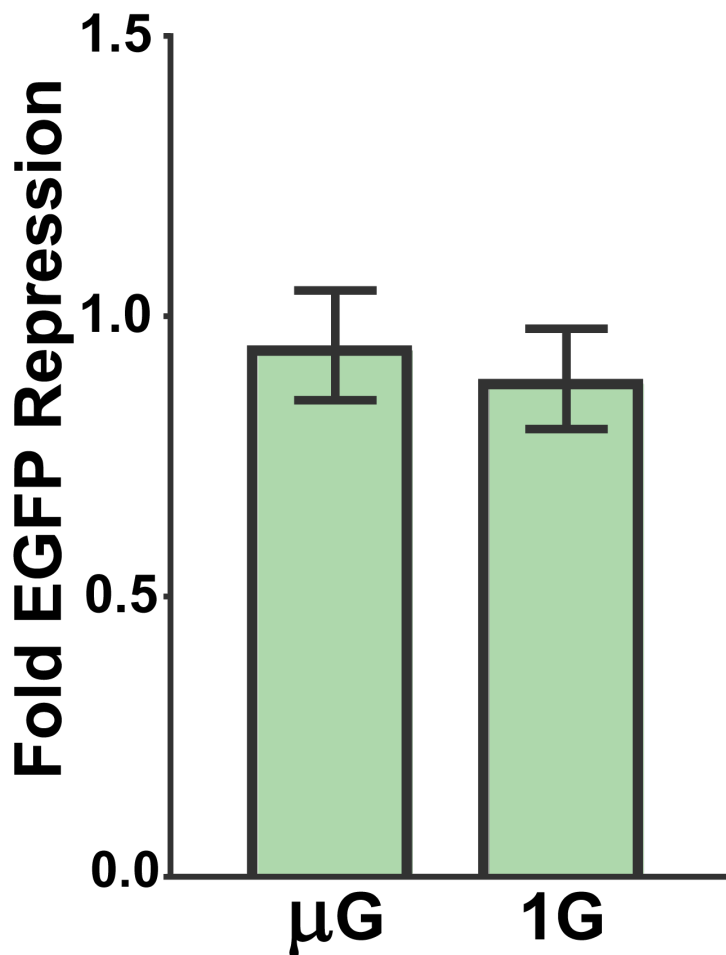


Figure 3.9 Control experiment.

To test if the difference in EGFP repression was the property of the device and not difference in EGFP expression itself we ran similar experiment as shown in Figure 3.8 without using plasmid containing anti-EGFP synsrRNA i.e $T_{antiEGFP_synsrRNA_UA}$. Difference in EGFP expression was compared (if any)

Chapter three

between microgravity and earth gravity Each bar represents the EGFP expression folds change between IPTG only and IPTG + ATC conditions. Error bar represent standard deviation from the mean. At least three biological replicates were used. No asterisk (*) means no significant difference in fold EGFP repression between microgravity and Earth gravity.

This suggested that our genetic device integrated the microgravity signal and responded accordingly. However, we observed a higher level of standard deviation in EGFP repression fold change in normal gravity than the microgravity (Figure 3.8). During IPTG + ATC induced state, the anti-EGFP SynsrRNAs were produced from the high copy number plasmid carrying a pUC origin of replication (pUC Ori). As this pUC Ori had a disrupted plasmid copy number regulation mechanism, it showed a higher degree of variation in plasmid borne gene expression (S *et al.*, 2008). Therefore, a moderate level of variation in the expression of anti-EGFP SynsrRNAs is highly probable, which was reflected in the higher standard deviation in normal gravity. However, in microgravity, Hfq was deregulated. This resulted a little or no repression of the EGFP expression by anti-EGFP SynsrRNAs. Therefore, the variation in anti-EGFP SynsrRNAs due to plasmid copy number variation had little or no effect. Hence, lower standard deviation was observed. It was interesting to note that a 50% change in Hfq in microgravity (Figure 3.5) was good enough to create substantial changes in the synthetic genetic device, using Hfq mediated silencing of target genes with SynsrRNAs. The Hfq was found downregulated in microgravity in a few other bacteria (Wilson *et al.*, 2007; Roy, Shilpa and Bagh, 2016). This inspires the possibilities of creating microgravity responsive genetic devices in other bacteria too. However, its functionality beyond *E.coli* would depend on two facts. First the nature and degree of Hfq mediated regulation of gene expression by small regulatory RNAs and the degree of the Hfq repression in microgravity.

3.2.3 The basic device design is general in nature for *E.coli*

As the SynsrRNA can be designed potentially against any gene (mRNA), our designed device should serve as a general design for creating biochemical microgravity responsive device. To test, we created a new genetic device (Figure 3.11) by replacing the target gene EGFP to TdTomato, an orange fluorescence protein, derived from DsRed, whose DNA sequence was completely different from EGFP and GFP derived fluorescent proteins (Shaner *et al.*, 2004). Further, we replaced the anti-EGFP SynsrRNA to anti-TdTomato SynsrRNA (Figure 3.10, Figure 3.19 d) and the IPTG inducible PLacO-1 promoter to a constitutive lambda promoter P_R (Figure 3.19 c). The device design was fabricated according to the design (Figure 3.11) and incorporated in two plasmids P_tdTomato_AC (Figure 3.19) and T_antitdTomato synsrRNA_UA (Figure 3.19), which carries td-tomato under constitutive promoter P_R and antitdTomato synsrRNA under P_LTetO-1, respectively.

5'**gcatgaactctttgatgacctctcgcccttgctcaccatTTTCTGTTGGGCCATTGCATT-**
GCCACTGATTTTCCAACATATAAAAAGACAAGCCCGAACAGTCGTCCG
GGCTTTTTTTCTCGAGCT3'

Figure 3.10 Anti-tdTomato synthetic small regulatory RNA encoding synthetic DNA sequence.

The sequence highlighted in green is the target (tdTomato mRNA) binding region and the MicC scaffold is highlighted in yellow.

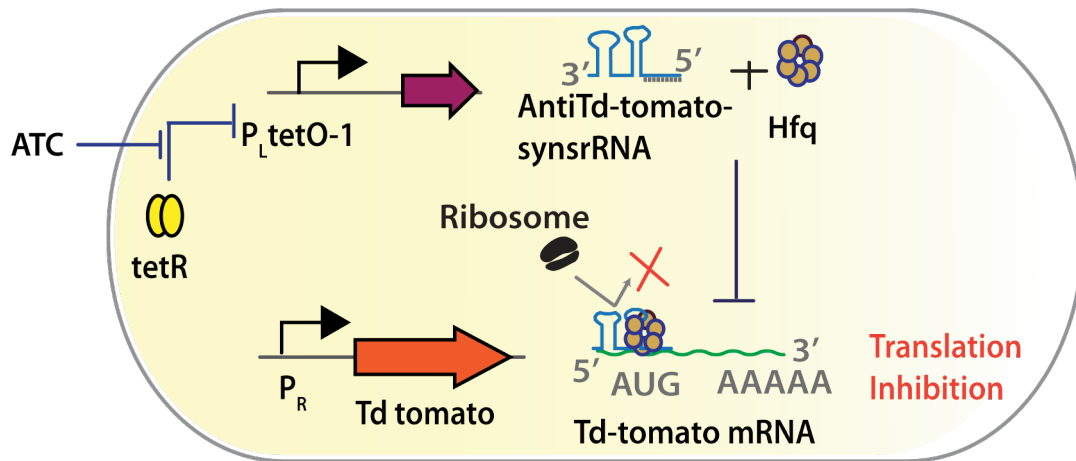


Figure 3.11 The design of the Td-tomato based microgravity sensor circuit.

The antiTd-tomato synsrRNA is expressed under ATC inducible promoter $P_{LtetO-1}$, the target protein is expressed under constitutive promoter P_R , translation inhibition of target mRNA by synsrRNA is mediated by Hfq protein. The device design was incorporated in two plasmids $P_tdTomato_AC$ and $T_antitdTomato_synsrRNA_UA$.

We performed the same microgravity experiments with this new system and found that it worked as a microgravity responsive device too with a 3.77 times differences between 1G and microgravity (Figure 3.12). Next, to test, whether the differences was due to the microgravity induced change in the inherent expression of TdTomato under P_R promoter, but not through microgravity integration in our device, we ran control experiments with cells having TdTomato under P_R promoter (Figure 3.19 c) but without the plasmid carrying anti-TdTomato SynsrRNA (Figure 3.19 d), in two conditions (without ATC and with ATC). We observed no significant differences (0.93 times, p value = 0.67) between microgravity and normal gravity (Figure 3.13).

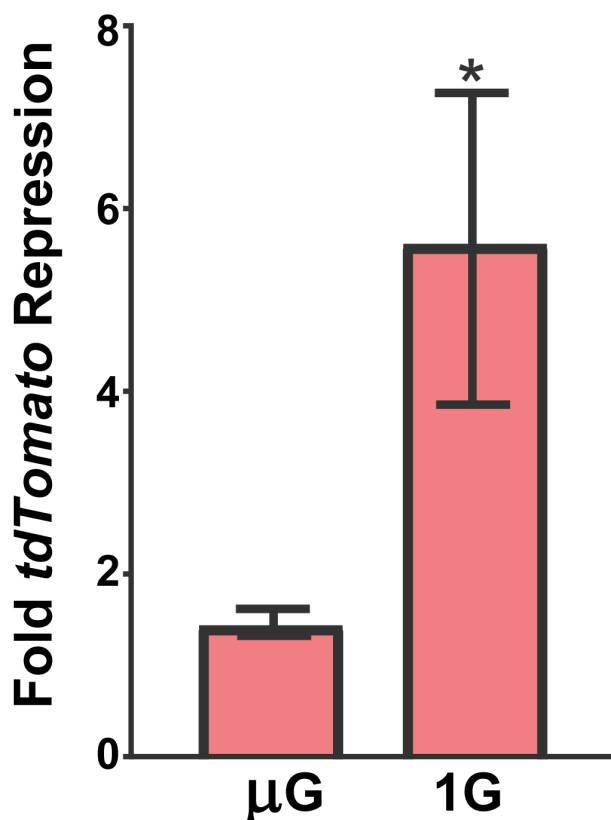


Figure 3.12 Experimental behavior of the TdTomato based microgravity sensor.

The device shown in Figure 3.11 was introduced and characterized by transforming two P_tdTomato_AC and T_antitdTomato synsrRNA_UA in *E.coli* cell strain DH5αZ1. The TdTomato repression capability of the antitdTomatosynsrRNA was compared between microgravity and earth gravity. Here, each bar represents the TdTomato expression folds change between IPTG only (absence of antitdTomatosynsrRNA) and IPTG + ATC (presence of antitdTomatosynsrRNA) conditions. At least three biological replicates were used. Unpaired t-test was used to measure the extent of statistical significance where (*) means P value lies between 0.01 to 0.05)

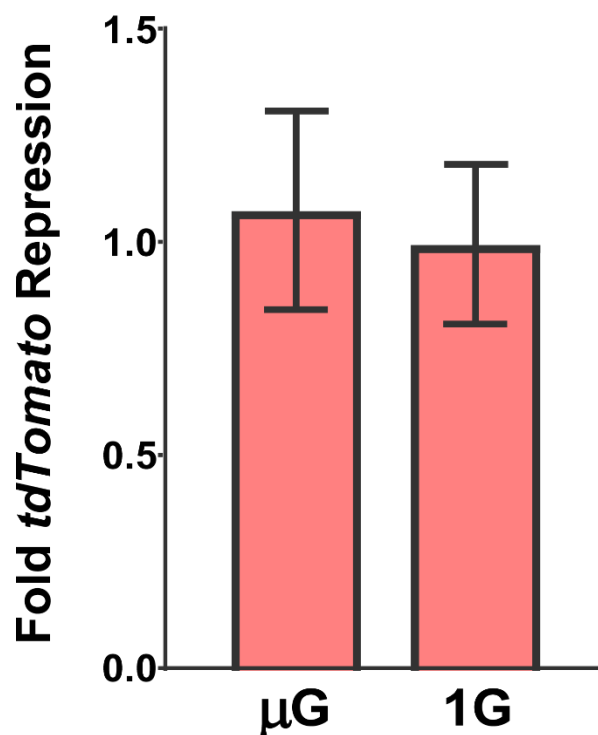


Figure 3.13 Control experiment.

To test if the difference in *tdTomato* repression was the property of the device and not difference in *TdTomato* expression itself we ran similar experiment as shown in Figure 3.12 without using plasmid containing anti-*tdTomato* *synsrRNA* i.e *T_antitdTomato synsrRNA_UA*. Difference in *tdTomato* expression was compared (if any) between microgravity and earth gravity Each bar represents the *tdTomato* expression folds change between IPTG only and IPTG + ATC conditions. Error bar represent standard deviation from the mean. At least three biological replicates were used. No asterisk (*) means no significant difference in fold *tdTomato* repression between microgravity and Earth gravity.

This suggested that the new genetic device responded against microgravity through integrating the microgravity as a signal within the genetic circuits. However, the difference (3.77 times) is several times less than the EGFP based device (where the signal difference between 1G and microgravity was 28.2 times). One of the possible explanations for this difference is differences in the efficacy of the corresponding synthetic small regulatory RNA. Since in case of EGFP there was a 30 times repression of EGFP expression, when *SynsrRNA* expression was fully induced (Figure 3.8, the column in 1G) compare to 5 times repression in *TdTomato* (Figure 3.12, the column in

1G). The inhibition efficiency of designed antisense domain of the SynsrRNA is a strong function of length, the position of target binding, and the target mRNA itself. In our design, we applied the rational strategy proposed by (Na *et al.*, 2013), which suggested to use at least the first 24 bases of the target mRNA. Though, this strategy was a good starting strategy, different efficiency was observed based on target mRNA and the scaffold region (H *et al.*, 2013; Sakai *et al.*, 2013; Zhao *et al.*, 2016). An efficient screening strategy may overcome this bias. Another possibility which could lead to the difference between repression of two system, is promoter strength differences, it may be possible that P_R -tdTomato expression is very higher than P_{LacO-1} -EGFP expression, as a result the sRNA is unable to neutralize the tdTomato mRNA.

3.2.4 Application of the device to control native cellular properties with microgravity

As this biochemical device could integrate microgravity with it and potentially target any gene as a microgravity responsive gene, it may be applied to target inherent cellular biochemical process, which can be controlled in a microgravity responsive way. Therefore, we applied this device to control the cell division process with microgravity. The effect of microgravity on the *E. coli*. cell division and growth were studied in space flight and simulated microgravity condition. An increase in cell growth was observed in microgravity (G *et al.*, 1994; Nickerson *et al.*, 2004; Arunasri *et al.*, 2013; Kim, Matin and Rhee, 2014). However, no mechanistic details are known. Here, using our device, we targeted a protein FtsZ, which directly involves in the cell division process. FtsZ is a tubulin family protein and its deregulation hinder the cell division process and shows elongated and filament-like shape (Sánchez-Gorostiaga *et al.*, 2016). Here we targeted native FtsZ by creating a microgravity responsive biochemical device (Figure 3.15).

5'**taatcaccgcgtcattggttaagttccattggtcaaaca****TTTCTGTTGGGCCATTGCATT-**
GCCACTGATTTTCCAACATATAAAAAGACAAGCCCGAACAGTCGTC
GGCTTTTTTTCTCGAGCT3'

Figure 3.14 AntiFtsz synthetic small regulatory RNA encoding synthetic DNA sequence.

The sequence highlighted in green is the target binding region and the MicC scaffold is highlighted in yellow.

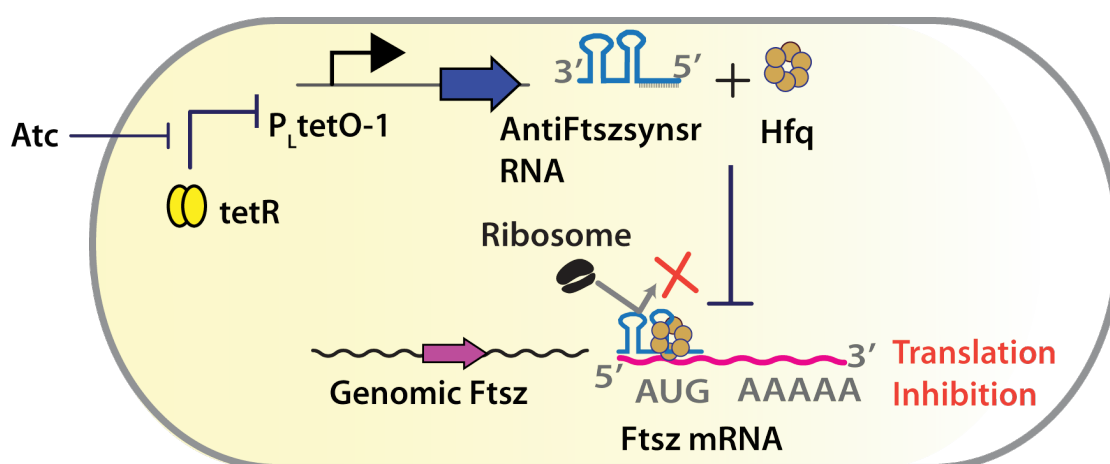


Figure 3.15 SynsrRNA based microgravity sensor for integrating microgravity with native FtsZ protein in Dh5αZ1 *E.coli*.

The synsrRNA is expressed under ATC inducible promoter P_LtetO-1, the target protein is a native cellular protein essential for cell division, translation inhibition of target mRNA by synsrRNA is mediated by Hfq protein.

Here, we expressed anti-FtsZ SynsrRNAs (Figure 3.14) under P_L-TetO-1 promoter from plasmid T_antiFtszsynsrRNA_UA (Figure 3.19 e) in DH5αZ1. We also co-transformed a constitutively expressing EGFP plasmid (Figure 3.19 f) for easy visualization of the cells under fluorescence microscope. In normal gravity, with the help of normal level of Hfq, in ATC induced condition, anti-FtsZ SynsrRNAs repressed the expression of the native FtsZ in the cell. The reduced repression of FtsZ resulted a deregulation in cell division process, which resulted deformed and elongated bacterial cell

Chapter three

shapes (Figure 3.16 a). The Hfq mediated repression of FtsZ by anti-FtsZ SynsrRNAs (in ATC induced condition) was disrupted due to decrease in the Hfq, when the cells were under microgravity. This allowed the native FtsZ to function normally and brings back the elongated cells to normal size (Figure 3.16 b). Thus, the *E.coli* with this device expressing SynsrRNA against FtsZ showed that the deformed and elongated cell shapes in earth gravity got rescued in microgravity as shown in Figure 3.16 and Figure 3.18. However, the cell sizes in microgravity were still bigger than the control (Figure 3.17 and Figure 3.18 a), where the engineered cells were grown in the absence of inducers in both gravity and microgravity conditions. We reasoned that in microgravity, the amount of hfQ decreased but not became zero. Thus, a basal level Hfq mediated repression was always present. This made a slight deregulation in FtsZ, which resulted a little increase in the cell shape in microgravity compare to the control (cells without FtsZ carrying plasmid, Figure 3.16 and Figure 3.18 b). All the individual fluorescence and DIC images are shown in Figure 3.18.

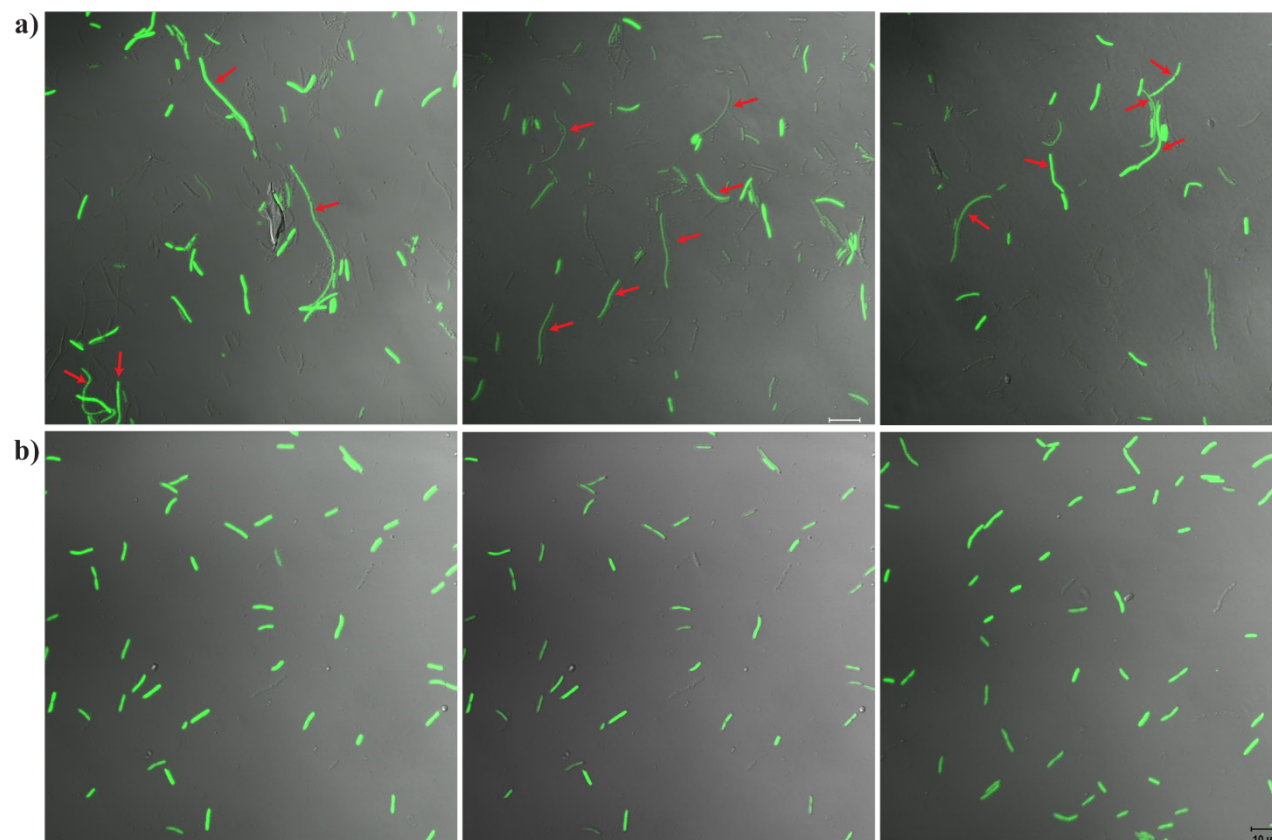


Figure 3.16 Merged fluorescence and DIC image of *E. coli* DH5αZ1 cells expressing SynsrRNA against FtsZ from microgravity responsive synthetic genetic device.

The cells were cultured in microgravity simulating reactor for 24 h in presence of inducer at saturated concentration in A) Earth's gravity condition (horizontal position), B) microgravity condition (vertical position). The elongated cells in the earth's gravity were shown with red arrows

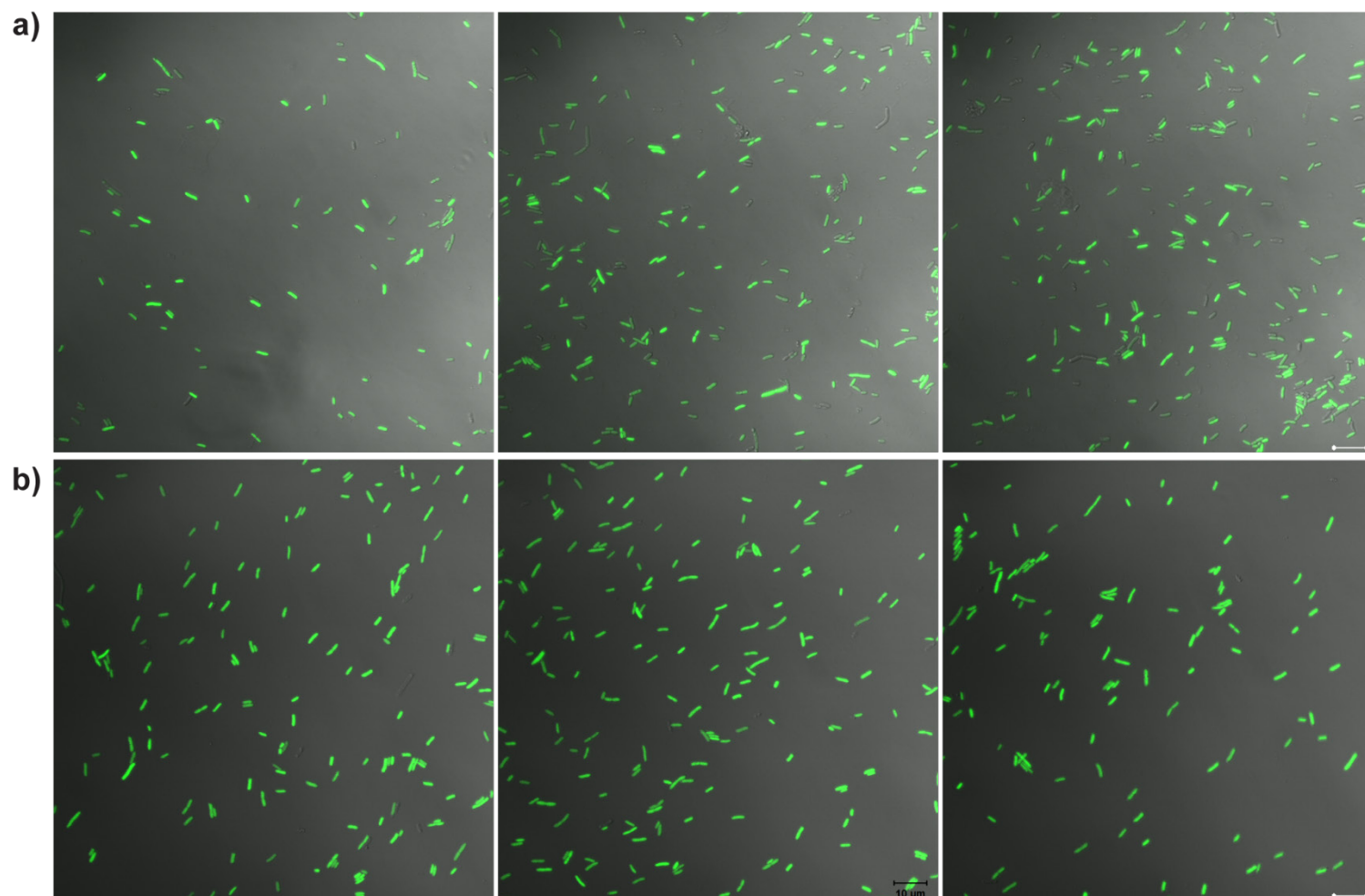


Figure 3.17 Control experiments for SynsrRNA against FtsZ from microgravity responsive synthetic genetic device.
Merged images of Dh5αZ1 cell co-transformed with anti-Ftsz synsr, which have SynsrRNA against genomic FtsZ under Atc inducible PltetO-1 promoter and P_EGFP_AC, expressing EGFP constitutively from PR promoter

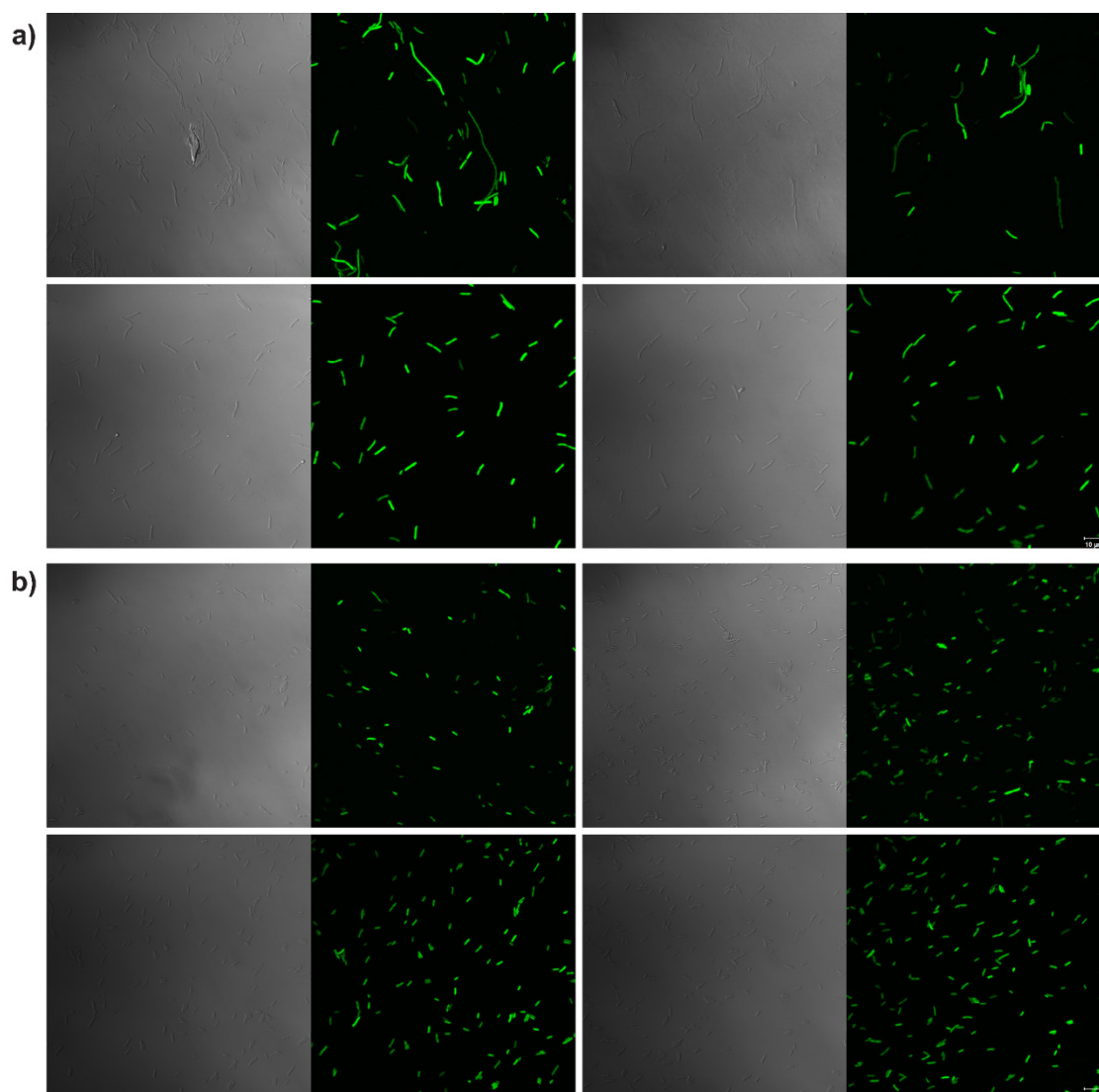


Figure 3.18 DIC and Fluorescence image of the merged images a) in Figure 3.16 and b) in Figure 3.17 (control experiments).

3.3 Discussion & Conclusion

Microbiological organisms are an important candidate for future space synthetic biology. The risk of radiation and microgravity not only affects astronauts but microorganisms also (Horneck, Klaus and Mancinelli, 2010). One of the concerning facts for NASA is increased virulence and antibiotic resistance in pathogens in response to microgravity (DM and HN, 2006; Wilson *et al.*, 2007). Microgravity is also responsible

Chapter three

for imbalance in gut microbiome (JS *et al.*, 2013; P *et al.*, 2015) and biofilm formation as a stress response of microbes (Mauclaire and Egli, 2010). Therefore, to engineer microorganisms in space, it is essential to understand their behavior in space, as these engineered microbes may not function as they are expected to do.

In this work, we created the first biochemical and cellular microgravity responsive device using an engineered genetic circuit in *E.coli*, which responded to microgravity by changing the expression of a target gene. The underlying principle of the device was based on the downregulation of Hfq protein in *E.coli* in microgravity, which was translated by coupling Hfq mediated silencing of target gene by synthetic small regulatory RNAs. This resulted a reduced silencing of the target gene in microgravity. Thus, the basic design of the device is general in nature for *E.coli*, which we demonstrated by altering its target genes (EGFP, TdTomato, and FtsZ) and the promoters (inducible and constitutive). We applied this device to control the cell division process, where the disrupted cell shape due to FtsZ deregulation in normal gravity was rescued in microgravity. However, for better results, we believe an extensive screening of SynsrRNA against target genes is required. As the device can integrate microgravity as a physical signal potentially with any downstream synthetic and native cellular processes in *E.coli*, it may have future space applications in microgravity responsive metabolic engineering and in developing cellular microgravity-health hazards monitoring systems. The human designed control of cell division processes and phenotype with microgravity suggested successful integration of microgravity as a physical signal with cellular biochemical processes and may have some applications in space bioengineering. However, in order to perform such experiments in space, like in international space station, specialized hardware is required. For example, the confocal microscopy

Chapter three

experiment of this work may be done in space by using FLUMIAS-DEA, a miniaturized fluorescence microscopy platform (Thiel *et al.*, 2019) and other cellular fluorescence measurements could be done with ‘Microflow1’, a sheath less fiber optic flow cytometer (G *et al.*, 2014) and Nanoracks fluorescence plate reader (*NanoRacks Honored with Space Station Innovation Award | NASA*, no date). Further, innovative superwetttable microchip platform was developed for biosensing applications in microgravity (Xu *et al.*, 2016). Such platform would be useful to conduct the experiments on cellular microgravity responsive device at the single cell level in space and to assess its real usefulness. Taken together, our work may have significance in space bioengineering and synthetic biology.

3.4 Materials and Method

3.4.1 Media, strains and growth conditions

We used DH5 α Z1 (laciq, SpR, deoR, supE44, Delta(lacZYA-argFV169), Phi80 lacZDeltaM15, hsdR17(rK- mK+), recA1, endA1, gyrA96, thi-1, relA1) strain of *E.coli* for our experiment as well as for cloning. The "Z1" strains carry two copies of the Lac Repressor and Tet Repressor encoding genes (lacI and tetR, respectively) at the attB locus, driven by the constitutive promoters Placiq and PN25, respectively. The plasmid(s) were transformed in chemically competent *E. coli* DH5 α Z1 cells and transformed cells were grown in LB-Agar, Miller (Difco, Beckton Dickinson) plates with appropriate antibiotics, followed by overnight liquid culture in LB broth from single colonies at 37 °C with appropriate antibiotics. As the microgravity simulator has 4 HARVs, during each run, we can run two biological replicates with and without inducers. Two colonies were picked and grown overnight in Luria-Bertani, Miller broth

Chapter three

(Difco, Beckton Dickinson) containing required antibiotic(s) (100 µg/mL Ampicillin sodium salt, Himedia; 34 µg/mL Chloramphenicol, Himedia) at 37°C & 180 rpm. Next day each overnight culture was diluted 1/100 times in two set of fresh LB; one containing appropriate antibiotic(s) and 1mM IPTG, (Sigma Aldrich); and another set containing required antibiotic(s), 1mM IPTG and 100 ng/ml ATC, (Sigma Aldrich). The cell suspensions were carefully transferred to individual 50 ml high aspect ratio rotating vessels (HARVs; model number: RCCS 4H Synthecon, Houston, TX, USA) without introducing any air bubble, which may disrupt the low shear stress condition. The HARVs were rotated in a vertical base platform to mimic the microgravity environment (Figure 3.4 a) at 37°C, 25 RPM for 24 hrs. (Castro *et al.*, 2011). We reoriented the previous setup in a horizontal manner using a custom platform as an earth gravity control (Figure 3.4 b). The vessels were cleaned and autoclaved after every run according to synthecon user manual. DH5αZ1 strain without any plasmid was grown in 10 ml LB broth without any antibiotic in 100 ml conical flasks for measurement of auto-fluorescence.

3.4.2 synsrRNA and Plasmid construction

All the oligonucleotides used for construction are listed in Table 3.1 and were synthesized from Integrated DNA Technologies, Singapore. The relevant plasmids maps are shown in Figure 3.19 as well as shown in a tabulated manner in Table 3.2. All the plasmids were constructed using standard cloning techniques. The *EGFP* gene, P_{LtetO-1} promoter, T1 terminator, p15A origin of replication (ori), ampicillin and chloramphenicol resistance genes were procured from plasmids: pOR-*EGFP*-12 and pOR-Luc-31 (a gift from Prof. David McMillen, University of Toronto, Toronto, Canada). The pUC origin of replication was amplified from pmCherry-N1 (Clontech) using

Chapter three

pUC_Lp and pUC_Rp (Table 3.1) flanked by *SpeI* and *AvrII*. The *tdTomato* gene was obtained from *ptdTomato* vector, clontech. P_R promoter along with RBS was constructed by overlap extension PCR using *Pr_LpI* and *Pr_RpI*. *EGFP* gene was fused to this promoter-RBS fragment by an overlap PCR with *Pr_LpII* and *EGFP_RpII* to construct full-length P_R -*EGFP*. This P_R -*EGFP* construct was flanked by *XhoI* and *XbaI*. The DNA encoding for small synthetic RNA against *EGFP* was chemically synthesized (Invitrogen GeneArt Gene Synthesis service, Thermo Fischer), the same was used as a template to construct *synsrRNA* against *tdTomato* using primer *antitom_Lp* and *antitom_Rp*. The *synsrRNA* against *Ftsz* was constructed by overlap pcr of primer *antiftsz_Lp* and *antiftsz_Rp*. All *synsrRNA* were cloned downstream of an inducible $P_{LtetO-I}$ promoter using *EcoRI* and *Xba-I* into plasmid harbouring pUC origin and ampicillin resistance. The target genes *EGFP* and *tdTomato* gene were introduced in a plasmid with p15a origin and chloramphenicol resistance gene, under inducible $P_{LlacO-I}$ promoter and a constitutive P_R promoter, respectively. The plasmid containing the *synsrRNA* e.g *antiEGFPsynsrRNA* and the corresponding target gene e.g. *EGFP* were co-transformed in DH5 α Z1. To aid visualization a plasmid P_{EGFP_AC} , where *EGFP* is constitutively expressed under P_R promoter, was used in case of *ftsZ* silencing experiment. All restriction enzymes, T4 DNA ligase and buffers were purchased from New England Biolabs. All PCR reactions were carried out by KOD Hot Start DNA polymerase (Merck Millipore), Pfu Turbo Hotstart PCR Master Mix (Agilent Technologies), or Phusion High-Fidelity PCR Master Mix (New England BioLabs). PCR and Gel purification were done using Qiaquick nucleotide removal kit (Qiagen) and Qiaquick Gel purification kit (Qiagen), respectively, all according to the manufacturer's instruction. Invitrogen SYBR Safe DNA gel stain was used for visualization of DNA in blue light

Chapter three

illuminator. All the plasmids were extracted using Qiaquick miniprep kit (Qiagen) and sent for sequencing to Eurofins Genomics India Pvt. Ltd., Bangalore, India.

Table 3.1 List of all the synthetic oligos used in this study

Primer name	Primer Sequence (5'-->3')	Description	Source
EGFP_LpI	caagggcgaggagctgtt	1st round amplification of <i>EGFP</i> (forward)	This study
EGFP_RpI	ccatgccgagagtgatcc	1st round amplification of <i>EGFP</i> (reverse)	This study
EGFP_LpII	Cttcagtcgaggtaccatggtgagcaagggcgag-gagctgtt	2nd round amplification of <i>EGFP</i> flanked by KpnI (forward)	This study
EGFP_RpII	Ctgattatgatctagattactgttacagctcgccatgccgagagtgatcc	2nd round amplification of <i>EGFP</i> flanked by XbaI (Reverse)	This study
pUC_Lp	gcaggaaagacctaggggctggcactctgtcgatac	Amplification of pUCori(Forward primer)	This study
pUC_Rp	gactccaagcactagtaggggataacgcaggaaaga	Amplification of pUCori (Reverse primer)	This study
Pr_LpI	tcgtcttcggctcgagtaacaccgtgcgtgttgactattttac-ctctggcggtgataatggtt	1 st round amplification of P _R promoter(Forward primer)	This study
Pr_RpI	tgaacagctcctcgcccttgctcaccatggtac-ctttctcctctttaatgaattcg-tacatgcaaccattatcaccgccagag	1 st round amplification of P _R promoter (Reverse primer)	This study
Pr_LpII	tcgtcttcggctcgagtaac	2 nd round amplification of P _R promoter (Forward primer)	This study
antifsz_Lp	cgtcttcacgaattctaataccgcgctcattggttaagttccatt-ggtcaaacattttctgttgggccattgcattgccactgattttcc	Overlap extension of antiFtszsynsrRNA flanked by EcoRI(Forward primer)	This study
antifsz_Rp	ttcttagatgtctagaagctcgagaaaaaagcccggac-gactgttcgggcttgctctttta-tatgttggaataatcagtggaatgc	Overlap extension of antiFtszsynsrRNAflanked by XbaI(Reverse primer)	This study

Chapter three

antitom_Lp	ccacaagttcgaattcgcacatgaactctttgatgac- ctcctcgcccttgctcaccattttctgttg	Generating synsrRNA against tdTomato, using T_antiEG-FPsynsrRNA_UA as a template (Forward primer)	This study
antitom_Rp	ctccagtcgtggatcccgcgaggattgtcctactc	Generating synsrRNA against tdTomato, using T_antiEG-FPsynsrRNA_UA as a template (Reverse primer)	This study
Ecoli_16s_LP	ccacggaagtttcagagatg	QPCR primer against 16s rRNA (Forward primer)	This study
Ecoli_16s_RP	gctggcaacaaaggataagg	QPCR primer against 16s rRNA (Reverse primer)	This study
HFQ_QPCR_LP	cgcgatttctactgttgcc	QPCR primer against Hfq (Forward primer)	This study
HFQ_QPCR_RP	attcggtttcttcgctgtcc	QPCR primer against Hfq (Reverse primer)	This study
synsrRNA_seqLp	cgaaaagtgccacgtgac	Sequencing Primer for plasmid containing synsrRNA	This study
target_seqLP	gcttttggcgaagaatgaaa	Sequencing Primer for plasmid containing fluorescent protein targets	This study
RBS 1_left	aattcacagcagacgagaaaggtac	RBS for reducing translation initiation rate of tdTomato (left hybridising oligo)	
RBS 1_right	ctttctcgtctgctgtg	RBS for reducing translation initiation rate of tdTomato (right hybridising oligo)	

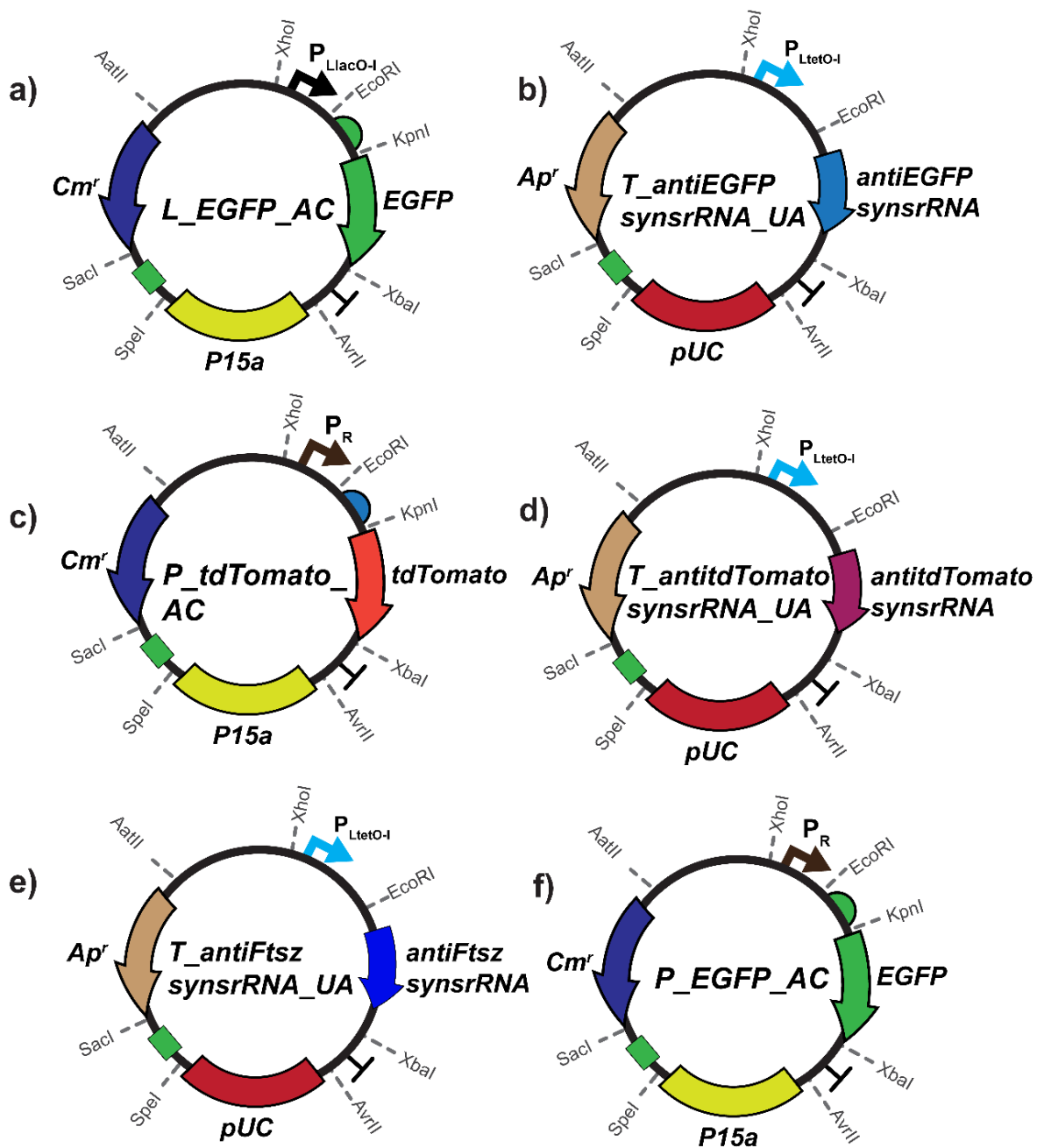


Figure 3.19 Plasmid maps for the plasmids developed for this study.

Table 3.2 List of all the plasmids used in this study.

Sr no	Plasmid Name	Origin of replication	Antibiotic selection	Description	Source
1	pOR-EGFP-12	ColE1	Amp	Source of <i>EGFP</i> , Ampicillin resistance, ColE1 origin	Kind gift from Prof. David McMillen
2	pOR-Luc-31	P15A	Cam	Source of P _L tetO-1 and Chloramphenicol resistance, P15A origin	Kind gift from Prof. David McMillen
3	pmCherryN-1	pUC	Amp	Source of pUC origin of replication	Clontech
4	ptdTomato	pUC	Amp	Source of <i>td-Tomato</i>	Clontech
5	L_EGFP_AC	p15A	Cam	<i>EGFP</i> under P _L lacO-1	This study
6	P_ <i>tdTomato</i> _AC	p15A	Cam	<i>td-Tomato</i> under P _R	This study
7	P_EGFP_AC	P15A	Cam	<i>EGFP</i> under P _R	This study
8	T_antiEGFPsynsrRNA_UA	pUC	Amp	antiEGFPsynsrRNA under P _L tetO-1 promoter	This study
9	T_antitdTomatosynsrRNA_UA	pUC	Amp	antitdTomatosynsrRNA under P _L tetO-1 promoter	This study
10	T_antiFtszsynsrRNA_UA	pUC	Amp	antiFtszsynsrRNA under P _L tetO-1 promoter	This study

3.4.3 Working principle of Synergy HTX Multi-Mode plate reader (Biotek Instruments, USA)

A multi-mode plate reader also known as microplate reader device is capable of detecting and quantifying two or more signals. The device is usually capable of detecting absorbance, luminescence, fluorescence, time-resolved fluorescence, and alpha-screen assays. They are routinely used in pharmaceutical and biotechnological industry and academic organisations for testing, drug development, bioassay confirmation, quality management, and manufacturing processes. It is capable of quantifying several sample assays in a relatively short period of time. It supports a variety of microtiter plate formats varying from six to three hundred and eighty-four wells; nevertheless, the most often used microtiter plate type in academic laboratories is the 96 well format. It employs a xenon lamp and a monochromator to pick a wavelength (between 200 and 999 nm) for absorbance quantification. It features a tungsten quartz halogen lamp with user-selectable system-compatible filters and a photo multiplier tube (PMT) for fluorescence quantification (Figure 3.20). It comes with Gen5 software platform manages data processing, review, and export. It has a shaking function that is used to prevent cell sedimentation. Additionally, it includes a specialized plate, take3, that is used to quantify nucleic acids in microliter volumes.

Multimode microplate readers based on optical filters (Figure 3.20) usually integrate several sets of optical filters. In the excitation line, white light is passed through a filter that usually transmits at least 60% of the target wavelength in the visible spectrum to the sample well. The light excites the sample, which responds by emitting a fluorescent signal unique to its properties. An emission filter is used in conjunction with a dichroic mirror to clean the sample's signal and usually transmits at least 60% of the

Chapter three

target wavelength to the detector. Sensitivity is increased due to the high transmission efficiency on both the excitation and emission channels (*Assay: Hybrid Bridges the Microplate Reader Gap*, no date).

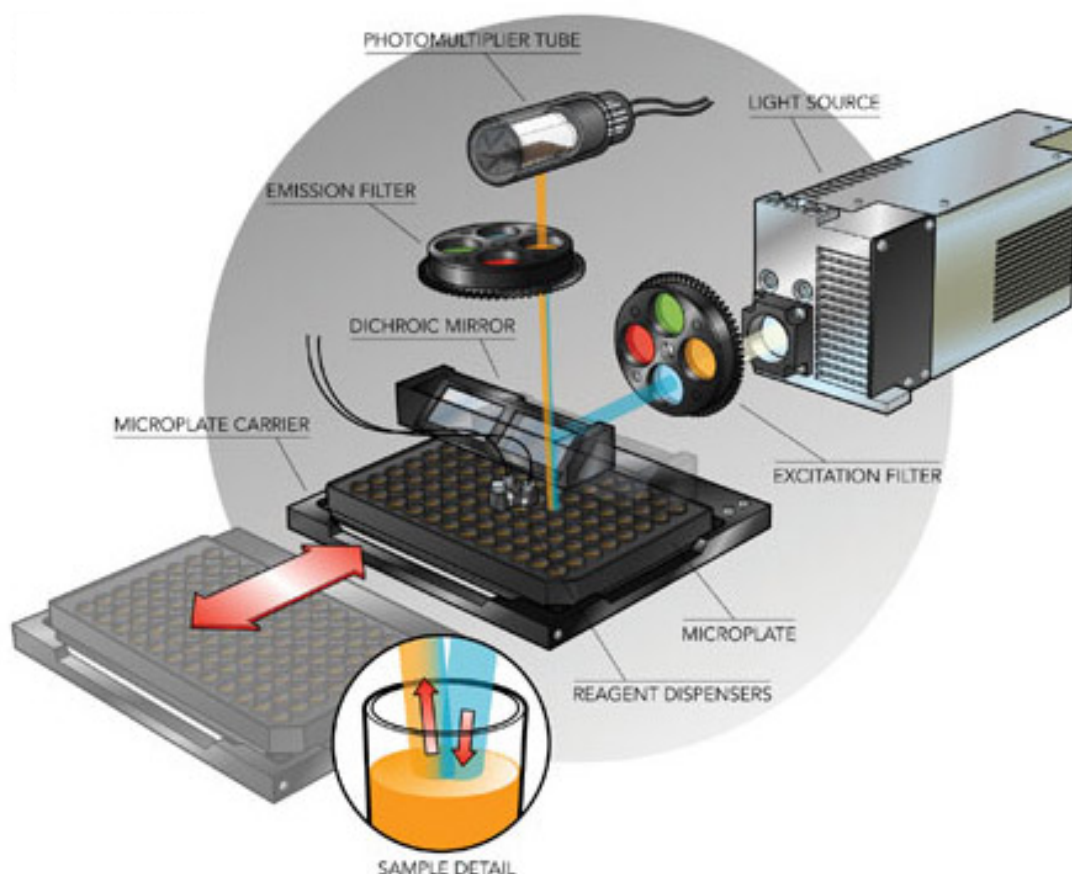


Figure 3.20 Schematic diagram of filter-based fluorescence detection in multimode reader.

(Reproduced with permission from (*Assay: Hybrid Bridges the Microplate Reader Gap*, no date))

3.4.4 Working principle of fluorescence microscope (confocal LSM 710, Ziess)

Fluorescence microscopy is a particularly effective instrument for biological experiments, since the procedure is nondestructive, and can have both high sensitivity and excellent spatial resolution. Fluorescence microscopy is dependent on the fluorescence

Chapter three

of the specimens of interest. Fluorescence is the emission of light from the specimen within nanoseconds after the absorption of shorter wavelength light. The principal behind fluorescence detection is based on Stoke's shift, or the difference between the excited and emitted wavelengths. Thus, by completely filtering out the exciting light and allowing the emitted fluorescence to pass through, it is possible to see only the objects that are fluorescent. Numerous fluorescence microscopy techniques are used, including epifluorescence, confocal, small field complete internal reflection, and near-field scanning.

Confocal microscopy Figure 3.21 illuminates the material with a guided laser beam directed at a specific point in the sample focal plane. The light from this point is observed after passing through a pinhole, ensuring that only light produced from the focal plane passes through and is registered on the detector. Since the pinhole blocks light from out-of-focus planes, the confocal records only light from the sample's focal plane. Scanning mirrors are used to raster the laser beam over the sample, gradually forming an image. Confocal microscopy can be considered for samples thicker than $\sim 20\text{ }\mu\text{m}$ (Thorn, 2016). Since we are interested in *E.coli* cells, we partially opened the pinhole and used it as an epifluorescence microscope.

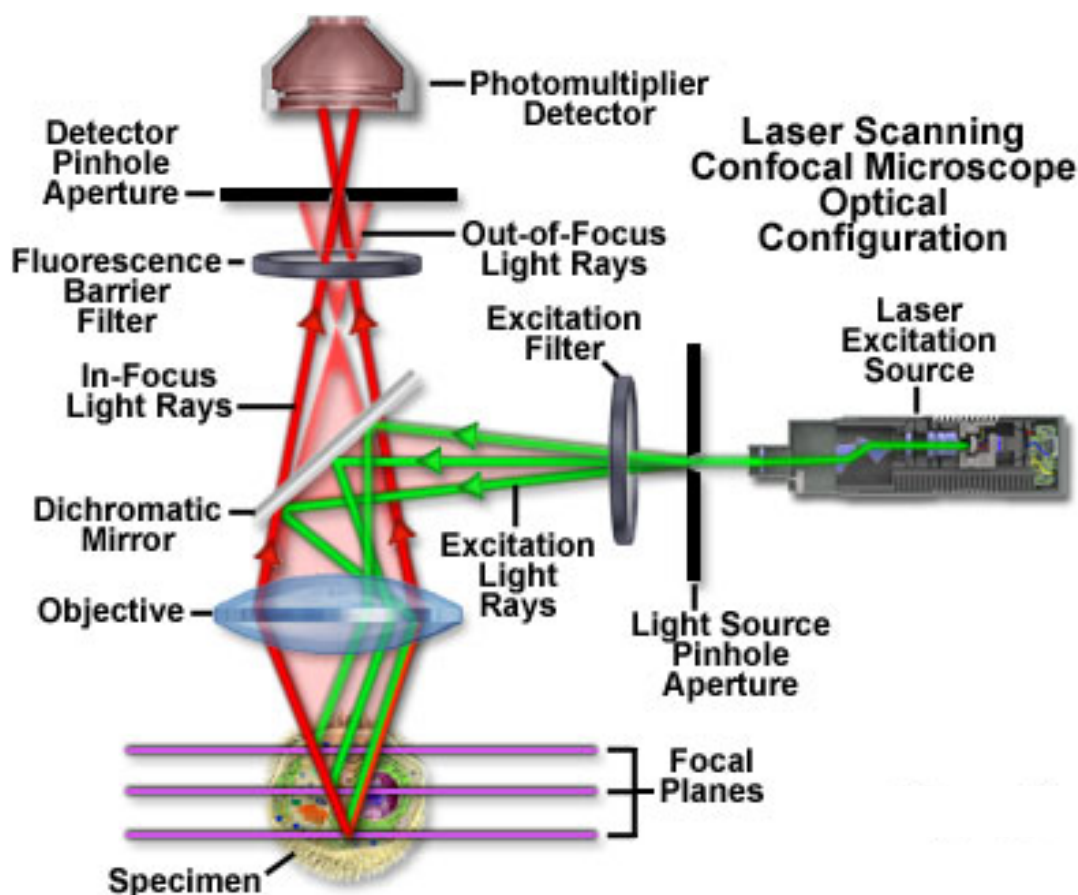


Figure 3.21 Schematic diagram of confocal microscope used in these experiments.

(Reproduced from (ZEISS Microscopy Online Campus | Live-Cell Imaging | Microscopy Techniques, no date) with permission from Zeiss.)

3.4.5 Fluorescence Measurements and data analysis

At least 3 biological replicates were used for comparing fold repression between microgravity and earth's gravity. Cell fluorescence and absorbance at 600 nm were measured after 24 hrs. We took out 2ml of cell suspension from each HARV and washed twice in 1ml cold PBS (pH 7.4), and resuspended in 500 μ l of PBS. Then the cells were diluted in PBS (pH 7.4) to reach around OD₆₀₀ 0.8, 200 μ l cell suspension was loaded onto 96-well multiwell plate (black, Greiner Bio-One) for measurement using Synergy HTX Multi-Mode reader (Biotek Instruments, USA). The measurement steps were

Chapter three

performed in following order: Absorbance at 600nm, followed by 15s linear shaking, final step was fluorescence measurement. For measuring fluorescence of *EGFP*, the cells were excited by a white light source that had been passed through an excitation filter 485/20 nm and emission was collected by 516/20 nm bandpass filter for with appropriate gain. Similarly, for fluorescence measurements of tdTomato an excitation filter of 540/35 nm and an emission filter of 590/20 nm was used. The raw fluorescence measurements were normalized using eqn1, where $Fluor_{norm}$ is the normalized fluorescence, $Fluor_{sample}$ is the raw fluorescence of the sample, $OD600_{sample}$ is the absorbance of the sample at 600nm, $Fluor_{control}$ is the auto-fluorescence of DH5 α Z1 strain, and $OD600_{control}$ is the absorbance of DH5 α Z1 strain at 600nm.

$$Fluor_{norm} = \frac{Fluor_{sample}}{OD600_{sample}} - \frac{Fluor_{control}}{OD600_{control}}$$

Equation 3.1

$$\overline{Fluor_{norm_{off}}} = \frac{1}{n} \sum_{i=1}^n Fluor_{norm_i}$$

Equation 3.2

The *average fold repression* between the induced and uninduced state of synsrRNA was calculated using Equation 3.3. The $\overline{Fluor_{norm_{off}}}$ & $\overline{Fluor_{norm_{on}}}$ (Equation 3.2) is the average *EGFP/tdTomato* normalized fluorescence when synsrRNA is in uninduced and induced state, respectively.

$$average\ fold\ repression = \frac{\overline{Fluor_{norm_{off}}}}{\overline{Fluor_{norm_{on}}}}$$

Equation 3.3

Unpaired t-test was used to measure the extent of statistical significance. All data were analyzed and plotted using GraphPad Prism 8.

3.4.6 RNA purification and cDNA synthesis

We collected 4ml of bacterial cell suspension after 24 hrs from the HARV's and added twice the volume of RNAProtect bacteria reagent (Qiagen) and vortexed the mix for 5s. This mix was incubated at room temperature for 5mins, centrifuged for 12min at 5800 x g and supernatant was decanted. The Bacterial cell pellets were kept at -80 until RNA extraction. The bacterial cells were lysed using 15mg/ml lysozyme (Sigma) and proteinase-k (Qiagen) treatment (according to manufacturer's instruction in RNAProtect bacteria reagent handbook). The total RNA was extracted using RNeasy mini kit (Qiagen), following manufacturer's instruction. To lower the chance of genomic DNA contamination the optional on-column DNase digestion step was implemented using RNase free DNase (Qiagen), following manufacturer's instruction. For checking the integrity of the extracted RNA, it was run on denaturing agarose gel. The RNA was denatured by heating it at 65 °C with formamide, loading dye and EtBr for 5mins, chilled at ice for 5mins followed by running it on 1.2% agarose gel. The concentration of the RNA was measured using UV vis absorbance function of Biotek synergy HTX plate reader in the provided Take3 plate. One microgram of RNA from each sample was used for cDNA synthesis using Qiagen quantitect reverse transcription kit according to the manufacturer's instruction; the optional step for genomic DNA elimination reaction was also carried out.

3.4.7 qPCR and Data analysis

All oligos used for qPCR are listed in, and were designed in primer 3. The oligos were synthesized from Integrated DNA Technologies, Singapore. The qPCR was carried out using Qiagen Quantitect SYBR green PCR kit, according to manufacturer's instruction in Applied Biosystems step one plus real time PCR system in a 96 well plate. Total

Chapter three

reaction volume was 10 μ l. Amplifications were performed under the following conditions: 95 for 15mins, 40 cycles of 94 °C for 15s, 55 °C for 30s and 72°C for 30s and final step 95°C for 15s. After this a melting analysis was performed, which showed the specificity of the reactions. Atleast three biological replicates and three technical replicates were used for qPCR. The qPCR was performed for hfq and 16s rRNA as an endogenous reference gene as used previously in (Castro *et al.*, 2011). The difference in gene expression was calculated using $\Delta\Delta C_t$ method, which was normalized on basis of 16s rRNA expression. The data are plotted as mean and standard deviation in Graphpad Prism 8. The p value was 0.0052. A no template control showed the absence of detectable transcripts for both primers set.

3.4.8 Confocal microscopy

After 24 hrs 2ml bacterial cells suspension was harvested at 4000 rpm for 3mins and washed thrice in 1ml ice cold PBS. The cells were finally diluted in 1ml PBS enough to prevent crowding. To visualize, ten microliters cell suspension was put onto a one percent thin agarose pad (SeaKem LE agarose) and was covered with a cover slip. The DIC and fluorescence imaging was performed in Zeiss in LSM 710 confocal microscope, where the pinhole was kept partially open and the images were captured using Zen 2008 software. The cell suspension slides were subjected to excitation by appropriate laser channels (458 nm Ar 488 nm Ar Laser for EGFP and 543 nm He-Ne Laser for TdTomato) and fluorescence emissions were captured through proper emission filters (BP500-520 nm for EGFP, and BP 561-591 nm for TdTomato) with a 63x Oil immersion objective and were detected through a T-PMT. At least two slides were analysed for each sample.

Chapter 4: Development of a process pipeline for synthetic genetic circuit fabrication and its application in creating complex genetic circuit in single cell.

4.1 Introduction

In third chapter we already discussed the application of synthetic biology in manned space mission. The goal of synthetic biology is to achieve noble function, by engineering organisms, at DNA/RNA level following a hierarchy similar to electronics (Andrianantoandro, Basu, David K Karig, *et al.*, 2006). Implementing this fundamental in synthetic biology has resulted in development of interesting and useful devices and pathways (Basu *et al.*, 2005; Ro *et al.*, 2006; Friedland *et al.*, 2009; Fernandez-Rodriguez *et al.*, 2017). For reliable construction of complex genetic circuits, for space synthetic biology and synthetic biology as a whole, efficient assembly of genetic elements is required. Despite the availability of various DNA assembly techniques, it is still a limiting factor for most small laboratories (Ellis, Adie and Baldwin, 2011) like ours. This chapter is based on development of a process pipeline for synthetic genetic circuit fabrication and its application in creating complex genetic circuits.

4.1.1 Bioparts Assembly Techniques

Synthetic biology borrows engineering principles of abstraction and characterization for engineering or designing biological systems. Just like transistors, resistors in electronics, biological genetic ‘part’ can be reduced to DNA encoded functional units like promoters and genes (Andrianantoandro, Basu, David K Karig, *et al.*, 2006; Ellis, Adie

and Baldwin, 2011). These parts are combined together to rationally engineer new functions (Wang, Barahona and Buck, 2013; Fernandez-Rodriguez *et al.*, 2017). The successful engineering of an artificial molecular genetic device requires good kinetic or signal match among the functional modules. The chance of getting desirable output lowers due to incompatibility between the functional modules of a particular device. The fabrication of the proposed design needs a process pipeline for assembly of multiple DNA cassettes, fast optimization of the rate of gene expression, and reduction of unwanted kinetic influence. Here, genetic-cassette is defined as a genetic construct with ‘promoter-ribosome binding site-gene-transcription terminator’ or equivalent, where each of the non-reducible DNA entity like promoter or ribosome binding sites (RBS) is defined as bioparts.

Gold standard for such assembly and library construction of bio parts is ‘Bio-brick’ assembly method (Shetty, Endy and Knight, 2008). Biobrick is a DNA unit (e.g., promoter, ribosome binding site) similar to LEGO, with standardized flanking sequences, that allows easy assembly using restriction/ligation method Figure 4.1.

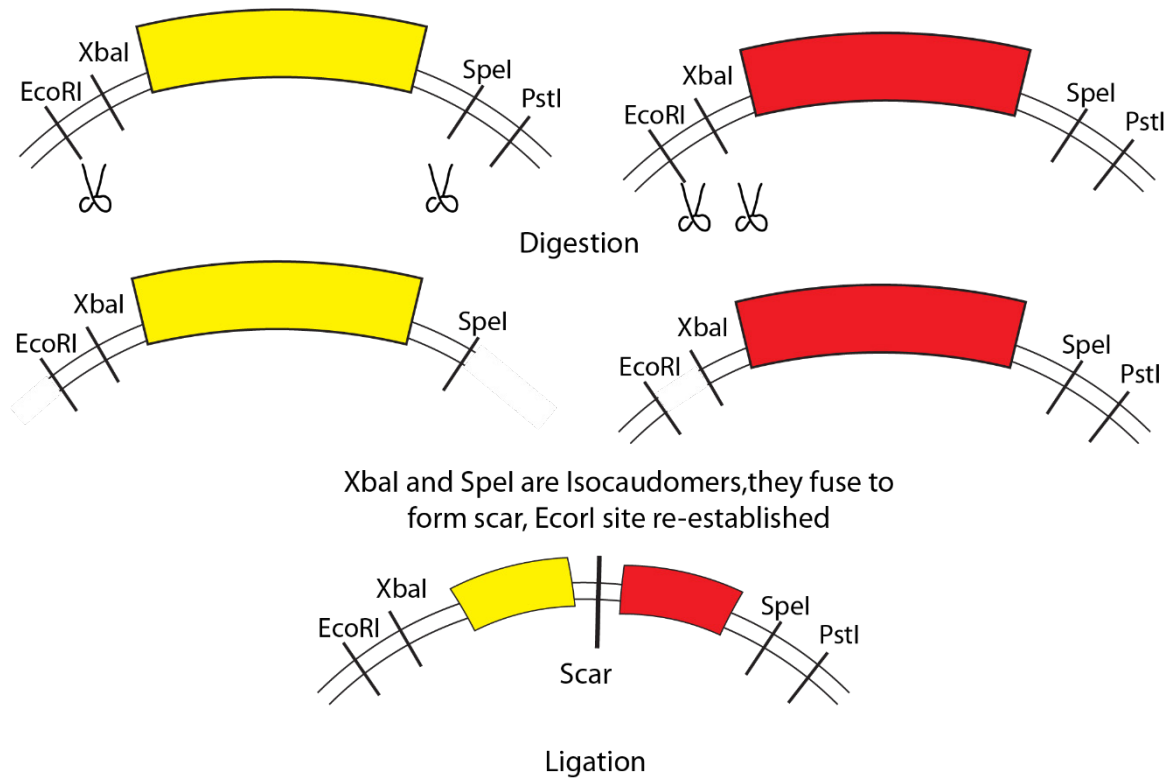


Figure 4.1 Biobrick assembly method

Demonstrates BioBrick cloning of parts through restriction digestion and ligation of the resulting compatible sticky ends, thus reforming the prefix sequence (EcoRI and XbaI) and leaving a 'scar' between parts. In BglBricks a modified scheme of biobrick replaces XbaI, SpeI, and PstI with BglII, BamHI, and XhoI, respectively, resulting in a scar that lacks a stop codon or frameshift.

This technique revolutionized the biological engineering by giving an easy and physical platform for engineers to apply the engineering paradigm of modularity and abstraction to develop new biological device at the genetic level (Cameron, Bashor and Collins, 2014). Several modified schemes of the original Biobrick system, including BglBricks (Anderson *et al.*, 2010), Silver Biofusion standard (Phillips and Silver, 2006) and Freiburg standard (Che, 2004) were developed. Main thrust of those modifications was to generate a proper biofusion for constructing multi-domain fusion proteins by optimizing the mixed site ('scar'). However, applying those strategies to optimize the kinetics of gene expression by altering RBS within an assembled cassette is not possible. This is due to the fact that the assembled cassette from biobrick principle does not have any

Chapter four

restriction site within bioparts (Anderson, Voigt and Arkin, 2007; Shetty, Endy and Knight, 2008). To address this optimization issue, option of customized inserts was introduced, but it is not that user-friendly (Norville *et al.*, 2010). Second, in biobrick strategy, two cassettes can only be assembled in a single direction. This uni-directionality results in a non-zero influence of a strong promoter in the upstream cassette on the kinetics of gene expression of the downstream cassette, even in the presence of transcription terminator (Chen *et al.*, 2013). Each step of an assembly generates scar between the parts; these are sometimes undesirable and can interfere with the functioning of parts together because of context-dependency (Salis, Mirsky and Voigt, 2009). Therefore, scarless assembly techniques were explored. One of which is Overlap extension polymerase chain reaction (OE-PCR), it uses overlapping PCR to create homologous ends between different DNA molecules, this homology is used to facilitate extension in a second round of PCR between the initial products (Horton *et al.*, 1989) and its variant (Quan and Tian, 2009). In a ground breaking work, JCVI successfully synthesized and assembled complete *Mycoplasma genitalium* genome, for such large scale assembly, they invented Gibson isothermal assembly, which is attractive because it can be automated using liquid handling systems (Gibson *et al.*, 2008, 2009; Alvarez, 2015). The method is shown in Figure 4.2. However, Gibson assembly doesn't allow reuse of parts between different experiments, which is not ideal for synthetic biology standards. Here, each part has to be flanked with overlapping extensions; thus, unique long primers have to be designed carefully for each individual part, therefore it is not possible to mix and match. For small labs like ours this process is inefficient, as full potential cannot be realized, in absence of automation, and will be too much for smaller routine projects.

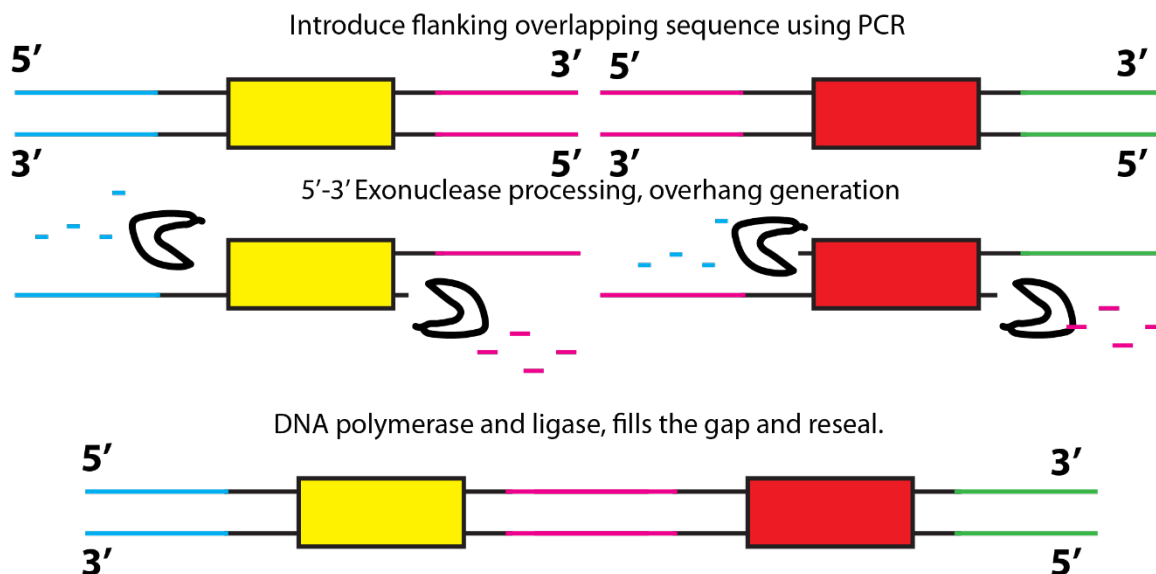


Figure 4.2 Gibson isothermal assembly

Parts are synthesized to overlap by 30+ bp and their ends are processed and fused together using an exonuclease, a ligase and a polymerase

4.1.2 Network brick

Here we introduce a new approach, which we named as Network Brick, it is a fabrication and optimization process pipeline which allows i) bidirectional assembly ii) parallel optimization. To demonstrate its features, we designed, fabricated, characterized and optimized a two input fully chemically controllable molecular genetic NOR, IMPLY gate in living single *Escherichia coli* (*E. coli*). The engineered *E. coli* can sense and process two environmental chemical signals and produced a fluorescent output by expressing a fluorescent protein following the IMPLY logic function. Our design is based on cascaded transcriptional regulation. The system is memory less and reversible. Further, following the same cascaded principle, we created and tested an integrated 2-input-2-output genetic circuits by cascading the IMPLY gate. We also developed a simple mathematical model captures the essential features of the device at the steady-state. We showed that those circuits are digital as well as mathematically predictive. Further a

Chapter four

PCR based Network Brick method, which does not require any gel separation step was developed and demonstrated by constructing a simple exchanger circuit. The name Network Brick was borrowed from the term ‘network plasmid’, first coined for constructing combinatorial genetic logic circuits in early days of synthetic biology (Guet *et al.*, 2002).

4.1.3 Synthetic Genetic IMPLY gate

The ability to compute the environmental chemical signals by bacteria with synthetic molecular genetic devices is paving the way of advanced molecular computation, future smart biosensors, micro-biorobotics and their numerous applications (Andrianantoandro, Basu, David K. Karig, *et al.*, 2006; Kim, Steager and Agung, 2012; Ashkenasy *et al.*, 2017; Goñi-Moreno and Nikel, 2019). The basic logic processing including NOT, BUFFER, AND, OR, NAND, NOR, XOR, and XNOR has been demonstrated in living bacterial cells and in the last 15 years, various designs and implementation strategies for almost all logic gates have been realized (Moon *et al.*, 2012; Bonnet *et al.*, 2013; Siuti, Yazbek and Lu, 2013; Nomura and Yokobayashi, 2015; Jusiak *et al.*, 2016; Green *et al.*, 2017). Here the inputs are environmental chemical signals, which are sensed, integrated and computed inside cells by various molecular modalities including transcription factors, recombinases, CRISPRi/dCas9, RNA and rybozymes and produced output by expressing proteins, nucleic acids or metabolites (Moon *et al.*, 2012; Bonnet *et al.*, 2013; Siuti, Yazbek and Lu, 2013; Nomura and Yokobayashi, 2015; Jusiak *et al.*, 2016; Green *et al.*, 2017).

IMPLY is one of the most fundamental four logic functions (AND, OR, NOT and IMPLY), which was proposed in 1910 by Whitehead and Russell in *Principia Mathematica* (Whitehead, 2005). Unlike the other three logic functions, IMPLY is a

Chapter four

constraint-based logic and is defined by its theorem as, "If 'P' is TRUE then output follows 'Q' otherwise it is TRUE (Whitehead, 2005; Borghetti *et al.*, 2010). IMPLY has applications in robotics, control systems, and artificial intelligence (Kampis, 1991; Swart H.C.M. and Nederpelt, 1992; Castelfranchi and Lespérance, 2001; Borghetti *et al.*, 2010). In principle, any circuit can be made from universal NOR or NAND gates. However, from a practical point of view, it is not always the case. For example, to create an IMPLY function, 3 NOT gates are required. Therefore, the application of IMPLY gates is more useful for circuits like decoder or adder, instead of IMPLY equivalent three NOR gates. This is specifically important for creating large and complex synthetic genetic circuits, where an increasing number of genetic components and cascaded steps are often hinder the correct functionality of the circuits. An IMPLY type behavior in bacterium, was reported before (Weiss, Weiss and Basu, 2002; Yokobayashi, Weiss and Arnold, 2002) where only one environmental chemical (Isopropyl β -D-1-thiogalactopyranoside, IPTG) signal was sensed and processed whereas the other input was a constitutively expressed intracellular protein (lacI) (Weiss, Weiss and Basu, 2002; Yokobayashi, Weiss and Arnold, 2002). This intracellular input protein (lacI) was always present in a high amount (input logic level 1) and thus it cannot be brought to logic level '0' (Weiss, Weiss and Basu, 2002; Yokobayashi, Weiss and Arnold, 2002). As a result, the 0,0 and 1,0 input logic levels, with respect to IPTG and lacI respectively could not be realized. Further, IPTG and lacI cannot be treated as independent inputs as the function of lacI depends on IPTG. Thus it makes it limited spectrum IMPLY gate. Full-spectrum IMPLY gates were constructed by combining multiple NOR gates by engineering bacterial communications, where 3 different cell populations were required to construct the IMPLY gates (Tamsir, Tabor and Voigt, 2011). N-IMPLY (inverted

output of IMPLY) gate was reported in bacterial and mammalian cells (Guet *et al.*, 2002; Ausländer *et al.*, 2012). Further, using a serine recombinase based system, IMPLY gates were constructed in single *E.coli* cells (Siuti, Yazbek and Lu, 2013). Whereas those constructs have important applications in state dependent circuits, it cannot be used to develop combinational circuits, which depend on the input values at a given instance.

Here, by using Network brick, we designed, fabricated, characterized and optimized a two input fully chemically controllable molecular genetic IMPLY gate in living single *Escherichia coli* (*E. coli*). The engineered *E. coli* can sense and process two environmental chemical signals and produced a fluorescent output by expressing a fluorescent protein following the IMPLY logic function. Our design is based on cascaded transcriptional regulation. Further, following the same cascaded principle, we created and tested an integrated 2-input-2-output genetic circuits by cascading the IMPLY gate. As such transcriptional circuit requires several bioparts in harmony, one of the fundamental challenges to make such circuit works is the optimization of various molecular genetic parts, such that they work together as a functional device. In order to fabricate the circuit and optimize its systems chemistry inside a cell, we used Network brick, for systematic fabrication, fast kinetic optimization and elimination of unwanted kinetic influence from one set of biochemical modules to other.

4.2 Results and Discussion

4.2.1 Network brick base plasmid design and construction

First, an empty plasmid vector (Figure 4.4), named as forward cassette empty vector (FCEV) was created. The basic architecture of the Network Brick plasmid was adapted

from the pZ vector system (Lutz and Bujard, 1997a), which was widely used for building synthetic genetic circuits, screening and gene expression studies (Stricker *et al.*, 2008; Bagh, Mandal and McMillen, 2010; Bagh *et al.*, 2011b; Genée *et al.*, 2016; Nyerges *et al.*, 2016; Rubens, Selvaggio and Lu, 2016).

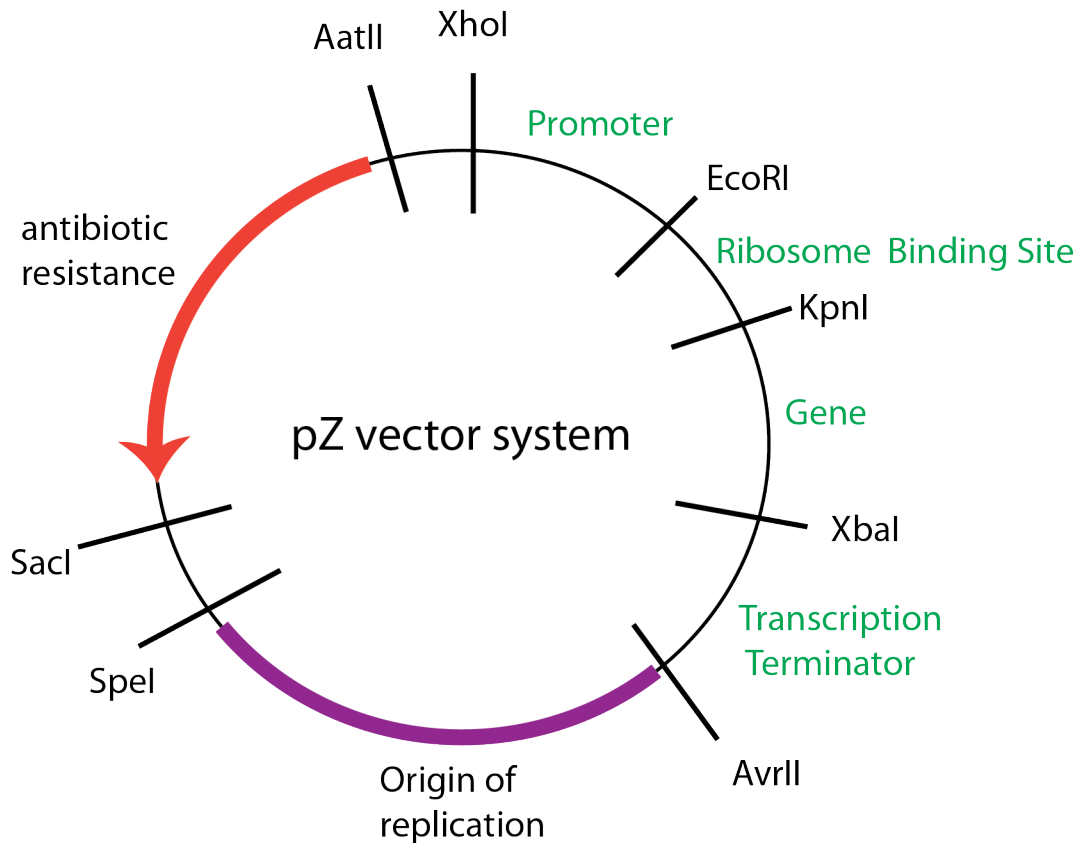


Figure 4.3 pZ vector system for cloning and expression (Lutz and Bujard, 1997a). The promoter, ribosome binding site, gene, replication origin, transcription terminator, and antibiotic resistance gene are all illustrated together with their flanking restriction sites.

The empty Network brick plasmid consists of three parts prefix, multiple cloning site (MCS) and suffix. The prefix and suffix fragment were chemically synthesized and cloned upstream and downstream of pZ vector MCS. The MCS harbors vacant positions flanked with unique restriction sites for insertion of bioparts namely promoter, ribosome binding site (RBS), coding sequence or gene of interest and transcription terminators. A complete genetic cassette consists of all the bioparts arranged in the above

Chapter four

aforementioned order. The bioparts can be chemically synthesized or amplified from other source and it must be flanked by the appropriate restriction sites. The biopart and the FCEV thus could be doubly digested with appropriate restriction site, gel purified and finally ligated. This architecture thus ensures that any of the bioparts within a cassette could be changed as many times as required by using the two appropriate restriction enzymes without starting from the scratch (Figure 4.4). Next, biobrick type prefix (PstI, BamHI) and suffix (BglI, Sall) restriction sites were introduced upstream and downstream of a DNA cassette, respectively (Figure 4.5). This allowed assembling multiple cassettes together into a single direction using only those 4 enzymes (Figure 4.5) by isocaudomeric principle (Shetty, Endy and Knight, 2008). FCEVs with three different origins of replication were created (Figure 4.4). A second plasmid system, named as reverse cassette empty vector (RCEV), was created (Figure 4.6). In RCEV, positions of the first and last restriction sites of the cassette (XhoI and AvrII respectively) were swapped, such that if a cassette from FCEV was incorporated in RCEV, the 5'-3' sequence direction of the cassette would be changed. Thus, assembling this reverse cassette from RCEV with forward cassette in FCEV resulted in the two cassettes in the opposite direction (Figure 4.6). Thus, Network brick system would allow integrating multiple DNA cassettes in any direction and parallel optimization of a DNA-biopart within a cassette. Here we showed a comparison between Network brick and biobrick for constructing, assembling and optimization of two genetic cassettes (Table 4.1). A detailed step-by-step process pipeline of Network brick for constructing IMPLY gates and other gates (next section) is already shown in Table 4.2.

Chapter four

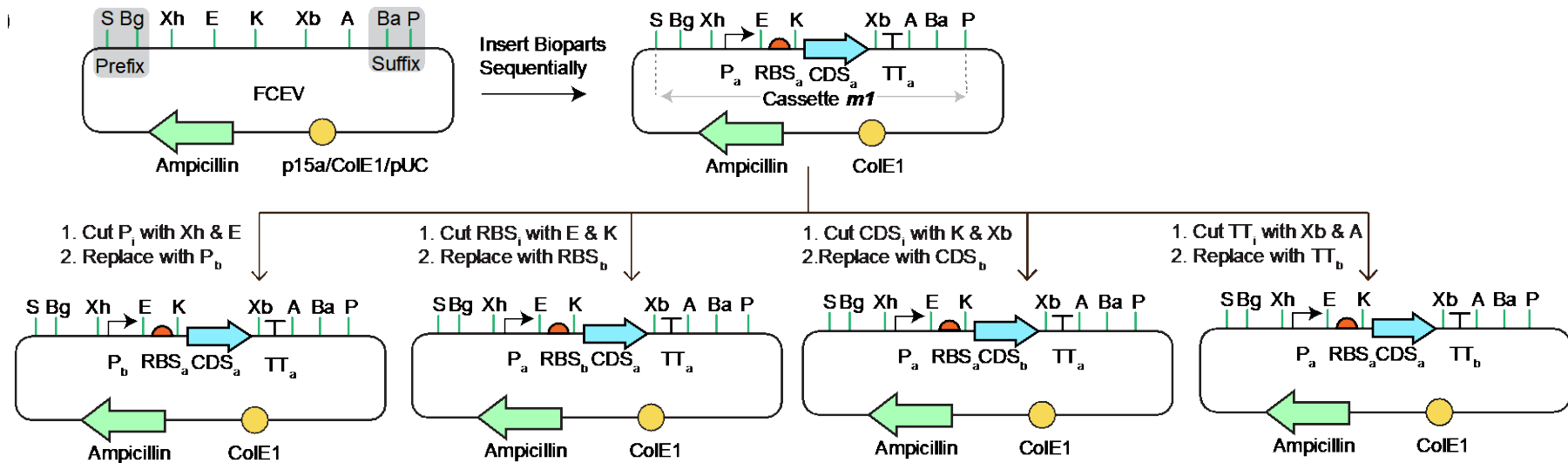


Figure 4.4. Network Brick, the new biochemical process pipeline.

FCEV with its restriction sites. The biobrick type prefix and suffix consist SalI (S), BglIII (Bg) and BamHI (Ba), PstI (P) respectively. The empty cassette consists of Xho-I (Xh), EcoRI (E), KpnI (K), XbaI (Xb) and AvrII (A). P, RBS, CDS and TT represent promoter, ribosome binding site, gene and transcription terminator DNA sequences respectively. The process pipeline of inserting individual biopart(s) in FCEV is also shown.

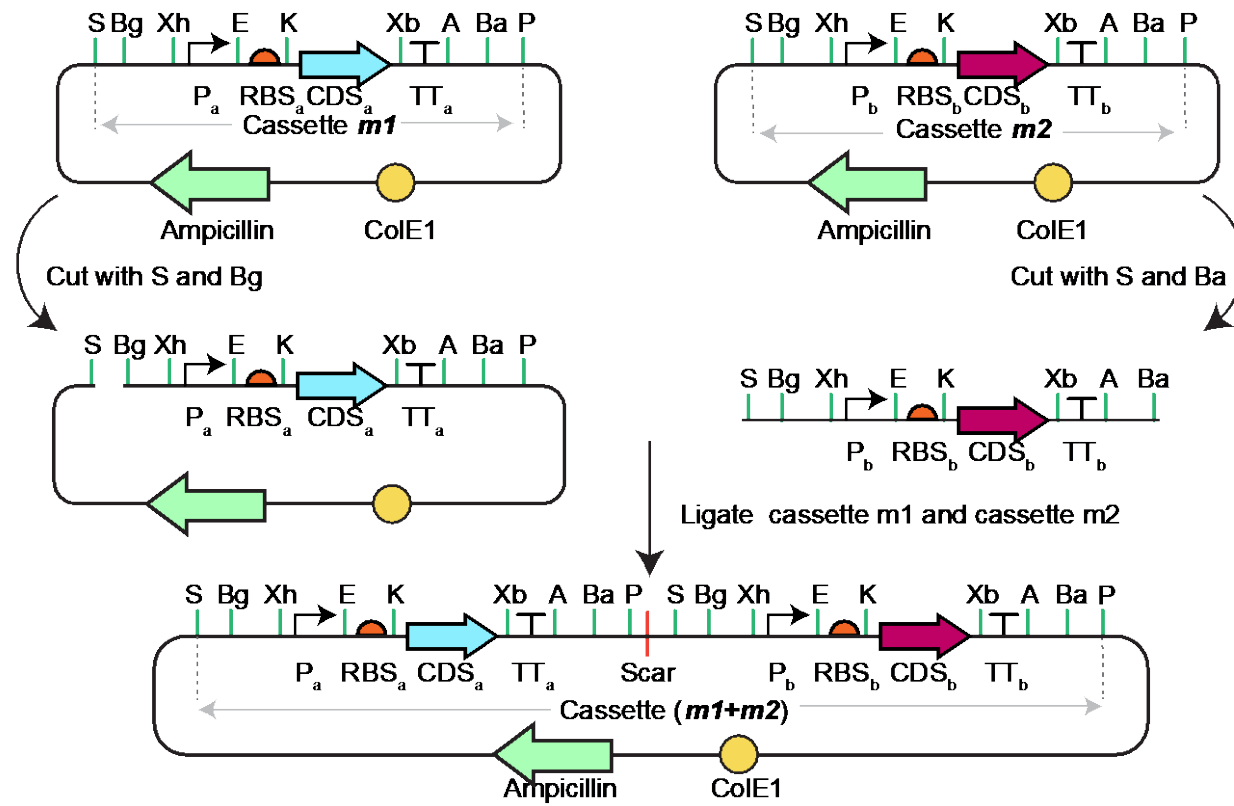


Figure 4.5 Process pipeline for assembling two genetic cassettes in a single direction.

Chapter four

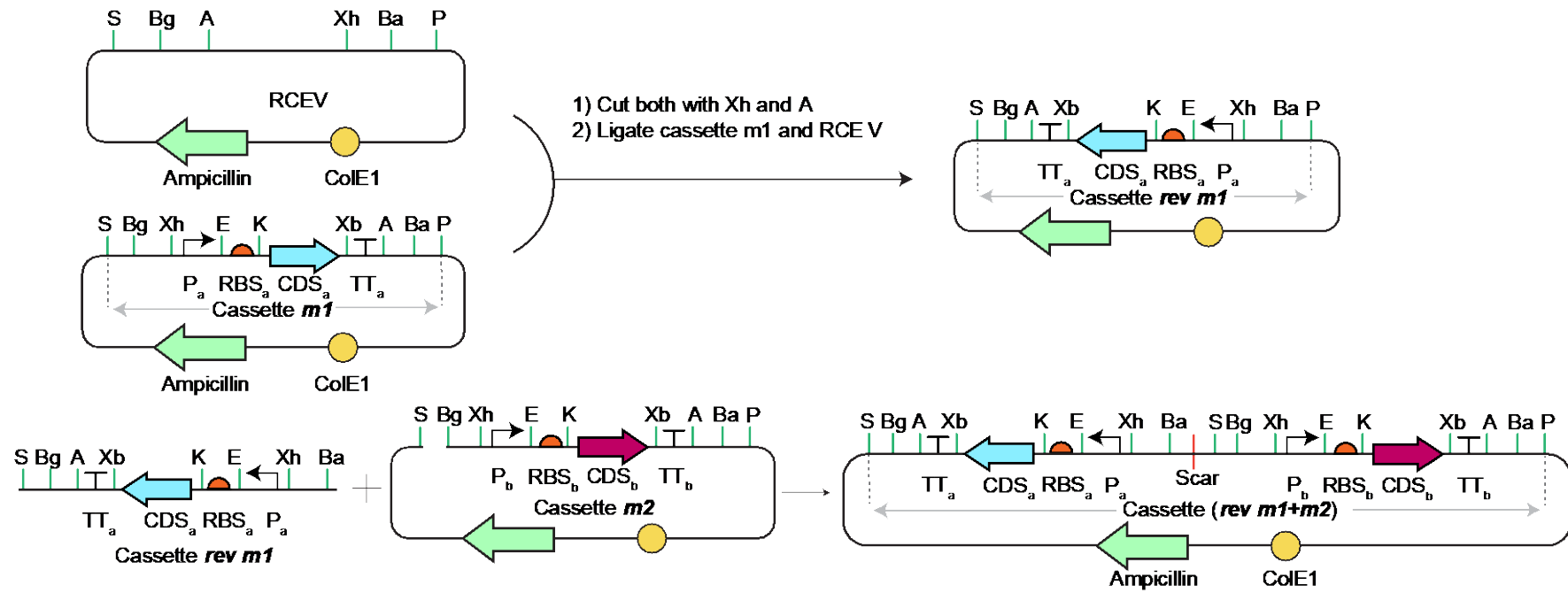


Figure 4.6 Schematic of RCEV and process pipeline for assembling two genetic cassettes in two opposite directions.

Table 4.1. Comparison between Network Brick and Biobrick for construction & high throughput optimization of hypothetical constructs in three cases.

Biobrick	Network brick	
Case I: Construction of the the construct --P₁-R₁-G₁-T₁-/- P₂-R₂-G₂-T₂-- (forward assembly), where P,R,G and T signify promoter, RBS, gene and transcript terminator respectively.		<i>Both Bio-brick and Network brick require same number of steps and reactions</i>
<i>I. Construction of P₁-R₁-G₁-T₁</i> Step1: T ₁ Step2: G ₁ -T ₁ Step3: R ₁ -G ₁ -T ₁ Step4: P ₁ -R ₁ -G ₁ -T ₁ <i>II. Construction of P₂-R₂-G₂-T₂</i> Step5: T ₂ Step6: G ₂ -T ₂ Step7: R ₂ -G ₂ -T ₂ Step8: P ₂ -R ₂ -G ₂ -T ₂ <i>III. Construction of -- P₁-R₁-G₁-T₁-/-P₂-R₂-G₂-T₂--</i> Step9: P ₁ -R ₁ -G ₁ -T ₁ -/- P ₂ -R ₂ -G ₂ -T ₂	<i>I. Construction of P₁-R₁-G₁-T₁</i> Step1: T ₁ Step2: G ₁ -T ₁ Step3: R ₁ -G ₁ -T ₁ Step4: P ₁ -R ₁ -G ₁ -T ₁ <i>II. Construction of P₂-R₂-G₂-T₂</i> Step5: T ₂ Step6: G ₂ -T ₂ Step7: R ₂ -G ₂ -T ₂ Step8: P ₂ -R ₂ -G ₂ -T ₂ <i>III. Construction of -- P₁-R₁-G₁-T₁-/-P₂-R₂-G₂-T₂--</i> Step9: P ₁ -R ₁ -G ₁ -T ₁ -/- P ₂ -R ₂ -G ₂ -T ₂	
Total number of steps and re- actions=9	Total number of steps and reactions=9	
Case II: High throuput optimization of RBS (R1) in the construct -P₁-(R₁)-G₁-T₁-/- P₂-R₂-G₂-T₂-- ,		
<i>IV. Optimisation of P₁-R₁-G₁-T₁</i>	<i>IV. Optimisation of P₁-R₁-G₁-T₁</i> Step10 High throughput parallel op- timization reactions	

<p>Step10 Construction of G1-T1 G1-T1</p> <p>Step 11 High throughput parallel optimization reactions Assembling R1 to R100 with G1-T1 (100 reactions)</p> <p>1: R_{1_1}-G₁-T₁ 2: R_{1_2}-G₁-T₁ 100: R_{1_100}-G₁-T₁</p> <p>Step12: Assembling P1 with each Ri-G1-T1 (100 reactions)</p> <p>1: P₁-R_{1_1}-G₁-T₁ 2: P₁-R_{1_2}-G₁-T₁ 100: P₁-R_{1_100}-G₁-T₁</p>	<p>Cut R1 from cassette P₁-R₁-G₁-T₁ from step 4 with two restriction enzymes on both of its sides (EcorI and KpNI)</p> <p>1: P₁-R_{1_1}-G₁-T₁ 2: P₁-R_{1_2}-G₁-T₁ 100: P₁-R_{1_100}-G₁-T₁</p>	<p>Network brick requires less steps and half the reactions in comparison to Bibrick.</p>
<p>Total No. of steps = 3 Total No. of reactions = 200</p>	<p>Total No. of steps = 1 Total No. of reactions = 100</p>	
<p>Case III Bidirectional assembly --T₁-G₁-R₁[*]-P₁-/- P₂-R₂-G₂-T₂--</p>		

<p><i>I. Construction and optimisation of $P_1-R_1-G_1-T_1$</i></p> <p>Step1: T_1</p> <p>Step2: G_1-T_1</p> <p>Step3: $R_1-G_1-T_1$</p> <p>Step4: $P_1-R_1-G_1-T_1$</p> <p><i>II. Reversing orientation of $P_1-R_1-G_1-T_1$</i></p> <p>Step5: Not possible</p>	<p><i>I. Construction and optimisation of $P_1-R_1-G_1-T_1$</i></p> <p>Step1: T_1</p> <p>Step2: G_1-T_1</p> <p>Step3: $R_1-G_1-T_1$</p> <p>Step4: $P_1-R_1-G_1-T_1$</p> <p><i>II. Reversing orientation of $P_1-R_1-G_1-T_1$ with reverse vector</i></p> <p>Step5: $T_1-G_1-R_1-P_1$</p> <p><i>III. Construction of $P_2-R_2-G_2-T_2$</i></p> <p>Step6: T_2</p> <p>Step7: G_2-T_2</p> <p>Step8: $R_2-G_2-T_2$</p> <p>Step9: $P_2-R_2-G_2-T_2$</p> <p><i>IV. Construction of $--T_1-G_1-R_1^*-P_1-/-P_2-R_2-G_2-T_2--$</i></p> <p>Step10: $T_1-G_1-R_1-P_1-/-P_2-R_2-G_2-T_2$</p>	<p><i>Bidirectional assembly is only possible with Network brick</i></p>
<p style="text-align: center;">Advantage of Network brick over Biobrick</p> <p>1. No. of steps: The no. of steps is less when same biopart(s) is used in different cassette. Thus, it takes <i>less time</i>.</p> <p>2. No. of reaction: Clearly, the no. of reaction is drastically less in Network brick during optimization. Thus, <i>less consumables</i> are required.</p>		

3. Bidirectionality: Two consecutive cassettes can be assembled in a opposite orientation. Thus, it provides *transcriptional insulation*. Bidirectional assembly is not possible in Biobrick

4.2.2 Design, fabrication and characterization of NOR logic gate

First, we used the Network brick principle and plasmids to construct a synthetic genetic NOR gate. The NOR or Negative OR gate is the inverse of OR gate and it is one of the most fundamental four logic functions (AND, OR, NOT and IMPLY). In principle, any circuit can be made from universal NOR and NAND gate. The Apollo guidance system in 1960's was one of the first technology to implement integrated circuits entirely made from NOR gates (Hall, 1996). We used the Network brick pipeline to fabricate the NOR gate as shown in Table 4.2. The design and experimental behavior of the NOT gate is shown in Figure 4.7. The electronic equivalence (Figure 4.7 A) of the molecular design (Figure 4.7 B) is based on modular DNA-parts and its specific interactions with transcription factors and RNA polymerase. One enhanced green fluorescent protein (EGFP) gene (output) was placed under a P_R promoter. The promoter P_R is inhibited by a mutated version of a lambda repressor CI protein (CI_m) (Sarkar *et al.*, 2019) .

The CI_m gene was under the control of ATC inducible PLtetO-1 promoter and PLlacO-1, both in distinct cassettes. Small molecule inducers IPTG and anhydrotetracycline (ATC) were sensed and interacted allosterically with LacI and tetR respectively and made a conformational change (Lutz and Bujard, 1997a), which could no longer inhibit the PLlacO-1 and PLtetO-1 promoters respectively. We characterized the NOR logic gate accordingly as mentioned in Materials and Methods. The experimental result (Figure 4.7 D) was in harmony with the truth table (Figure 4.7 C).

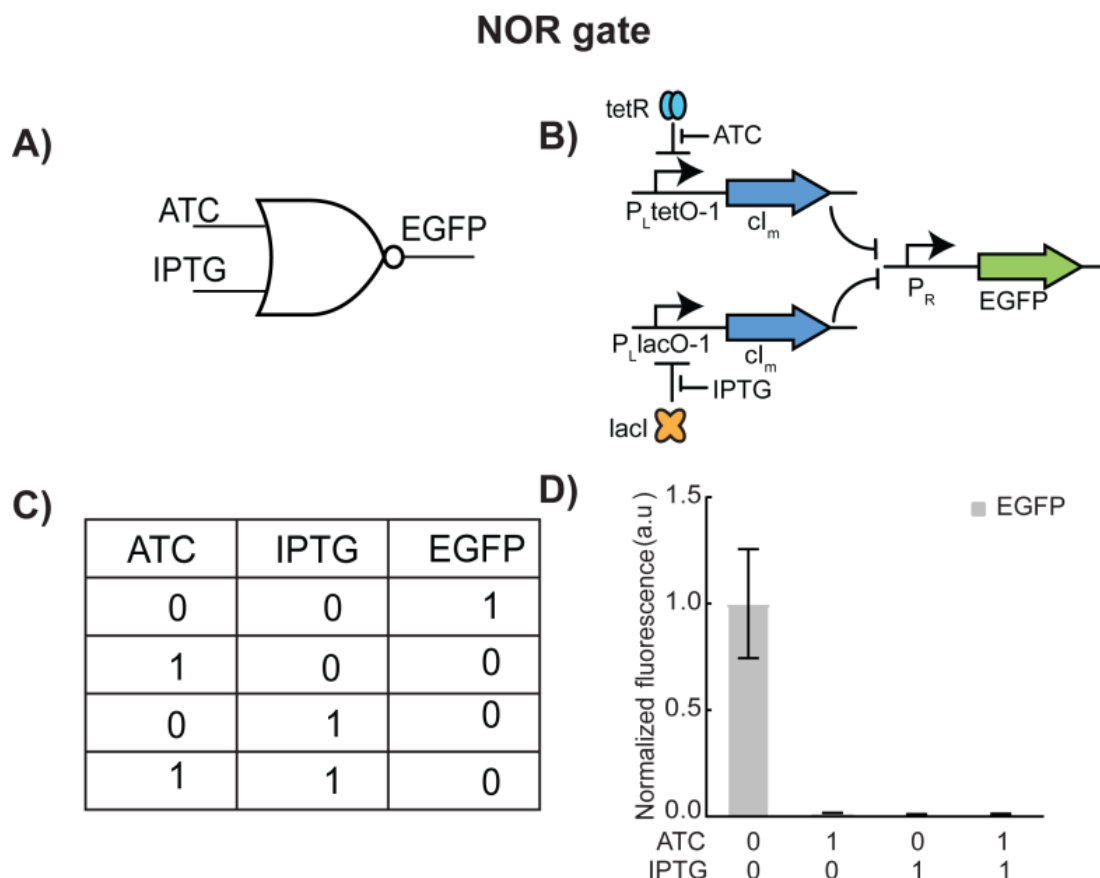
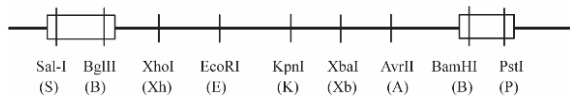
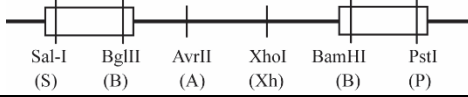
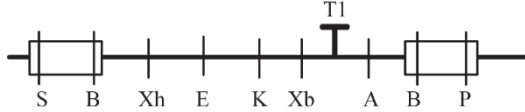
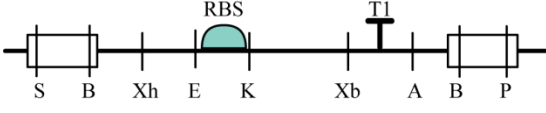
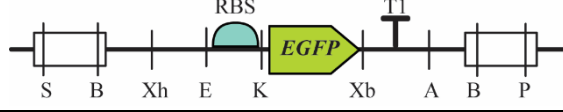
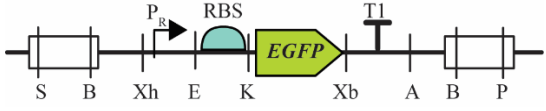


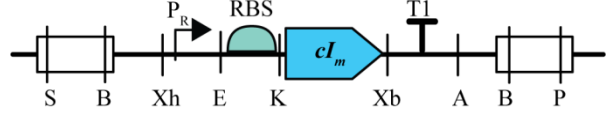
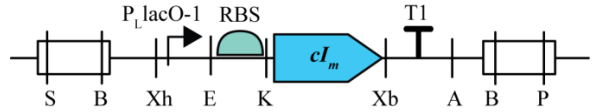
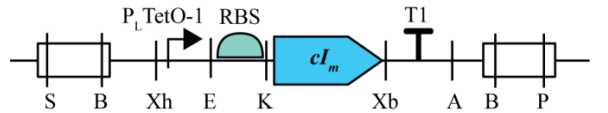
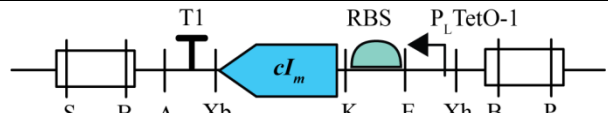
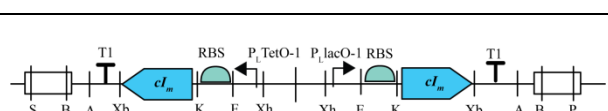
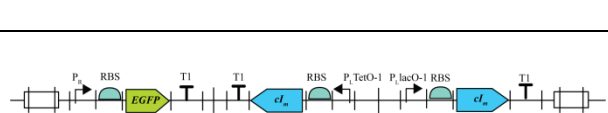


Figure 4.7. Design and testing of molecular NOR gate.

A) Electronic equivalent representation and truth table. **B)** Biomolecular design. One mutated lambda repressor (cIm) is kept under ATC inducible $P_{LtetO-1}$ promoter another cIm is kept under IPTG inducible $P_{LlacO-1}$ promoter and the reporter gene EGFP is kept under lambda repressible P_R promoter. The tet repressor (tetR) and lac inhibitor (lacI) is constitutively expressed in DH5 α 1 *E.coli* cell strain. **C)** Experimental behaviour before optimization and **D)** experimental behaviour. Where each bar represents mean normalized EGFP fluorescence at different combinations of inducers at saturating concentrations. Error bar represents standard deviation from the mean. At least three biological replicates were used.

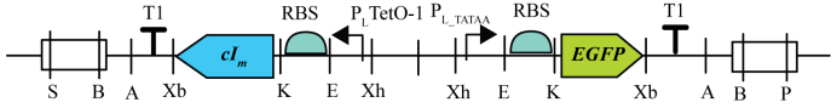
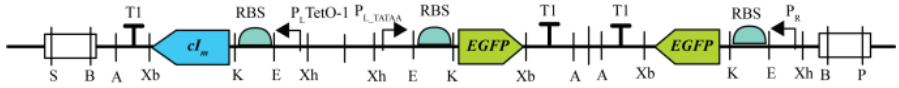
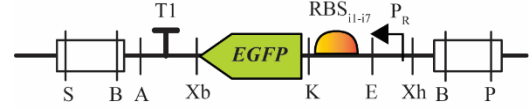
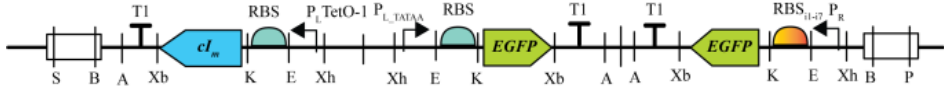
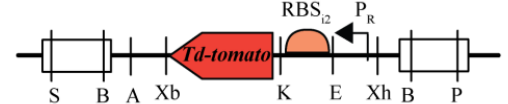
Table 4.2. Construction steps of gene circuits using Network Brick

Device	Biopart or Cassette number	Design	Construction steps	Remarks
Forward empty (FCEV)	cassette vector	Forward cassette empty vector (FCEV) 	Synthesized	
Reverse empty (RCEV)	cassette vector	Reverse cassette empty vector (FCEV) 	Synthesized	
For Creating Synthetic NOR Gate	Biopart 1		Inserted T1 terminator in FCEV between XbaI and AvrII	
	Biopart 2		Inserted Ribosome binding site (RBS) in biopart 1 between EcoRI and KpnI	
	Biopart 3		Inserted <i>EGFP</i> in Biopart 2 between KpnI and XbaI	
	Cassette 1		Inserted P_R promoter in Biopart 3 between XhoI and EcoRI	

Chapter four

	Cassette 2		Replaced <i>EGFP</i> in Cassette 1 with <i>cI</i> (mutant) between KpnI and XbaI	
	Cassette 3		Replaced P _R in Cassette 2 with P _L -lacO-1 between Xho and EcoRI	
	Cassette 4		Replaced P _R in Cassette 2 with P _L -tetO-1 between XhoI and EcoRI pair	
	Cassette 5		Inserted Cassette 4 in Reverse cassette empty vector (RCEV) between XhoI and AvrII	
	Cassette 6		Assembled SalI and BamHI digested cassette 5 with SalI and BglII digested cassette 3	
	Cassette 7		Assembled SalI and BamHI digested cassette 1 with SalI and BglII digested cassette 6	
For Creating Synthetic Gate IMPLY	Cassette 8		Inserted Cassette 1 in RCEV between XhoI and Avr II	
	Cassette 9		Replaced 'P _R ' in Cassette 1 with P _L TATAA between XhoI and EcoRI	

Chapter four

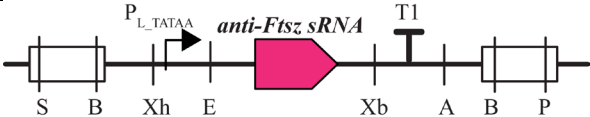
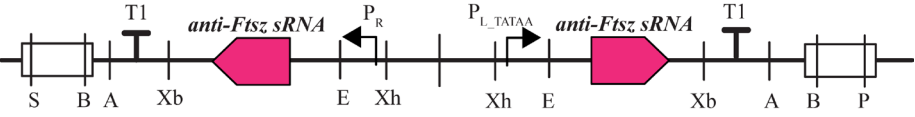
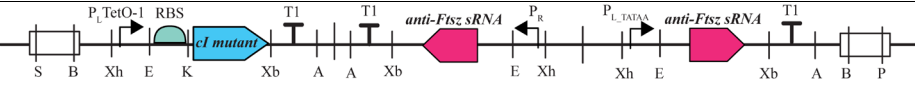
	Cassette 10		Assembled Sall and BamHI digested cassette 5 with Sall and BglII digested vector containing cassette 9	
	Cassette 11		Assembled BglII and PstI digested cassette 8 with BamHI and PstI digested cassette 10	Bug: Unequal Outputs: EGFP expression from P _R promoter is very high. Debug step: Lower EGFP translation rate between designing RBS with lower predicted length
	Cassette(s) 12i1 to 12i7		Inserted Modified RBS i1 to i7 in cassette 8 between EcorI and KpnI	
	Cassette(s) 13i1 to 13i7		Assembled BglII and PstI digested Cassette(s) 12i1 to 12i7 with BamHI and PstI digested cassette 10	
For Creating Synthetic NOT Gate &	Cassette 14		Replaced EGFP with Td tomato in cassette 12i2 between KpnI and XbaI	

Chapter four

For creating Integrated circuit (Two input two output)	Cassette 15		Assembled SalI and BamHI digested cassette 4 with SalI and BglII digested vector containing cassette 14	
	Cassette 16		Assembled SalI and BamHI digested cassette 15 with SalI and BglII vector containing cassette 13 i2	Bug: Leakage of cI protein leading to unequal outputs Debug: Reduce the copy number
Library Vector			synthesized	
For Creating PCR Based Ex-changer circuit	Promoter Library 1		Inserted P _L lacO-1 in library vector between XhoI and EcoRI	
	Promoter Library 2		Inserted P _L tetO-1 in library vector between XhoI and EcoRI	
	CDS library 1		Inserted EGFP in library vector between KpnI and XbaI	
	CDS library 2		Inserted Td-tomato in library vector between KpnI and XbaI	

Chapter four

	Biopart 5		Inserted <i>Td-tomato</i> in Biopart_2 between KpnI and XbaI digestion	
	Biopart 6		Inserted <i>EGFP</i> in Biopart_2 between KpnI and XbaI digestion	
	Cassette 17		Inserted P _L lacO-1 in Biopart 5 between XhoI and EcoRI	
	Cassette 18		Inserted P _L tetO-1 in Biopart 6 between XhoI and EcoRI	
	Cassette 19		Inserted Cassette 18 in RCEV between XhoI and Avr II digestion	
	Cassette 20		Assembled SalI and BamHI digested cassette 19 with SalI and BglII digested vector containing cassette 17	
For Creating Synthetic IMPLY Gate with anti Ftsz sRNA	Cassette 21		Replaced <i>EGFP</i> with anti Ftsz sRNA in cassette 8 between EcoRI and Xba-I	

	Cassette 22		Replaced <i>EGFP</i> with anti Ftsz sRNA in cassette 9 between EcoRI and Xba-I	
	Cassette 23		Assembled SalI and BamHI digested cassette 21 with SalI and BglIII digested vector containing cassette 22	
	Cassette 24		Assembled SalI and BamHI digested cassette 4 with SalI and BglIII digested vector containing cassette 23	

4.2.3 Design and fabrication of IMPLY logic for sensing and processing two extracellular environmental chemicals

The stepwise construction steps of the IMPLY gate using Network Brick are shown in (Table 4.2). The design and experimental behavior of the IMPLY gate is shown in Figure 4.8. The electronic equivalence (Figure 4.8 A) of the molecular design (Figure 4.8 B) is based on modular DNA-parts and its specific interactions with transcription factors and RNA polymerase (cassette 11). One enhanced green fluorescent protein (EGFP) gene (output) was placed under a synthetic hybrid promoter PL_{TATAA} (Sarkar *et al.*, 2019), made earlier in our lab. The promoter PL_{TATAA} worked such a way that lac repressor (LacI) or Tet repressor (tetR) proteins could bind and inhibit its function. Small molecule inducers IPTG and anhydrotetracycline (ATC) were sensed and interacted allosterically with LacI and tetR respectively and made a conformational change (Lutz and Bujard, 1997b), which could no longer inhibit the promoter PL_{TATAA} . Another EGFP gene was placed under P_R promoter, which got inhibited by a mutated version of a lambda repressor CI protein (CI_m) (Sarkar *et al.*, 2019). The CI_m gene was under the control of ATC inducible PL_{tetO-1} promoter (Sarkar *et al.*, 2019).

The biomolecular design of IMPLY gate (Figure 4.8) is generic and can be used to create IMPLY gates, which may sense and process other environmental signals. Next, we fabricated the design using Network Brick. The fabrication of the design and its optimization (Section 4.2.4) showed imperfect (Figure 4.8 C) and matching (Figure 4.8 D) behavior. The details can be found in the next section.

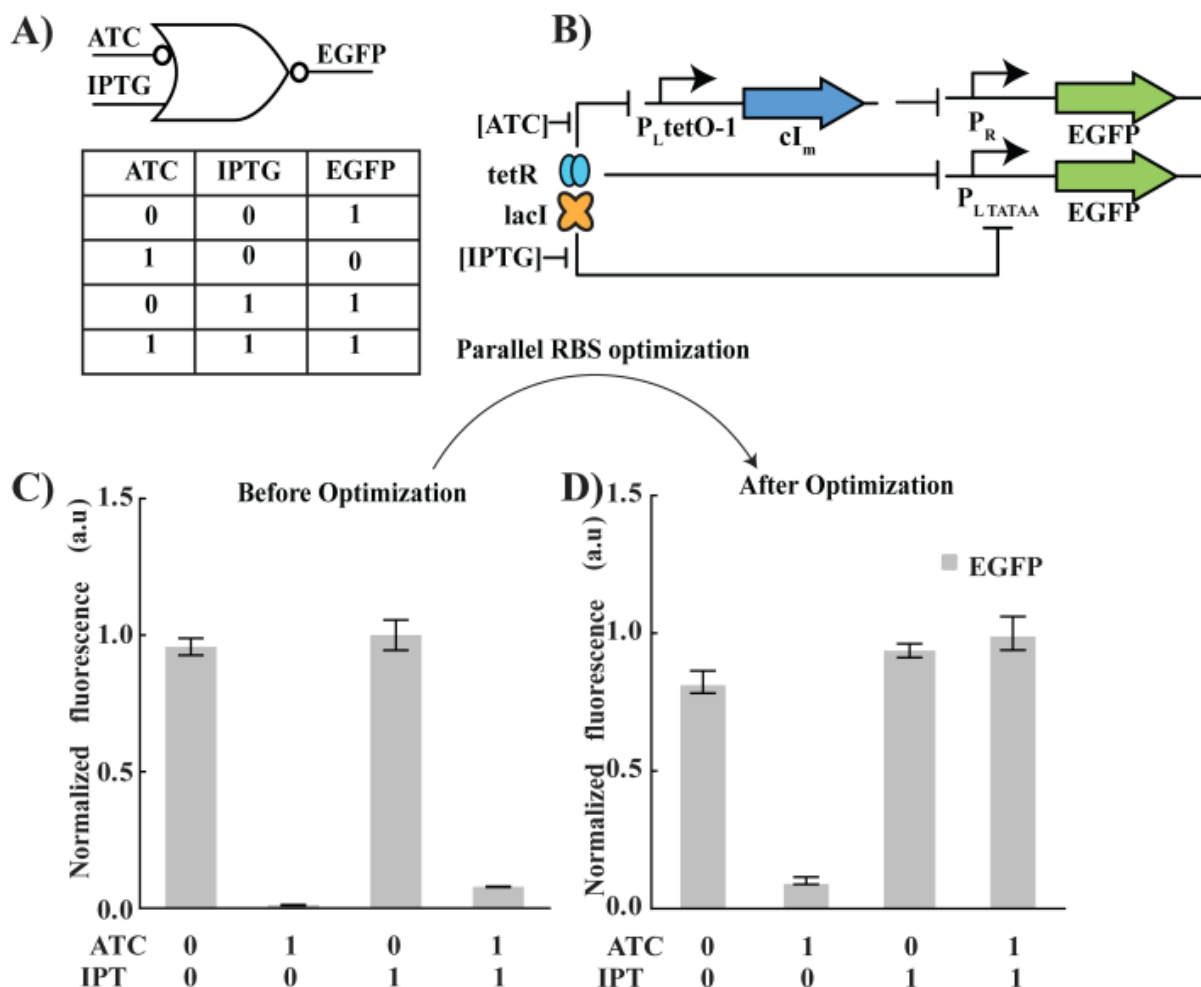


Figure 4.8. Design and testing of molecular IMPLY gate.

A) Electronic equivalent representation and truth table. **B)** Biomolecular design. The mutated lambda repressor (cI_m) is kept under ATC inducible $P_{LtetO-1}$ promoter and the reporter gene EGFP is kept under lambda repressible P_R promoter as well as P_{LTATAA} (hybrid promoter, with operator sites for lac repressor and tet repressor both). The tet repressor ($tetR$) and lac inhibitor ($lacI$) is constitutively expressed in DH5 α 1 *E.coli* cell strain **C)** Experimental behaviour before optimization, Where each bar represents mean normalized EGFP fluorescence at different combinations of inducers at saturating concentrations. Error bar represents standard deviation from the mean. At least three biological replicates were used. **D)** Experimental behaviour after optimization (RBS weakening of P_R -EGFP cassette).

4.2.4 Characterization and optimization of IMPLY gate

To characterize the behaviour of the IMPLY gate we transformed the plasmid into *E. coli* DH5 α Z1 cell strain, which constitutively produced TetR and LacI proteins (Lutz

Chapter four

and Bujard, 1997a). Next, cells were grown in appropriate condition (Section 4.4.2) and experimental characterization of IMPLY logic circuit (cassette 11) was performed in the absence or presence of input chemical signals, ATC (0 and 200 ng/mL) and IPTG (0 and 10 mM), highest concentration is saturating in *E.coli* Dh5 α Z1. The EGFP fluorescence was measured as a logic output. The experimental behavior was plotted in Figure 4.8 C, which showed an imperfect behavior for an IMPLY gate. Failure analysis on this device suggested that the sensing of the two environmental chemicals worked. However, during the signal processing the gene expression rate of output EGFP from P_R promoter was much higher than the hybrid promoter P_{LTATAA} (Figure 4.9). Therefore, one of the crucial design criteria was to reduce the EGFP formation rate from the P_R promoter to match the expression from the hybrid promoter P_{LTATAA} .

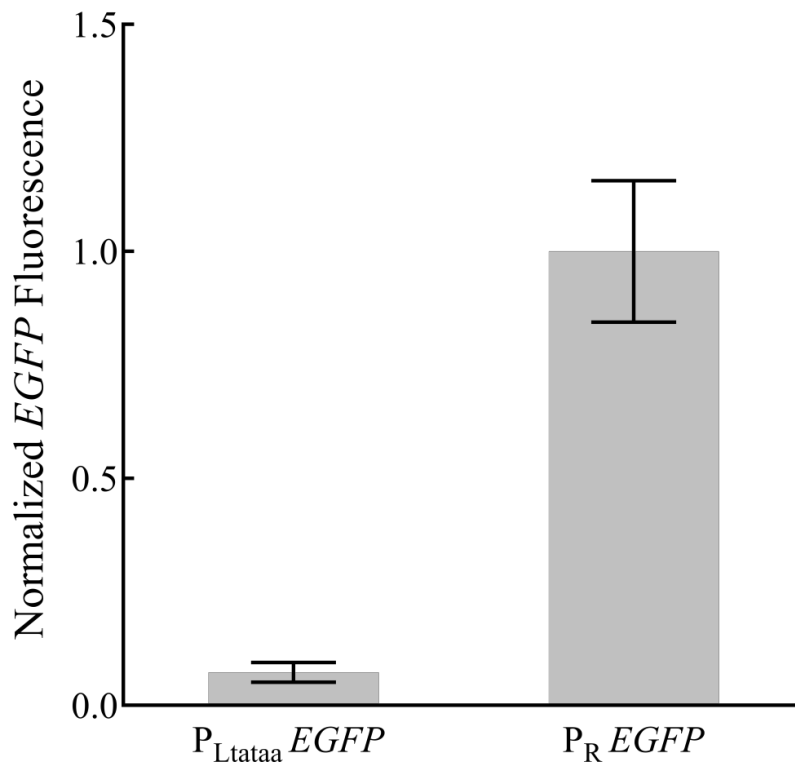


Figure 4.9 Comparison of EGFP expression from P_{Ltataa} and P_R promoters with identical RBS and identical plasmid vector.

Chapter four

Where each bar represents mean normalized EGFP fluorescence. Error bar represents standard deviation from the mean. At least three biological replicates were used.

The Network brick pipeline, in this study, was developed to meet such challenges in the fabrication line by allowing easy optimization. In this direction, we designed 7 new RBSs and estimated their relative strength using thermodynamic calculation (RBS calculator) (Salis, Mirsky and Voigt, 2009; Espah Borujeni, Anirudh S Channarasappa and Salis, 2014). Applying Network brick, we replaced the RBS between P_R promoter and EGFP gene of the cassette 11 (Table 4.2), a constituent cassette of the IMPLY gate, with seven different RBSs of varying strength (Table 4.3) in parallel. The resulting cassettes were named as cassette 12i1 to 12i7. IMPLY gate circuits containing cassette 13i2 and 13i4, which contains RBS-i2 and i4 respectively showed matching behavior of an IMPLY gate (Figure 4.8 D & Figure 4.10) truth table.

Table 4.3. RBS used in the study and their estimated translation rates from RBS calculator

RBS Number	RBS Sequences	Predicted translation Initiation	la-
		Rate (au)	
RBS	ATTAAAGAGGAGAAA	5007.43	la-
RBS i1	TGATAGGCGGCACAG	294.67	
RBS i2	ACAGCAGACGAGAAA	434.4	
RBS i3	AACTTCGTAGACAGG	395.68	
RBS i4	AAGCATATCGTAGCG	571.6	
RBS i5	GGCAGGGAAGGACCA	842.76	
RBS i6	AGGCTGTAAGTAAGG	1069	
RBS i7	TGGCATAGGCGGTAC	1259.5	

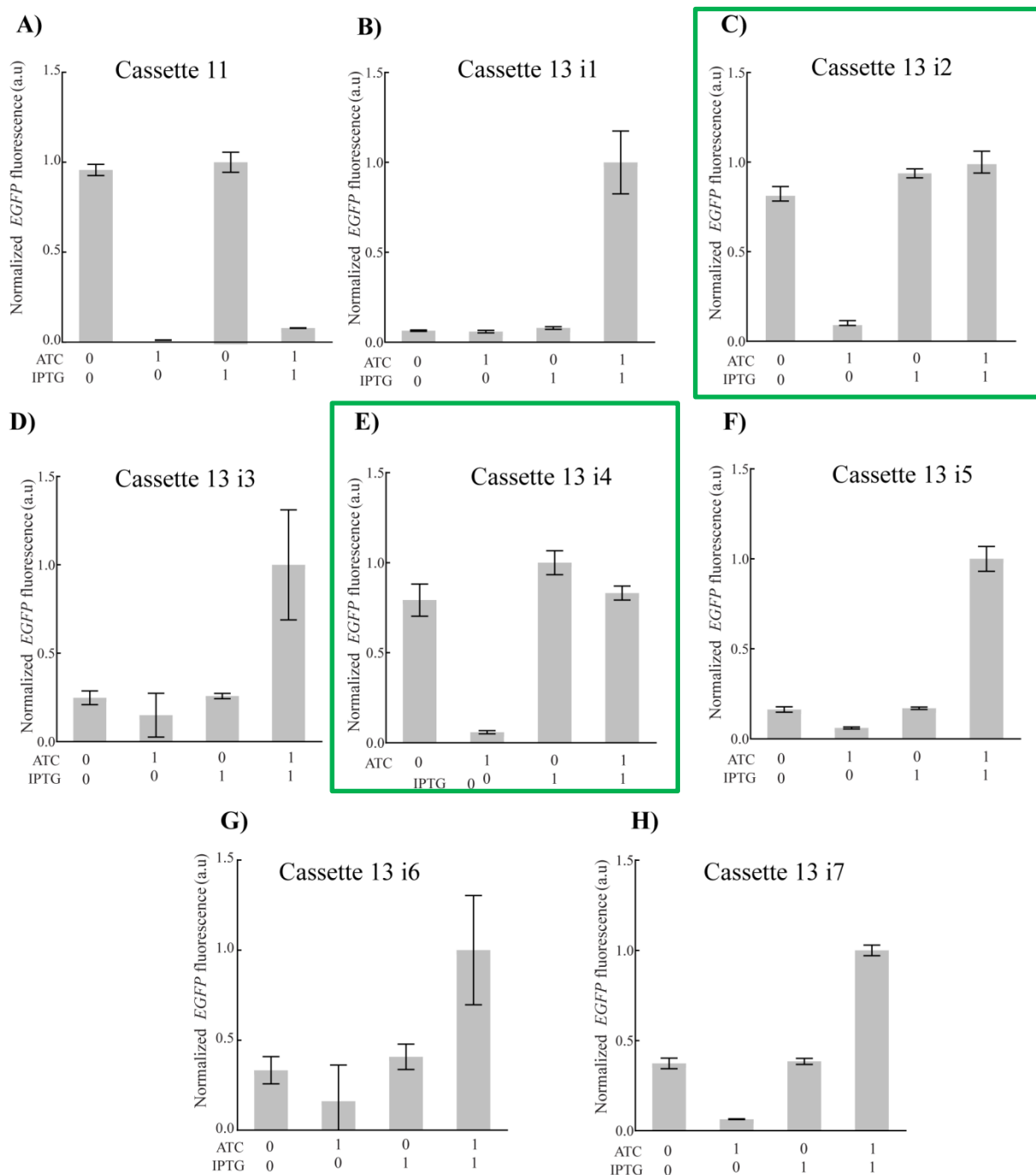


Figure 4.10 IMPLY gate logic behavior with various RBSs (i1-i7).

To optimize the IMPLY gate logic behavior, we replaced the native RBS in IMPLY gate with seven new engineered RBS. The desired behaviour obtained with RBS i2 and i4 (boxed). Where each bar represents mean normalized EGFP fluorescence at different

combinations of inducers at saturating concentrations. Error bar represents standard deviation from the mean. At least three biological replicates were used.

4.2.5 Mathematical modeling, digitality, simulation and experimental validation of IMPLY device.

A device's transfer function is an important quantitative characteristic that is commonly utilized in engineering disciplines. Similar to an electric circuit, a genetic circuit's transfer feature can be described as the system's input-output relationship at steady state. A genetic device should ideally be mathematically predictable and digital in its response. To demonstrate such device physics in our systems, we developed a simple ordinary differential equation (ODE) based kinetic model which was applied to explain genetic circuits (Roquet and Lu, 2014). The model was based on a phenomenological Hill function. The numerical values of the Hill coefficient “ n_i ” estimate the degree of sensitivity of the “output” as a function of “input.” For “ n_i ” greater than 1, the gene circuit is conventionally considered as digital-like or ultrasensitive (Roquet and Lu, 2014). Here, the EGFP expression (output) as a function of two inputs (IPTG and ATC) of the IMPLY gate was modeled through following equation.

$$\begin{aligned} \frac{d[EGFP]}{dt} = & k_a \left[\left(b_1 + \frac{\left(\frac{[IPTG]}{K_1} \right)^{n_1}}{1 + \left(\frac{[IPTG]}{K_1} \right)^{n_1}} \right) \left(b_2 + \frac{\left(\frac{[ATC]}{K_2} \right)^{n_2}}{1 + \left(\frac{[ATC]}{K_2} \right)^{n_2}} \right) \right. \\ & \left. + \left(b_3 + \frac{1}{1 + \left(\frac{[ATC]}{K_3} \right)^{n_3}} \right) \right] - k_d[EGFP] \end{aligned}$$

Equation 4.1

Chapter four

Where, k_a is the scaling rate constant, which represents maximum additional production rate due to up-regulation; K_i and n_i represent the Hill constant and coefficients (cooperativity); k_d is represent degradation rate of *EGFP* and the b_i is the basal level expression from the respective promoters.

At steady state

$$[EGFP]_{ss} = \left(\frac{k_a}{k_d}\right) \left[\left(b_1 + \frac{\left(\frac{[IPTG]}{K_1}\right)^{n_1}}{1 + \left(\frac{[IPTG]}{K_1}\right)^{n_1}} \right) \left(b_2 + \frac{\left(\frac{[ATC]}{K_2}\right)^{n_2}}{1 + \left(\frac{[ATC]}{K_2}\right)^{n_2}} \right) + \left(b_3 + \frac{1}{1 + \left(\frac{[ATC]}{K_3}\right)^{n_3}} \right) \right]$$

Equation 4.2

Or,

$$[EGFP]_{ss} = \left(\frac{k_a}{k_d}\right) \left(b_1 + \frac{\left(\frac{[IPTG]}{K_1}\right)^{n_1}}{1 + \left(\frac{[IPTG]}{K_1}\right)^{n_1}} \right) \left(b_2 + \frac{\left(\frac{[ATC]}{K_2}\right)^{n_2}}{1 + \left(\frac{[ATC]}{K_2}\right)^{n_2}} \right) + \left(\frac{k_a}{k_d}\right) \left(b_3 + \frac{1}{1 + \left(\frac{[ATC]}{K_3}\right)^{n_3}} \right)$$

Equation 4.3

Or,

$$[EGFP]_{ss} = c \left(b_1 + \frac{\left(\frac{[IPTG]}{K_1}\right)^{n_1}}{1 + \left(\frac{[IPTG]}{K_1}\right)^{n_1}} \right) \left(b_2 + \frac{\left(\frac{[ATC]}{K_2}\right)^{n_2}}{1 + \left(\frac{[ATC]}{K_2}\right)^{n_2}} \right) + c \left(b_3 + \frac{1}{1 + \left(\frac{[ATC]}{K_3}\right)^{n_3}} \right)$$

Equation 4.4

Where, $c = \left(\frac{k_a}{k_d}\right)$

The above equation can be divided in two components or modules as follows.

$$[EGFP]_{ss} = [EGFP]_{ss}^{mod1} + [EGFP]_{ss}^{mod2}$$

Equation 4.5

$$[EGFP]_{ss}^{mod1} = c \left(b_1 + \frac{\left(\frac{[IPTG]}{K_1}\right)^{n_1}}{1 + \left(\frac{[IPTG]}{K_1}\right)^{n_1}} \right) \left(b_2 + \frac{\left(\frac{[ATC]}{K_2}\right)^{n_2}}{1 + \left(\frac{[ATC]}{K_2}\right)^{n_2}} \right)$$

Equation 4.6

$$[EGFP]_{ss}^{mod2} = c \left(b_3 + \frac{1}{1 + \left(\frac{[ATC]}{K_3}\right)^{n_3}} \right)$$

Equation 4.7

Equation 4.6 and Equation 4.7 also represent two distinct genetic modules within the IMPLY gate [module 1 (P_{LTATAA}–EGFP in DH5alphaZ1) and module 2 (Pltet_o_1–CI_m,

Chapter four

P_R -EGFP in DH5alphaZ1)]. We are assuming the behavior of each module within the whole IMPLY gate would be practically similar as each module in separation, we performed experimental characterization of individual genetic module in cells. To estimate the parameter values, experimentally we generated dose response curves (Figure 4.11). Now, in case of module I, keeping the ATC at constant saturated concentration, we can rewrite Equation 4.6 as

$$[EGFP]_{ss} = c' \left(b_1 + \frac{\left(\frac{[IPTG]}{K_1} \right)^{n_1}}{1 + \left(\frac{[IPTG]}{K_1} \right)^{n_1}} \right)$$

Equation 4.8

where

$$c' = c \left(b_2 + \frac{\left(\frac{[ATC]}{K_2} \right)^{n_2}}{1 + \left(\frac{[ATC]}{K_2} \right)^{n_2}} \right)$$

Equation 4.9

Similarly, at constant saturated concentration of IPTG, Equation 4.6 becomes

$$[EGFP]_{ss} = c'' \left(b_2 + \frac{\left(\frac{[ATC]}{K_2} \right)^{n_2}}{1 + \left(\frac{[ATC]}{K_2} \right)^{n_2}} \right)$$

Equation 4.10

where

$$c_1'' = c \left(b_1 + \frac{\left(\frac{[IPTG]}{K_1} \right)^{n_1}}{1 + \left(\frac{[IPTG]}{K_1} \right)^{n_1}} \right)$$

Equation 4.11

Chapter four

Thus Equation 4.8 and Equation 4.10 allowed performing experiments for module 1 (Figure 4.11 A and B), where the EGFP expression was measured as a function of IPTG and ATC concentration respectively. When IPTG dose response was performed, ATC was kept constant its saturated concentration. Similarly, during ATC dose response curve, IPTG was kept constant at its saturated concentration. The fitting was done using in-built data fitting algorithm in OriginPro using custom equations. The extracted parameter values are tabulated in Table 4.4. Whereas, Equation 4.7 allowed performing experiments for module 2 (Figure 4.11 C).

At highest concentration of both IPTG and ATC, taking the normalized *EGFP* fluorescence (at steady state) is 1 a.u, putting the value and arranging in Equation 4.8, Equation 4.7 and Equation 4.11. we obtained the value of $c = 0.86$ (for module I) and 0.93774 (for module II). We took the average value of $c = 0.89$

Simulation of IMPLY gate behavior: Assuming the hierarchical principle of electronics held well in in-cell systems chemistry, we created reduced forms of the equations (Equation 4.4) based on an individual set of molecular modules i.e., module I and module II, fitted with experimental data (Figure 4.11) to estimate the parameters (Table 4.4). We plugged the parameter values obtained from individual experiment of distinct modules in the Equation 4.4 and performed a complete simulation by varying the both IPTG and ATC concentration for total 4902 data points. We experimentally profiled the three-dimensional behavior of IMPLY gate across different concentrations of IPTG and ATC (132 data points). The simulated input-output behavior for the IMPLY gate in Figure 4.12 (A) showed a close topological match with the experimental profile Figure 4.12 (B), further validating the input-output behavior of the IMPLY gate.

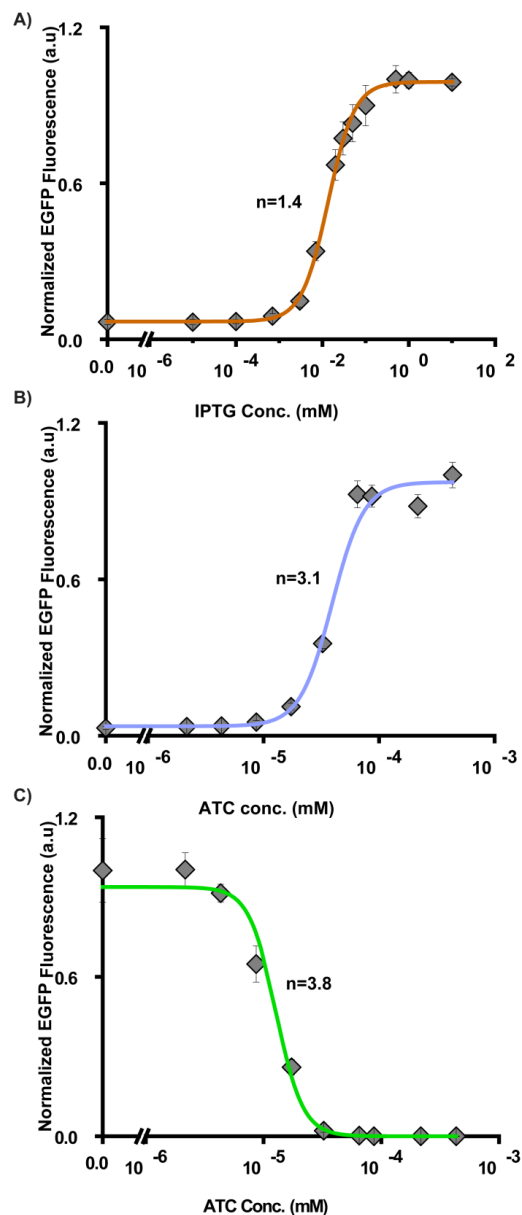


Figure 4.11. Dose-response curves

(A and B) for module 1 (PLTATAA –EGFP in DH5αZ1), where for **A)** ATC concentration was kept constant at its saturated concentration 200ng/mL and for **B)** IPTG concentration was kept constant at its saturated concentration, 10mM. **C)** Dose-response curve of module 2 (Pltet0_1-Clm, PR-EGFP in DH5αZ1). The ‘n’ signifies the Hill coefficient for corresponding dose-response curves.

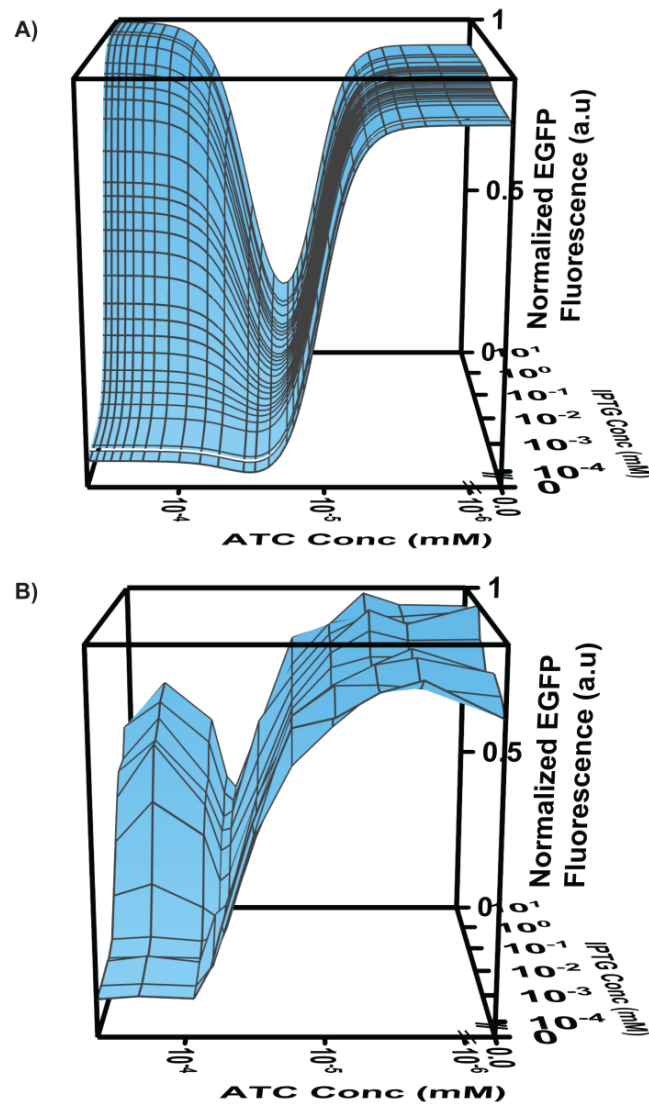


Figure 4.12. Simulated (A) and experimental (132 data points) (B) behavior of IMPLY gate. Plugging the parameter values in the Equation 4.4 we performed a complete simulation by varying the both IPTG and ATC concentration for total 4902 data points. We compared it with experimental behavior of IMPLY gate across different concentrations of IPTG and ATC (132 data points).

4.2.6 Fabrication and optimization of a 2-input-2output integrated circuit that integrates IMPLY and a NOT gate

To test, if the chemical signal processing by IMPLY gate could work as a part of a larger signal processing device within a single cell and to create a device out of IMPLY

Chapter four

with higher signal differentiation capability, we created a 2-input-2-output integrated circuit by connecting a NOT gate with the input line of the IMPLY gate (Figure 4.13 A). Whereas IMPLY is a powerful constraint-based logic, in terms of decoding, it can only differentiate between 1,0 and 00/01/11 input states by showing 0 and 1 logic output respectively. The NOT gate sensed environmental ATC and responded inversely by producing EGFP. We constructed this NOT gate by placing a TdTomato gene (produce an orange fluorescent protein) under PR promoter, which can be repressed by CIm protein coming from an ATC/TetR inducible PltetO-1 promoter. This NOT gate was integrated with the IMPLY gate (Figure 4.13 A, B). The integrated circuit consisted of five genetic cassettes (Table 4.2) in a single plasmid and total of seven genes were required for the functionality of the circuit. However, the experimental behavior of the first construct did not match with the truth table (Figure 4.13 C).

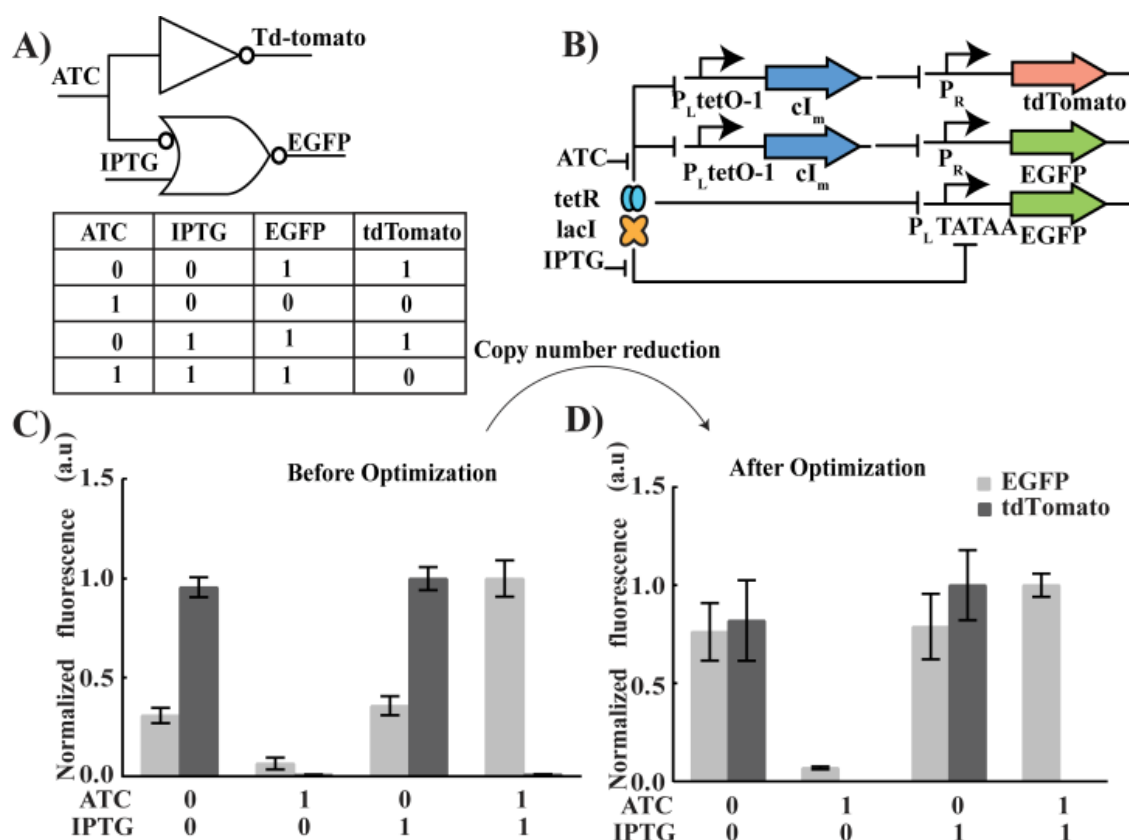


Figure 4.13. Design and testing of 2-input-2-output integrated logic circuit.

Chapter four

A) Electronic equivalent representation and truth table. **B)** Biomolecular design. It consists of two distinct module, one NOT gate where the mutated lambda repressor (cIm) is kept under ATC inducible $P_{LTetO-1}$ promoter and the reporter gene Td-tomato is kept under lambda repressible P_R promoter. The second module is IMPLY gate shown earlier. The tet repressor (tetR) and lac inhibitor (lacI) is constitutively expressed in DH5 α 1 *E.coli* cell strain. **C)** Experimental behaviour before optimization, Where each light colored bar represents mean normalized EGFP fluorescence and dark colored bar represents mean normalized tdTomato fluorescence, both at different combinations of inducers at saturating concentrations. Error bar represents standard deviation from the mean. At least three biological replicates were used. **D)** Experimental behaviour after optimization (copy number).

It was reasoned that as the molecular circuits inside cells did not have any insulation, the leakage of CIm (basal level expression) from the NOT gate $P_{LTetO-1}$ promoter added up with the leakage from $P_{LTetO-1}$ promoter of module 2 of IMPLY gate and increased the cellular availability of CIm protein, which resulted in the partial repression of the EGFP expression from the P_R promoter. Therefore, a difference in output EGFP expression levels between synthetic P_{LTATAA} and P_R promoter was observed. To reduce the CIm concentration, the number of available plasmids in the cell was reduced by changing the origin of replication from ColE1 (high copy) to p15A (low copy) (Lutz and Bujard, 1997a). The copy number was reduced so that the leakage of CIm can be minimized from $P_{LTetO-1}$ promoter, as it was assumed that the available pool of Tet repressor is constant in cell and it is insufficient for maintaining a tight regulation of $P_{LTetO-1}$ promoter. This also showed that origin of replication can be easily changed in Network brick. The optimized circuit showed matching behavior with the truth table (Figure 4.13 D). The 2-input-2-output integrated circuit sensed both the extracellular chemicals IPTG and ATC however it was able to differentiate a higher number of input states combination (11, 10 and 00/01) than that of IMPLY gate alone.

4.2.7 Mathematical modeling, digitality, simulation and experimental validation of 2input-2output-integrated device.

We also developed a mathematical model for this integrated logic circuit. The integrated circuit consists of IMPLY gate expressing EGFP and a NOT gate expressing *tdTomato*.

The overall behavior of the integrated circuit can be explained by two ODEs Equation 4.12 and Equation 4.13.

$$\frac{d[EGFP]}{dt} = k_a' \left[\left(b_1' + \frac{\left(\frac{[IPTG]}{K_1'} \right)^{n_1'}}{1 + \left(\frac{[IPTG]}{K_1'} \right)^{n_1'}} \right) \left(b_2' + \frac{\left(\frac{[ATC]}{K_2'} \right)^{n_2'}}{1 + \left(\frac{[ATC]}{K_2'} \right)^{n_2'}} \right) + \left(b_3' + \frac{1}{1 + \left(\frac{[ATC]}{K_3'} \right)^{n_3'}} \right) \right] -$$

$$k_d'[EGFP]$$

Equation 4.12

$$\frac{d[tdTomato]}{dt} = k_a'' \left(b_4 + \frac{1}{1 + \left(\frac{[ATC]}{K_4} \right)^{n_4}} \right) - k_d''[tdTomato]$$

Equation 4.13

Where, k_a' and k_a'' is the scaling rate constant, which represents maximum additional production rate of EGFP and tdTomato respectively due to up-regulation; K_i and n_i represent the Hill constant and coefficients(cooperativity); k_d' and k_d'' represent degradation rate of EGFP and tdTomato. The b_i is the scaling factor.

At steady state, Equation 4.12 and Equation 4.13 can be expressed as Equation 4.14 and Equation 4.15 respectively

$$[EGFP]_{ss} = c''' \left(b_1' + \frac{\left(\frac{[IPTG]}{K_1'} \right)^{n_1'}}{1 + \left(\frac{[IPTG]}{K_1'} \right)^{n_1'}} \right) \left(b_2' + \frac{\left(\frac{[ATC]}{K_2'} \right)^{n_2'}}{1 + \left(\frac{[ATC]}{K_2'} \right)^{n_2'}} \right) + c''' \left(b_3' + \frac{1}{1 + \left(\frac{[ATC]}{K_3'} \right)^{n_3'}} \right)$$

Equation 4.14

Where, $c''' = \left(\frac{k_a'}{k_d'} \right)$

$$[tdTomato]_{ss}^{mod3} = c'''' \left(b_4 + \frac{1}{1 + \left(\frac{[ATC]}{K_4} \right)^{n_4}} \right)$$

Equation 4.15

Where, $c'''' = \left(\frac{k_a''}{k_d''} \right)$

We again assumed that the various modules and IMPLY gate would work similar way as it works individually and the behaviors of the modules, in terms of relative fold change in expression and nature of its dose response curve were conserved at different copy number plasmids. We further assume that both EGFP and tdtomato worked similarly when they were under same promoter. Therefore, $n_1' = n_1, K_1' = K_1, b_1' = b_1, c''' = c, n_2' = n_2, K_2' = K_2, b_2' = b_2, n_3' = n_3, K_3' = K_3, b_3' = b_3, c'''' = c, b_4 = b_3, n_4 = n_3, K_4 = K_3$. Therefore, Equation 4.14 and Equation 4.15 can be modeled as Equation 4.4 and Equation 4.7 respectively. The quantitative characterization, fitting and parameter estimation (Table 4.4) and simulation were done similar as in IMPLY

gate. The numerical values of the Hill coefficient ' n_i ' estimate the degree of sensitivity of the 'output' as a function of 'input'. For ' n_i ' greater than 1, the gene circuit is conventionally considered as digital like or ultrasensitive (Roquet and Lu, 2014).

Table 4.4. Parameter values generated by curve fitting.

Parameter	Estimated value	unit
b_1	0.07388	
b_2	0.0388	
b_3	-8.79×10^5	
c	0.89	
K_1	0.01358	mM
K_2	3.913×10^5	mM
K_3	1.274×10^5	mM
n_1	1.417	
n_2	3.123	
n_3	3.806	

Assuming the hierarchical principle of electronics held well in in-cell systems chemistry, we created reduced forms of the equations based on an individual set of molecular modules, fitted with experimental data to estimate the parameters (Table 4.4). We further simulated the input-output behavior for this integrated circuit and Figure 4.14 showed a close topological match between simulation and experiments further validating the input-output behavior of the integrated circuit. We believed that the little discrepancies observed between the model and the experiment arose from the assumption of unaltered behavior of molecular modules within an integrated circuit. The hierarchy

Chapter four

principle in in-cell molecular computation is challenging due to the crosstalk of DNA modules and background cellular context (Falk et al., 2019). Both the IMPLY and integrated circuit were designed and fabricated assembling module by module. Similarly, the mathematical model was constructed by summing up the equations represent the modules. This suggested the successful application of the hierarchy principle of electronics in terms of abstraction, decoupling, and standardization at the level of in-cell systems chemistry (Andrianantoandro, Basu, David K Karig, *et al.*, 2006).

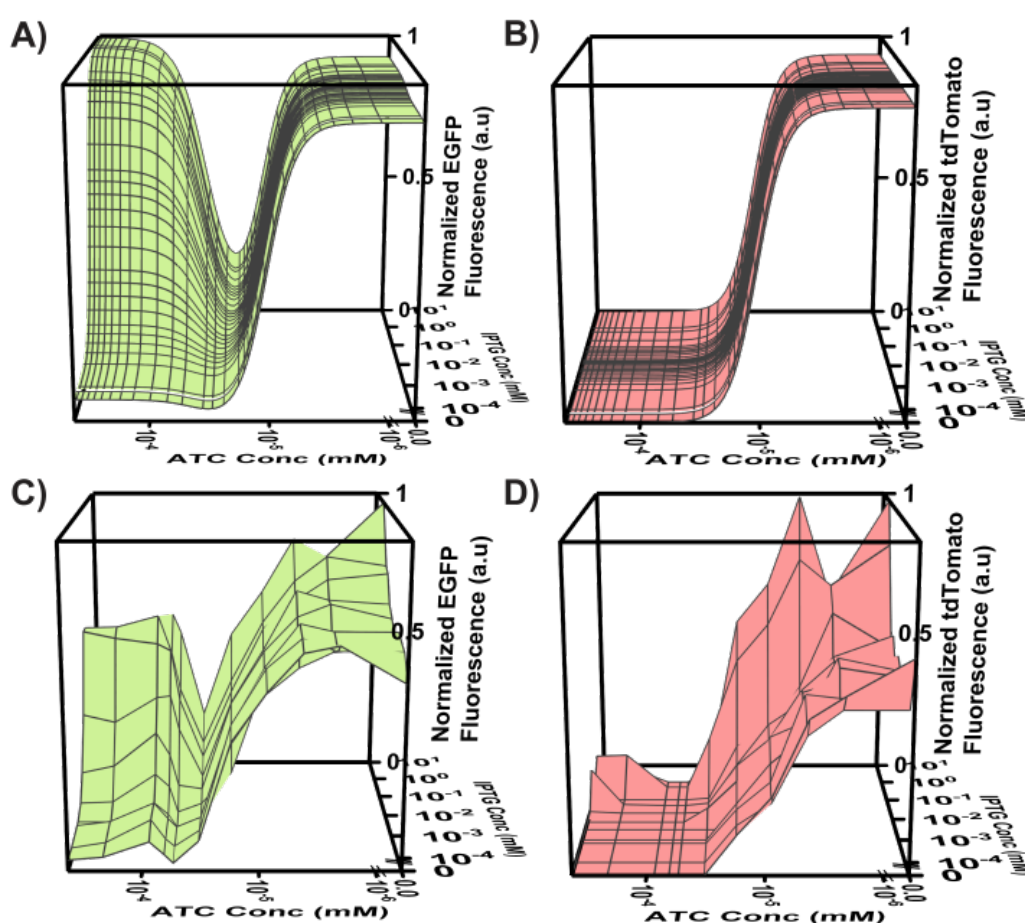
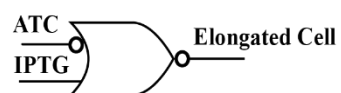


Figure 4.14. Simulated and experimental behaviour of 2-input-2-output integrated logic circuit. Simulated behavior of A) output 1 (EGFP) and B) output 2 (TdTomato). Experimental behaviour (121 data points each) of C) output 1 (EGFP) and D) output 2 (TdTomato) as a function of 2 input chemicals in the full integrated circuit inside cell.

4.2.8 Integrating synthetic circuits with native cellular process of *E.coli*.

Next, we wanted to show a simple application of this circuit using some output other than fluorescence, where we can demonstrate the output of the device using a cellular phenotype. Therefore, we integrated this IMPLY (Figure 4.15 and Figure A.1) and 2-input-2-output circuit (Figure 4.16 and Figure A.2) with inherent cellular process. *ftsZ* is a native protein, which helps in cell division in *E. coli* (Sánchez-Gorostiaga *et al.*, 2016). Deregulation of *ftsZ* results a non-divided, elongated shaped *E. coli*. This model system has been used in multiple studies as an output (Mückl *et al.*, 2018; Wang *et al.*, 2021). Here we used a synthetic small regulatory RNA (synRNA) against the *ftsZ* mRNA (Mukhopadhyay and Bagh, 2020). SynRNA can be designed to control the expression of target protein by binding with the target mRNA to hinder its translation (Na *et al.*, 2013). We replaced the EGFP with anti- *ftsZ* synRNA of the IMPLY gate, such that the new IMPLY gate had a phenotypic output (elongated cell shape) Figure 4.15. Using this approach shows the output phenotypically other than that of fluorescence and further validates the utility of this genetic circuit. Cell length is controlled by various factors and is very complex, using *ftsZ*. Experimental characterization of the engineered cell showed a matching behaviour with the expected IMPLIES truth table (Figure 4.15 B). In the second construct, we replaced the tdTomato gene of the 2-input-2-output circuit with the anti- *ftsZ* synRNA to create a new circuit, which had two phenotypic outputs, namely ‘elongated shape’ (resulted from anti- *ftsZ* synRNA) and ‘green colour’ (resulted from EGFP expression) (Figure 4.16 B). Experimental characterization with the engineered cells showed matching behaviour with the corresponding truth table (Figure 4.16 A).

A)



ATC	IPTG	Elongated Cell
0	0	1
1	0	0
0	1	1
1	1	1

B) ATC IPTG

0 0

1 0

0 1

1 1

Elongated Cell

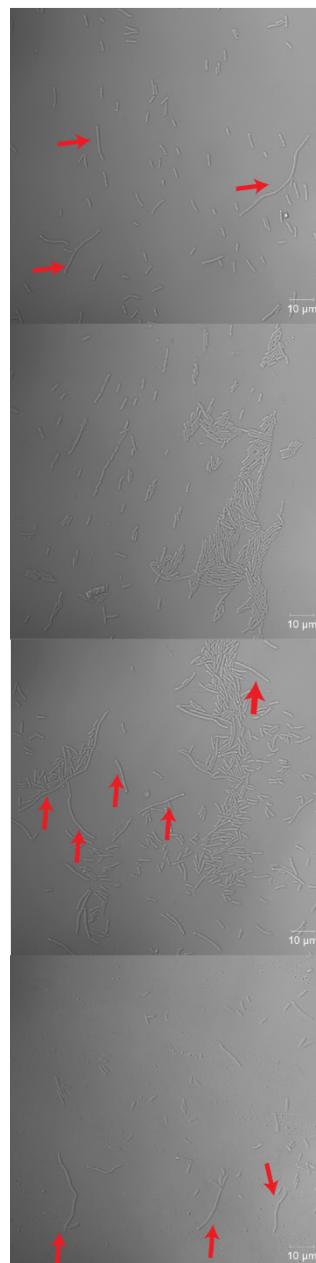


Figure 4.15. A) The IMPLY circuit with phenotypic output. B) Experimental behaviour of the engineered cell. The red arrows are pointing towards elongated cells.

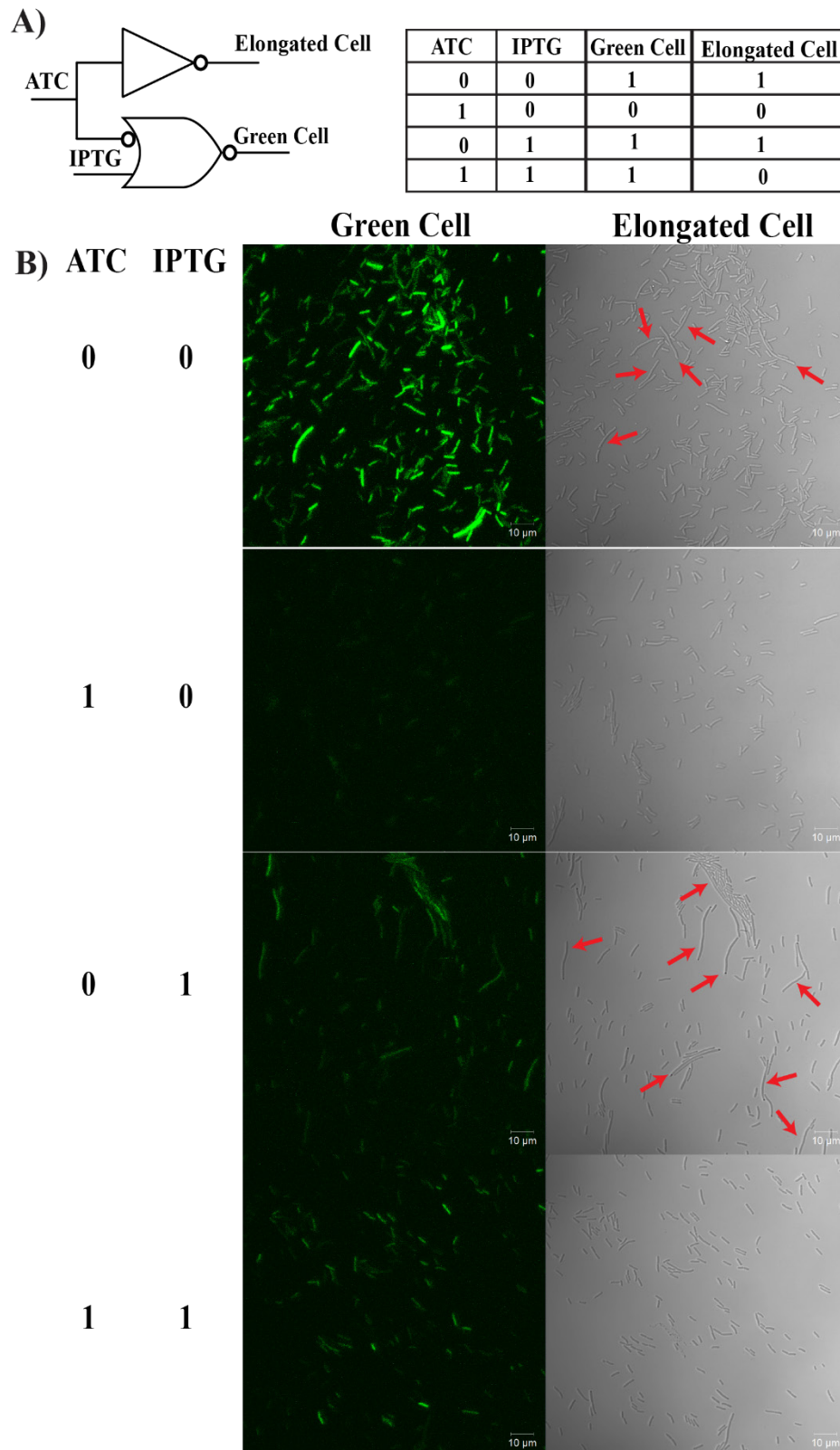


Figure 4.16. A) The 2-input-2-output circuit with phenotypic outputs. B) Experimental behaviour of the engineered cell. The red arrows are pointing towards elongated cells.

4.2.9 PCR based Network Brick for fabrication and characterization of Exchanger circuit

Assembly techniques like Gibson assembly can be automated as there is no gel purification step. Therefore, we developed a modified PCR based Network Brick, which did not require any gel separation step and successfully demonstrated construction and characterization of a genetic exchanger gate. As the whole process can be done without any gel extraction system, it may be automated in a liquid handling system. Network Brick system and its PCR version will be useful for systematic construction and optimization of bioparts library, large synthetic gene circuits and any genetic constructs with multiple cassettes in a plasmid. The PCR based version consists of a Library vector and FCEV. The library vector can be used for construction of libraries of individual bioparts. The library vector serves dual purpose i) storage of biopart ii) ease of amplification of biopart Figure 4.17. The bioparts can later on be amplified by using the appropriate primer pair. Each kind of biopart has a different set of primer as shown in Table 4.5. For inserting a biopart, the Blank FCEV can be digested and PCR purified as it will generate DNA fragment of less than 20bp. We then used the PCR based Network brick platform for fabrication and assembly of Exchanger circuit (cassette 21) as shown in Table 4.2 according to the design shown in Figure 4.18 B. We further characterized the circuit in *E.coli* (Figure 4.18 D) and the behavior was in accordance to the truth table (Figure 4.18 C)

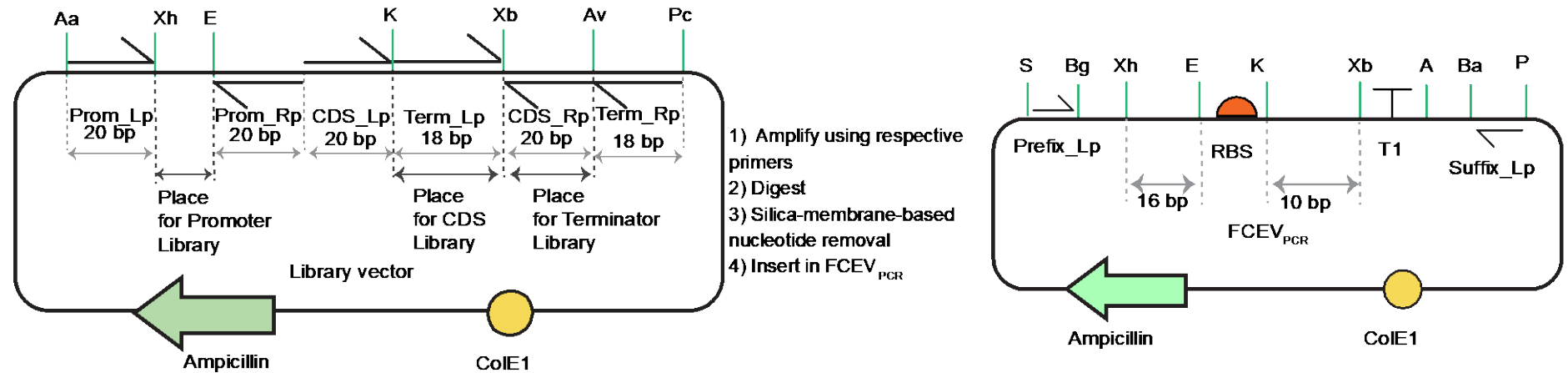


Figure 4.17 Modified PCR based Network Brick vector and assembly pipeline.

The library vector provides Multiple cloning sites for promoter, coding sequence (CDS) and terminator flanked by unique restriction sites and primer binding site for amplifying segment of interest using unique primers for promoter, coding sequence (CDS) and terminator thus avoiding gel purification of bioparts. FCEV_{PCR} allows assembly of gene circuits, the assembly MCS provided for promoter, coding sequence (CDS) and terminator generates very short fragments on restriction digestion of FCEV_{PCR}, thus it can be purified using PCR purification kit.

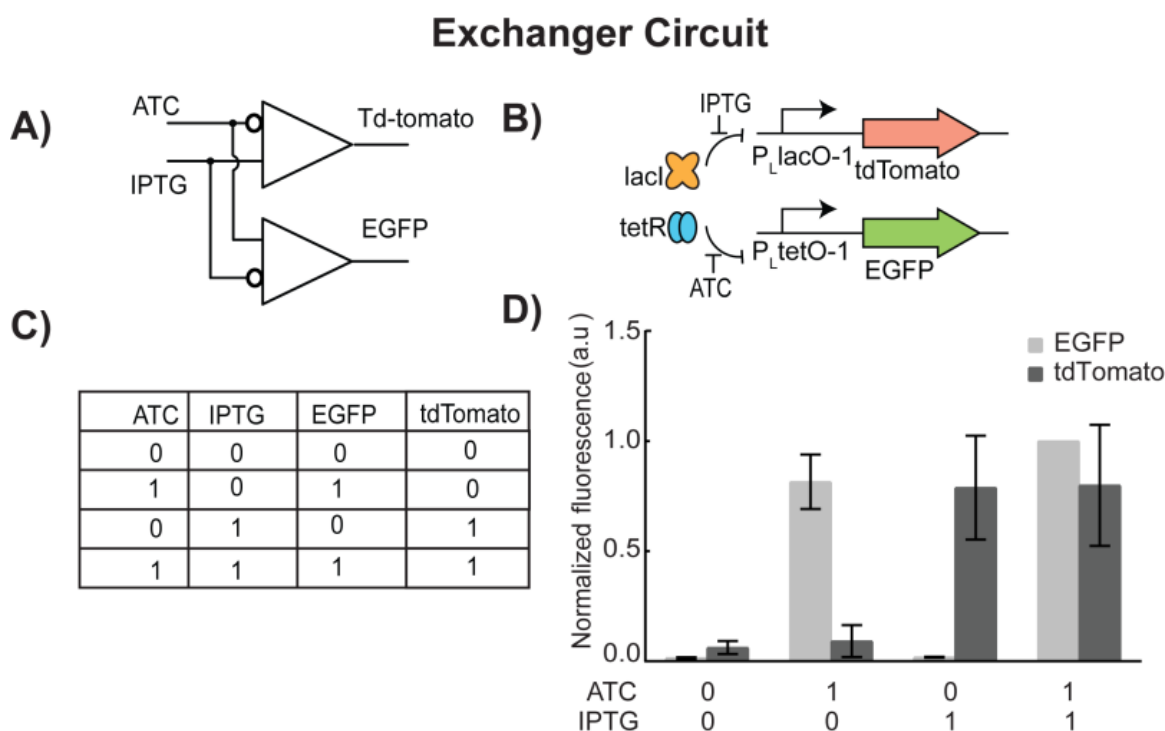


Figure 4.18 Design and experimental characterization of Exchanger circuit constructed using PCR based Network Brick.

A) Electronic equivalent representation **B)** Biomolecular design. Td tomato is kept under IPTG inducible P_{LacO-1} promoter and EGFP is kept under lambda repressible PR promoter. The tet repressor (tetR) and lac inhibitor (lacI) is constitutively expressed in DH5 α 1 *E.coli* cell strain. **C)** truth table **D)** Experimental behaviour, Where each light colored bar represents mean normalized EGFP fluorescence and dark colored bar represents mean normalized tdTomato fluorescence, both at different combinations of inducers at saturating concentrations. Error bar represents standard deviation from the mean. At least three biological replicates were used.

Table 4.5. Predefined set of primers required for Network PCR Brick

Sr No.	Primer Name	Primer Sequence (5'-->3')	Description
1	Prom_Lp	TTCCGTATTTGGCTCCTTTG	Primer pair for amplifying Promoter from library
2	Prom_Rp	CCCGCTCTGTGAGTTTCTTC	
3	CDS_Lp	GCGAATCTCCGACTATGACC	Primer pair for amplifying CDS from library
4	CDS_Rp	CCGACCCGTATGTATGAACC	

5	Term_Lp	CGCCAATCCAGCAATACC	Primer pair for amplifying Terminator from library
6	Term_Rp	CGCATCATCGTCAGCAAG	
7	Prefix_Lp	TACTGGGACGAAGACGAACA	Primer pair for amplifying entire cassette from Network brick plasmid and sequencing.
8	Suffix_Rp	GTTGTTTTGGAGCACGGAAC	

4.3 Discussion and Conclusion

In conclusion, a new plasmid vector system named as Network Brick, in which any biopart can be replaced easily within a single genetic cassette independent of its position in the cassette and multiple cassettes can be assembled in a single plasmid at any desired direction was created. Using this method, we demonstrated the construction and optimization of fully and chemically controllable in-cell molecular genetic NOR and IMPLY gate in a single *E. coli*, which can sense, integrate and process 2 extracellular environmental chemical signals.

Completion of the full spectrum IMPLY gate adds constraint-based fundamental logic in the in-cell molecular computation and completes the toolbox. Our design is generic and can be used to create other IMPLY gates with other equivalent inducers and repressor interactions. The successful function of IMPLY gate within a 2-input- 2-output integrated circuit, suggests its ability as a part of a larger circuit. The 2-input-2-output integrated circuit, due to its higher differentiation capability between combinations of input signals, has significance as a cell-based chemical signal-decoding device. As the output genes of IMPLY and the 2-input-2-output circuits can be changed to any genes, this could be useful to control cellular or metabolic pathways in *E.coli* from outside, in a specific logical manner. In the 2-input-2-output circuit, which was made from

Chapter four

the IMPLY gate, it is possible to reverse the expression of the tdTomato gene more than once by ‘adding’ inducers sequentially to the system (Figure 4.13 D). This property could be useful to control the ON and OFF of a gene multiple times just by adding the inducers in sequence, instead of flushing out the inducers from the system. This property can be very useful to control a cellular or metabolic process in a running reactor, where it is not convenient to flush out any inducer from the media during the run to reverse the expression of a gene. Due to the significance of constraint-based logic in robotics and control, the constraint-based genetic IMPLY may have significance in the future development of microbiorobotics and chemical control. Further, as the IMPLY gate and the combinatorial circuit made from it were implemented in a single cell, in fact in a single plasmid, the design may be applied to control intracellular process in a specific logical manner by changing its output proteins as shown in Figure 4.15 and Figure 4.16. Such intracellular IF-THEN logic may be applied to control other metabolic and cellular processes in future, which cannot be achieved by multi-cellular implementation (Tamsir, Tabor and Voigt, 2011). The function of those circuits is digital like and mathematically predictive. The modules and its integration in both experiments and models suggested the successful application of the electronic hierarchy principle in in-cell molecular circuitry. This may serve as a general design criterion in creating future complex whole-cell smart biosensing and signal processing device. We further developed a modified PCR based Network Brick, which did not require any gel separation step and successfully demonstrated construction and characterization of a genetic exchanger gate. As the whole process can be done without any gel extraction system, it may be automated in a liquid handling system. Network Brick system and its PCR version will be useful for systematic construction and optimization of bioparts library, large

synthetic gene circuits and any genetic constructs with multiple cassettes in a plasmid. Taken together, the work has implications in in-cell complex molecular computation, smart whole-cell biosensors, signal decoding device, micro-biorobotics, and space synthetic biology.

4.4 Materials and Methods

4.4.1 Promoters, primers, genes, plasmids, and bacterial cell strains

The full lists of plasmids, primers, and RBSs are shown in Table 4.6 Table 4.7 and Table 4.3 respectively. All empty vectors were made by incorporating double-stranded DNA fragments according to our designed framework with appropriate restriction sites in two backbone plasmids pOR-EGFP-12 and pOR-Luc-31 (gift from Prof. David McMillen, University of Toronto, Toronto, Canada). The DNA fragments were synthesized from Invitrogen. The required genes and promoters are PCR'd with appropriate primers and incorporated accordingly in the empty plasmids. The Network Brick pipeline was followed to make the logic circuits as shown in Table 4.2. PCRs were performed using KOD hot-start DNA polymerase (Merck Millipore) and primers were synthesized from Integrated DNA Technologies, Singapore. The bioparts were assembled into the Network brick plasmid vectors using sticky end ligation as generated according to the Network brick principle Figure 4.4. Plasmids were transformed in chemically competent DH5 α Z1 *E. coli* strains. The genotype is laciq, SpR, deoR, supE44, Delta(lacZYA-argFV169), Phi80 lacZDeltaM15, hsdR17(rK- mK+), recA1, endA1, gyrA96, thi-1, relA1 (Lutz and Bujard, 1997a). Sequencing of the plasmids was done from Eurofins Genomics India Pvt Ltd., Bangalore, India. Restriction endonuclease and

Chapter four

T4 DNA ligase were bought from New England BioLabs; plasmid isolation, gel extraction and PCR purification kits were from Qiagen.

Table 4.6 List of plasmids

Sr no	Plasmid Name	Origin of repli- cation	Antibiotic selection	Description	Source
1	pOR- <i>EGFP</i> -12	ColE1	Amp	Source of <i>EGFP</i> , Ampicillin resistance, ColE1 origin	Kind gift from Prof. David McMillen, Uni of Toronto
2	pOR-Luc-31	P15A	Cam	Source of P _{LtetO} -1 and Chloramphenicol resistance, P15A origin	Kind gift from Prof. David McMillen, Uni. Of Toronto
3	pZA31 <i>EGFP</i>	p15A	Cam	<i>EGFP</i> under P _{LtetO} -1	This study
4	pZE12 <i>EGFP</i>	ColE1	Amp	<i>EGFP</i> under P _{LlacO} -1	This study
5	ptdTomato	pUC	Amp	Source of <i>td-Tomato</i>	Clontech

Chapter four

6	pNB_A_FCEV	p15A	Amp	Empty Network brick vector for assembly in forward direction	This study
7	pNB_E_FCEV	ColE1	Amp	Empty Network brick vector for assembly in forward direction	This study
8	pNB_U_FCEV	pUC	Amp	Empty Network brick vector for assembly in forward direction	This study
9	pNB_RCEV	p15A,ColE1, pUC	Amp	Empty Network brick vector for assembly in Reverse direction	This study
10	pNB_EA_BP_t1	ColE1	Amp	biopart 1	This study
11	pNB_EA_BP_t1rbs	ColE1	Amp	biopart 2	This study
12	pNB_EA_BP_t1rbs <i>EGFP</i>	ColE1	Amp	biopart 3	This study
13	pNB_EA_3 <i>EGFP</i>	ColE1	Amp	cassette 1	This study
14	pNB_EA_R3 <i>EGFP</i>	ColE1	Amp	cassette 8	This study

Chapter four

15	pNB_EA_R3EGFPi1	ColE1	Amp	cassette 8i1	This study
16	pNB_EA_R3EGFPi2	ColE1	Amp	cassette 8i2	This study
17	pNB_EA_R3EGFPi3	ColE1	Amp	cassette 8i3	This study
18	pNB_EA_R3EGFPi4	ColE1	Amp	cassette 8i4	This study
19	pNB_EA_R3EGFPi5	ColE1	Amp	cassette 8i5	This study
20	pNB_EA_R3EGFPi6	ColE1	Amp	cassette 8i6	This study
21	pNB_EA_R3EGFPi7	ColE1	Amp	cassette 8i7	This study
22	pRA1SEGFP <i>TCIm</i>	pUC	Amp	Source of <i>CI_m</i>	(Sarkar <i>et al.</i> , 2019)
23	pDA1EGFP	pUC	Amp	Source of P _{LTATAA}	(Sarkar <i>et al.</i> , 2019)
24	pNB_EA_3 <i>Cim</i>	ColE1	Amp	cassette 2	This study
25	pNB_EA_2 <i>Cim</i>	ColE1	Amp	cassette 3	This study
26	pNB_EA_1 <i>Cim</i>	ColE1	Amp	cassette 4	This study

Chapter four

27	pNB_EA_R1 <i>Cim</i>	ColE1	Amp	cassette 5	This study
28	pNB_EA_4 <i>EGFP</i>	ColE1	Amp	cassette 9	This study
29	pNB_EA_R1 <i>Cim_4EG</i> <i>FP</i>	ColE1	Amp	cassette 10	This study
30	pNB_EA_R1 <i>Cim_</i> <i>EA_2Cim_4EGFP</i>	ColE1	Amp	cassette 7, NOR gate	This study
31	pNB_EA_R1 <i>Cim_4EG</i> <i>FP_R3EGFP</i>	ColE1	Amp	IMPLY gate (Poor), cassette 11	This study
32	pNB_EA_R1 <i>Cim_4EG</i> <i>FP_R3EGFPi1</i>	ColE1	Amp	Optimized IMPLY gate (poor), cassette 13i1	This study
33	pNB_EA_R1 <i>Cim_4EG</i> <i>FP_R3EGFPi2</i>	ColE1	Amp	Optimized IMPLY gate (best), cassette 13i2	This study
34	pNB_EA_R1 <i>Cim_4EG</i> <i>FP_R3EGFPi3</i>	ColE1	Amp	Optimized IMPLY gate (poor), cassette 13i3	This study

Chapter four

35	pNB_EA_R1 <i>CI</i> m_4 <i>EG</i> <i>FP_R3EGFPi4</i>	ColE1	Amp	Optimized IMPLY gate (good), cassette 13i4	This study
36	pNB_EA_R1 <i>CI</i> m_4 <i>EG</i> <i>FP_R3EGFPi5</i>	ColE1	Amp	Optimized IMPLY gate (poor), cassette 13i5	This study
37	pNB_EA_R1 <i>CI</i> m_4 <i>EG</i> <i>FP_R3EGFPi6</i>	ColE1	Amp	Optimized IMPLY gate (poor), cassette 13i6	This study
38	pNB_EA_R1 <i>CI</i> m_4 <i>EG</i> <i>FP_R3EGFPi7</i>	ColE1	Amp	Optimized IMPLY gate (poor), cassette 13i7	This study
39	pNB_EA_R3tdToma- toi2	ColE1	Amp	cassette 14	This study
40	pNB_EA_1 <i>CI</i> m_R3tdT omatoi2	ColE1	Amp	synthetic NOT gate, cassette 15	This study

Chapter four

41	pNB_EA_1 <i>CI</i> m_R3tdT oma- toi2_R1 <i>CI</i> m_4 <i>EGFP</i> _R 3 <i>EGFP</i> i2	ColE1	Amp	Two input two output integrated circuit II (poor), cassette 16	This study
42	pNB_AA_1 <i>CI</i> m_R3tdT oma- toi2_R1 <i>CI</i> m_4 <i>EGFP</i> _R 3 <i>EGFP</i> i2	p15a	Amp	Optimized Two input two output circuit II (best), cassette 16	This study
43	Library vector	ColE1	Amp		This study
44	PL_EA_2	ColE1	Amp	Promoter library 1	This study
45	PL_EA_1	ColE1	Amp	Promoter library 1	This study
46	CL_EA_EGFP	ColE1	Amp	CDS library 1	This study
47	CL_EA_td-tomato	ColE1	Amp	CDS library 2	This study

Chapter four

48	pNB_EA_BP_t1rbs <i>stdT</i> <i>omato</i>	ColE1	Amp	Biopart 5	This study
49	pNB_EA_BP_t1rbs <i>EGFP</i> <i>FP</i>	ColE1	Amp	Biopart 6 (similar to biopart 3, but constructed using PCR)	This study
50	pNB_EA_2 <i>td-tomato</i>			Cassette 17	This study
51	pNB_EA_1 <i>EGFP</i>			Cassette 18	This study
52	pNB_EA_R1 <i>EGFP</i>			Cassette 19	This study
53	pNB_EA_R1 <i>EGFP_2td</i> <i>-tomato</i>			Cassette 20, Exchanger circuit	This study
54	pNB_EA_R3 <i>Anti-fts</i> <i>sRNA</i>	ColE1	Amp	Cassette 21	This study
55	pNB_EA_4 <i>Anti-fts</i> <i>sRNA</i>	ColE1	Amp	Cassette 22	This study

Chapter four

56	pNB_EA_R3 <i>Anti-ftsZ</i> <i>sRNA_4Anti-ftsZ sRNA</i>	pUC	Amp	Cassette 23	This study
57	pNB_EA_R3 <i>Anti-ftsZ</i> <i>sRNA</i>	pUC	Cam	n/a	This study

Table 4.7 List of primers

Primer name	Primer Sequence (5'-->3')	Description	Remarks
Primer 1	Caagggcgaggagctgtt	1st round amplification of <i>EGFP</i> (forward)	This study
Primer 2	Ccatgccgagagtgatcc	1st round amplification of <i>EGFP</i> (reverse)	This study
Primer 3	cttcagtcgaggtaccatggtgagcaagggcgag-gagctgtt	2nd round amplification of <i>EGFP</i> flanked by KpnI (forward)	This study
Primer 4	ctgattatgatctagattactt-gtacagctcgtccatgccgagagtgatcc	2nd round amplification of <i>EGFP</i> flanked by XbaI (Reverse)	This study
Primer 5	Ctgattatgactcgagaattgtgagcgggataacaaga-tactgagc	For creating Pllaco-1 promoter flanked by XhoI (forward)	This study
Primer 6	Ctccagtcgtgaattcggtcag-tgcgtcctgctgatgtgctcagtatcttgttatccgctc	For creating Pllaco-1 promoter flanked by EcoRI (Reverse)	This study

Chapter four

Primer 7	Cctgtgccgcctatcag	Annealing oligos for RBS i1 flanked by EcoRI and KpnI (antisense)	This study
Primer 8	Aattctgatagcgccacagggtac	Annealing oligos for RBS i1 flanked by EcoRI and KpnI (sense)	This study
Primer 9	Ctttctcgtctgctgtg	Annealing oligos for RBS i2 flanked by EcoRI and KpnI (antisense)	This study
Primer 10	Aattcacagcagacgagaaagggtac	Annealing oligos for RBS i2 flanked by EcoRI and KpnI (sense)	This study
Primer 11	Aattcaacttcgtagacaggggtac	Annealing oligos for RBS i3 flanked by EcoRI and KpnI(antisense)	This study
Primer 12	Ccctgtctacgaagttg	Annealing oligos for RBS i3 flanked by EcoRI and KpnI(sense)	This study
Primer 13	Ccgctacgatatgcttg	Annealing oligos for RBS i4 flanked by EcoRI and KpnI (antisense)	This study
Primer 14	Aattcaagcatatcgtagcgggtac	Annealing oligos for RBS i4 flanked by EcoRI and KpnI (sense)	This study
Primer 15	Ctggtccttcctgccg	Annealing oligos for RBS i5 flanked by EcoRI and KpnI (antisense)	This study
Primer 16	aattcggcaggggaaggaccaggtac	Annealing oligos for RBS i5 flanked by EcoRI and KpnI (sense)	This study
Primer 17	cccttagttacagcctg	Annealing oligos for RBS i6 flanked by EcoRI and KpnI (antisense)	This study
Primer 18	aattcaggtgtaactaaggggtac	Annealing oligos for RBS i6 flanked by EcoRI and KpnI (sense)	This study
Primer 19	cgtaccgcctatgccag	Annealing oligos for RBS i7 flanked by EcoRI and KpnI (antisense)	This study
Primer 20	aattctggcataggcggtacgggtac	Annealing oligos for RBS i7 flanked by EcoRI and KpnI (sense)	This study

Chapter four

Primer 21	<i>gcaggaaaga</i> acatgtctggaagatgccaggaa-gat	Amplifying p15A origin flanked by PciI (forward)	This study
Primer 22	ttggtgagaatccaagcact	Amplifying p15A origin (reverse)	This study
Primer 23	aacgaaaggctcagtcgaaa	Origin sequencing primer (forward)	This study
Primer 24	gcaggaaagaactagtggtggcactctgtcgatac	Amplifying puc origin flanked by AvrII (forward)	This study
Primer 25	gactccaagcactagtaggggataacgcag-gaaaga	Amplifying pUC origin (reverse)	This study
Anti-ftsZ synRNA sequence	5'-TCGTCTTCACGGATCCTAATCACCGCGTCATTGGTAAGTTCCATT- GGTTCAAACATTTTCTGTTGGGCCATTGCATTGCCACTGATTTTCCAACATATAAAAAGACAAGC CCGAACAGTCGTCCGGGCTTTTTTCTCGAGCTTCTAGACATCTAAGAA-3'		Chapter 3

4.4.2 Cell growth for genetic constructs characterization

Cells from overnight culture in LB media with appropriate antibiotics (100 $\mu\text{g/ml}$ and 34 $\mu\text{g/ml}$ for ampicillin, and chloramphenicol respectively) were diluted 100 times in fresh LB media with antibiotics and inducers IPTG (0 mM and 10 mM) or aTc(0 ng/ml of aTc and 200 ng/ml) (Sigma Aldrich) as appropriate. It was grown for 10 hrs at 37°C, ~ 250 rpm followed by dilution in fresh media, regrown for another 6 hours at 37°C, ~ 250 rpm before measurements.

4.4.3 Fluorescence and optical density (OD) measurements

Cells were washed 2 times and diluted in phosphate buffer saline pH around 7.4. and loaded onto 96-well black microtiter plate (Greiner Bio-One). The fluorescence was measured in a Synergy HTX Multi-Mode reader (Biotek Instruments, USA). For EGFP and TdTomato the cells were excited with a white light passed through excitation filters 485/20 nm and 540/35 nm respectively and emission was passed through 516/20 and 590/20nm bandpass filters respectively. The density of the cell samples was measured as OD600. First, OD normalized value of raw fluorescence data was calculated by dividing raw fluorescence values with respective OD600. Similarly, OD normalized autofluorescence was measured with cells without the plasmids divided by respective OD600. Next, we subtracted the OD normalized auto-fluorescence from the OD normalized fluorescence to get a single experiment normalized fluorescence value for each experiment. Average normalized fluorescence for each data point (each specific induced or uninduced condition) was determined by averaging single experiment normalized fluorescent values over a minimum of 3 biological replicates. For a set of

Chapter four

experiment the highest average normalized fluorescence value obtained at a specific data point (e.g. in presence of both ATC and IPTG) within the set of experiments (i.e. experiments with combination of ATC and IPTG) was taken as 1 and all other average normalized fluorescence values within that specific set of experiment were scaled accordingly. These scaled values were denoted as normalized fluorescence (a.u.) in the figures. For the 3D input-output relationship experiments single biological replicate was characterized. Here both the IPTG and ATC concentration were varied simultaneously and EGFP and TdTomato expression were measured against around 130 different IPTG and ATC concentration points, which are different than the ATC/IPTG concentration in dose-response curves.

4.4.4 Working principle of Synergy HTX Multi-Mode Reader (Biotek Instruments, USA)

See section 3.4.3 on page number 117

4.4.5 Calculating translation initiation rate in RBS Library Calculator

Translation initiation rate for EGFP was calculated from "evaluate RBS library" design methods in RBS Library Calculator v2.0 considering EcoRI site and its upstream PR promoter sequence as the pre-sequence, EGFP as the protein coding sequence, normal RBS along with conserved linker GGTACC (KpnI site) as RBS constraint sequence and *E. coli* str. K-12 substr. MG1655 as the organism. The RBS calculator is used to computationally generate synthetic ribosome binding site sequences to rationally control the production rate of any protein in bacteria from 0.1 to 100,000+ on a proportional scale (Salis, Mirsky and Voigt, 2009; Espah Borujeni, Anirudh S. Channarasappa and Salis,

2014). Translation initiation is the rate-limiting step in the bacterial mRNA to protein translation process, and its rate is regulated by a number of molecular interactions that occur throughout the process, including 16SrRNA hybridization and folding on mRNA, start codon-tRNA^{fmet} anticodon loop interaction, the distance between the 16S rRNA binding site and the start codon on the mRNA sequence, and stability. A thermodynamic model has been developed using free energy equations for both of these molecular interactions and previously documented observations of RNA hybridization and folding-induced energy adjustments in a sequence-dependent manner. Additionally, another genetic algorithm was developed to control the search mode, which includes multiple search choices such as RBS sequence constraints and search range. RBS Library Calculator was created by integrating this thermodynamic model and genetic algorithm, which iteratively performs four steps (in silico mutation followed by recombination, estimation, and finally selection) to produce RBS libraries with RBS sequences. All iterative processes have two goals: (1) to optimise search coverage across a user-defined range of translation rates, and (2) to ensure that the library contains the fewest possible RBS entries. This tool has a drawback in that it cannot identify protein-coding sequences that begin with any other codon than the most popular start codon ATG (Salis, Mirsky and Voigt, 2009; Espah Borujeni, Anirudh S. Channarasappa and Salis, 2014).

4.4.6 Working principle of fluorescence microscope (confocal LSM 710, Ziess)

Refer to chapter 3.

4.4.7 Microscopy

The bacterial cells were harvested at 4000 rpm, washed and finally diluted in phosphate buffered saline (PBS, pH= 7.4). To visualize, the cell suspension was put onto a 1% thin agarose pad (SeaKem LE agarose) and was covered with a cover slip. The differential interference contrast (DIC) and fluorescence imaging were performed in Zeiss LSM 710 confocal microscope in an epifluorescence mode, keeping the pin-hole completely open. The cells expressing EGFP were excited using 488nm laser and fluorescence was collected through appropriate dichroic and filters sets for EGFP. The DIC images were taken using a DIC attachment. The images were captured and post-processed using Zen 2009 software.

Conclusions and future directions

In this thesis I explored systems and synthetic biology strategies for applications in microgravity related space biology and gene circuits. Microgravity is one of the challenging aspects in manned space missions. The effect of microgravity at molecular genetic level is still not properly understood. In second chapter we have developed a systems biology pipeline to understand the pathway and network-level picture of the human cell under microgravity, which cannot be done by conventional data analysis method. Microgravity is a weak perturbation; thus, the differential gene expression is low for most of the genes, which prevents the identification of altered molecular pathways. Here we used Gene set enrichment analysis, to determine which molecular pathways were altered by microgravity directly from the global gene expression data in a statistically significant way. We found signatures related to immunity, cancer and other diseases. We further used an unsupervised machine learning algorithm to identify plausible functional and regulatory networks among affected genes. Our findings indicate a set of specific hypotheses that can be tested directly in an earth-based microgravity simulator or in space flight conditions in future. These findings can aid in risk assessment and the development of new medication (Morabito *et al.*, 2020) or countermeasures to address microgravity-induced health hazards. Moreover, we established a systems biology pipeline, which can be implemented in gene expression studies where the perturbation is modest and leads to a very low fold change in gene expression and can be used for other model systems. This approach is very important for space experiments where maximum amount of insight is required. Later on this systems biology analysis in our lab revealed new putative pathways and networks related to increased virulence, survivability and

Conclusion

antibiotic resistance of two Proteobacteria (Roy, Shilpa and Bagh, 2016). Recently, similar approach has been used in a multiomic study of human cell in space (da Silveira *et al.*, 2020).

In another part of my work (third chapter) we successfully designed, fabricated, and characterized the first biochemical and cellular microgravity responsive device in *E.coli*. Synthetic biology is being considered as an alternative technology for future space mission (Montague, George H. McArthur, *et al.*, 2012; Menezes, Cumbers, *et al.*, 2015; Menezes, Montague, *et al.*, 2015; Verseux *et al.*, 2015; Hall, 2017; Vitug, 2020). Apart from the chemical signals, synthetic cellular devices have been created to sense and integrate physical signals like temperature (Bagh *et al.*, 2011a) and light (Olson and Tabor, 2014; Liu *et al.*, 2018). However, no molecular or biological microgravity responsive device has been created. The presence of microgravity responsive device may lead to new applications relevant to space and microgravity. The device integrates microgravity to control the expression of a fluorescent gene (EGFP, Td-tomato). The design was universal and we demonstrated that by controlling the cell division process using microgravity. Owing to the device's ability to integrate microgravity as a physical signal with any downstream synthetic or native cellular processes in *E.coli*, it may find future space applications in microgravity responsive metabolic engineering and in the development of cellular microgravity-health hazards monitoring systems. However, to determine the device's true usefulness, it must be tested in space, since we used simulated microgravity. As a future direction, using this understanding we can engineer bacterial strains that are resistant to microgravity, because the effects of microgravity can induce stress in bacterial cell, and may lead to failure of any engineered genetic

Conclusion

programs (*How Synthetic Biology Will Solve Biological Mysteries and Make Humans Safer in Space* by Lucas Hartsough - *The Official PLOS Blog*, no date).

Another part of my work focusses on a process pipeline ‘Network-Brick’ for synthetic circuit fabrication and demonstrated its application in creating complex genetic circuit in single cell. The efficient assembly of genetic elements is needed for the reliable construction of complex genetic circuits for space synthetic biology, and synthetic biology as a whole. In conclusion, a new plasmid vector device named NETWORK Brick was developed in which any biopart inside a single genetic cassette can be conveniently substituted regardless of its location within the cassette and several cassettes can be inserted in any desired direction within a single plasmid. This approach can be applied to the construction of complex synthetic genetic circuits by small academic laboratories lacking automation. Inspired from the physical hierarchy of electronics, we demonstrated the construction and optimization of a totally and chemically controllable in-cell molecular genetic NOR and IMPLY gate in a single *E.coli* strain capable of sensing, integrating, and processing two extracellular environmental chemical signals using this approach. The completion of the full spectrum IMPLY gate incorporates constraint-based fundamental logic into in-cell molecular computation, complete the toolbox. Our architecture is generic and can be used to build additional IMPLY gates using equivalent inducer-repressor interactions. The IMPLY gate's efficient operation inside a two-input-two-output integrated circuit demonstrates its capability as a component of a broader circuit. Due to its superior differentiation capabilities between combinations of input signals, the two-input-2-output integrated circuit is significant as a cell-based chemical signal decoding system. Since the output genes of IMPLY and the two-input two-output circuits may be manipulated to any genes, this may be useful

Conclusion

for externally controlling cellular or metabolic pathways in *E.coli* in a logical manner. Due to the importance of constraint-based logic in robotics and regulation, the constraint-based genetic IMPLY algorithm could have implications for the potential growth of microbiorobotics and chemical control in future. Additionally, since the IMPLY gate and the combinatorial circuit it generates is implemented in a single cell, and indeed in a single plasmid, the architecture may be used to regulate intracellular processes in a logical manner by altering its output proteins. Additionally, we built a modified PCR-based NETWORK Brick that eliminated the need for gel separation and successfully demonstrated the construction and characterization of a genetic exchanger gate. Due to the fact that the whole procedure would not need a gel extraction system, it can be streamlined using a liquid handling method. Taken together, the research has implications for in-cell complex molecular computing, intelligent whole-cell biosensors, signal decoding devices, microbiorobotics, and space synthetic biology.

Bibliography

(*No Title*) (no date a). Available at: https://www.nasa.gov/pdf/185052main_FSB2000-2002REPORT.pdf (Accessed: 19 May 2021).

(*No Title*) (no date b). Available at: https://www.nasa.gov/sites/default/files/atoms/files/space_radiation_ebook.pdf (Accessed: 5 April 2021).

(*No Title*) (no date c). Available at: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un_2017_world_population_prospects-2017_revision_databooklet.pdf (Accessed: 9 February 2021).

Abel, A. J. *et al.* (2020) 'Systems-informed genome mining for electroautotrophic microbial production', *bioRxiv*. bioRxiv, p. 2020.12.07.414987. doi: 10.1101/2020.12.07.414987.

Alvarez, C. H. (2015) *DNA assembly and cloning in an overnight run with the BioXpTM 3200 system*, Nature Publishing Group. doi: 10.1038/nmeth.f.390.

Anderson, J. C., Voigt, C. A. and Arkin, A. P. (2007) 'Environmental signal integration by a modular AND gate.', *Molecular systems biology*, 3(1), p. 133. doi: 10.1038/msb4100173.

Anderson, Jc. *et al.* (2010) 'BglBricks: A flexible standard for biological part assembly', *Journal of Biological Engineering*, 4(1), p. 1. doi: 10.1186/1754-1611-4-1.

Andrianantoandro, E., Basu, S., Karig, David K, *et al.* (2006) 'Synthetic biology: new engineering rules for an emerging discipline', *Molecular Systems Biology*, 2(1), p. 2006.0028. doi: 10.1038/msb4100073.

Bibliography

Andrianantoandro, E., Basu, S., Karig, David K., *et al.* (2006) ‘Synthetic biology: New engineering rules for an emerging discipline’, *Molecular Systems Biology*. European Molecular Biology Organization, p. 2006.0028. doi: 10.1038/msb4100073.

Arunasri, K. *et al.* (2013) ‘Effect of Simulated Microgravity on E. coli K12 MG1655 Growth and Gene Expression’, *PLOS ONE*, 8(3), p. e57860. doi: 10.1371/JOURNAL.PONE.0057860.

Ashkenasy, G. *et al.* (2017) ‘Systems chemistry’, *Chemical Society Reviews*, 46(9), pp. 2543–2554. doi: 10.1039/C7CS00117G.

Assay: Hybrid Bridges the Microplate Reader Gap (no date). Available at: <https://www.genengnews.com/magazine/84/assay-hybrid-bridges-the-microplate-reader-gap/> (Accessed: 3 May 2021).

Aunins, T. R. *et al.* (2018) ‘Spaceflight Modifies Escherichia coli Gene Expression in Response to Antibiotic Exposure and Reveals Role of Oxidative Stress Response’, *Frontiers in Microbiology*, 9(MAR), p. 310. doi: 10.3389/fmicb.2018.00310.

Ausländer, S. *et al.* (2012) ‘Programmable single-cell mammalian biocomputers’, *Nature*, 487, p. 123. Available at: <https://doi.org/10.1038/nature11149>.

Averesch, N. J. H. and Rothschild, L. J. (2019) ‘Metabolic engineering of *Bacillus subtilis* for production of *para* -aminobenzoic acid – unexpected importance of carbon source is an advantage for space application’, *Microbial Biotechnology*, 12(4), pp. 703–714. doi: 10.1111/1751-7915.13403.

Bagh, S. *et al.* (2011a) ‘An active intracellular device to prevent lethal disease outcomes in virus-infected bacterial cells’, *Biotechnology and Bioengineering*, 108(3), pp. 645–654. doi: 10.1002/bit.22969.

Bagh, S. *et al.* (2011b) ‘An active intracellular device to prevent lethal disease outcomes

Bibliography

in virus-infected bacterial cells', *Biotechnology and Bioengineering*, 108(3), pp. 645–654. doi: 10.1002/bit.22969.

Bagh, S., Mandal, M. and McMillen, D. R. (2010) 'Minimal genetic device with multiple tunable functions', *Physical Review E*, 82(2), p. 21911. doi: 10.1103/PhysRevE.82.021911.

Barzegari, A. and Saei, A. A. (2012) 'An update to space biomedical research: Tissue engineering in microgravity bioreactor', *Bioimpacts*, 2, pp. 23–32.

Basu, S. *et al.* (2005) 'A synthetic multicellular system for programmed pattern formation', *Nature*, 434(7037), pp. 1130–1134. doi: 10.1038/nature03461.

Binzel, R. P. (2014) 'Human spaceflight: Find asteroids to get to Mars', *Nature*, 514, pp. 559–561.

Bizzarri, M. (2014) 'Systems Biology and Microgravity Effects on Living Organisms'. doi: 10.4172/2332-0737.1000e111.

Böhmer, M. and Schleiff, E. (2019) 'Microgravity research in plants', *EMBO reports*, 20(7). doi: 10.15252/embr.201948541.

Bonnet, J. *et al.* (2013) 'Amplifying Genetic Logic Gates', *Science*, 340(6132), pp. 599 LP – 603. doi: 10.1126/science.1232758.

Borghetti, J. *et al.* (2010) "'Memristive" switches enable "stateful" logic operations via material implication', *Nature*, 464, p. 873. Available at: <https://doi.org/10.1038/nature08940>.

Briegleb, W. (1992) 'Some qualitative and quantitative aspects of the fast-rotating clinostat as a research tool.', *ASGSB bulletin : publication of the American Society for Gravitational and Space Biology*, 5(2), pp. 23–30.

Brunet, J. P. *et al.* (2004) 'Metagenes and molecular pattern discovery using matrix

Bibliography

factorization', *Proceedings of the National Academy of Sciences of the United States of America*, 101(12), pp. 4164–4169. doi: 10.1073/pnas.0308531101.

Cameron, D. E., Bashor, C. J. and Collins, J. J. (2014) 'A brief history of synthetic biology', *Nature Reviews Microbiology*, 12(5), pp. 381–390. doi: 10.1038/nrmicro3239.

Carlson, R. (2009) 'The changing economics of DNA synthesis', *Nature Biotechnology*. Nature Publishing Group, pp. 1091–1094. doi: 10.1038/nbt1209-1091.

Castelfranchi, C. and Lespérance, Y. (2001) *Intelligent agents VII: agent theories architectures and languages: 7th International Workshop, ATAL 2000, Boston, MA, USA, July 7-9, 2000: proceedings, Proceedings of the 7th International Workshop on Intelligent Agents VII. Agent Theories Architectures and Languages*. Springer. Available at: <https://dl.acm.org/citation.cfm?id=749630> (Accessed: 21 May 2019).

Castro, S. L. *et al.* (2011) 'Induction of attachment-independent biofilm formation and repression of Hfq expression by low-fluid-shear culture of *Staphylococcus aureus*.', *Applied and environmental microbiology*, 77(18), pp. 6368–78. doi: 10.1128/AEM.00175-11.

Çelen, İ. *et al.* (2019) 'Comparative Transcriptomic Signature of the Simulated Microgravity Response in *Caenorhabditis elegans*', *bioRxiv*. bioRxiv, p. 531335. doi: 10.1101/531335.

Censi, F. *et al.* (2018) 'System biology approach: Gene network analysis for muscular dystrophy', in *Methods in Molecular Biology*. Humana Press Inc., pp. 75–89. doi: 10.1007/978-1-4939-7374-3_6.

Chang, T. T. *et al.* (2012) 'The Rel/NF- κ B pathway and transcription of immediate early genes in T cell activation are inhibited by microgravity.', *Journal of leukocyte biology*,

Bibliography

92(6), pp. 1133–45. doi: 10.1189/jlb.0312157.

Che, A. (2004) ‘BioBricks++: Simplifying Assembly of Standard DNA Components’. Available at: <https://dspace.mit.edu/handle/1721.1/39832> (Accessed: 21 May 2019).

Checinska Sielaff, A. *et al.* (2019) ‘Characterization of the total and viable bacterial and fungal communities associated with the International Space Station surfaces’, *Microbiome*, 7(1), pp. 1–21. doi: 10.1186/s40168-019-0666-x.

Chen, Y.-J. *et al.* (2013) ‘Characterization of 582 natural and synthetic terminators and quantification of their design constraints’, *Nature Methods*, 10(7), pp. 659–664. doi: 10.1038/nmeth.2515.

Cichan, T. *et al.* (2017) ‘Mars Base Camp: An Architecture for Sending Humans to Mars’, *New Space*, 5(4), pp. 203–218. doi: 10.1089/space.2017.0037.

Clément, G. (2006) ‘Introduction to Space Biology’, in *Fundamentals of Space Biology*. Springer New York, pp. 1–50. doi: 10.1007/0-387-37940-1_1.

Cook, M. E. and Croxdale, J. G. (2003) ‘Ultrastructure of potato tubers formed in microgravity under controlled environmental conditions’, *Journal of Experimental Botany*, 54(390), pp. 2157–2164. doi: 10.1093/jxb/erg218.

Crucian, B., Stowe, Raymond P, *et al.* (2015) ‘Alterations in adaptive immunity persist during long-duration spaceflight’, *Npj Microgravity*, 1, p. 15013. Available at: <http://dx.doi.org/10.1038/npjmgrav.2015.13>.

Crucian, B., Stowe, Raymond P., *et al.* (2015) ‘Alterations in adaptive immunity persist during long-duration spaceflight’, *npj Microgravity*, 1, p. 15013. doi: 10.1038/npjmgrav.2015.13.

Cucinotta, F. a *et al.* (2001) ‘Space radiation cancer risks and uncertainties for Mars missions.’, *Radiation research*, 156(5 Pt 2), pp. 682–8. doi: 10.1667/0033-

Bibliography

7587(2001)156[0682:SRCRAU]2.0.CO;2.

Devarajan, K. (2008) ‘Nonnegative matrix factorization: An analytical and interpretive tool in computational biology’, *PLoS Computational Biology*, p. e1000029. doi: 10.1371/journal.pcbi.1000029.

DM, K. and HN, H. (2006) ‘Antibiotic efficacy and microbial virulence during space flight’, *Trends in biotechnology*, 24(3), pp. 131–136. doi: 10.1016/J.TIBTECH.2006.01.008.

Drake, B. G. (2009) *Human Exploration of Mars Design Reference Architecture 5.0 Mars Architecture Steering Group NASA Headquarters*. Available at: <http://ston.jsc.nasa.gov/collections/TRS/> (Accessed: 24 February 2021).

Ellis, T., Adie, T. and Baldwin, G. S. (2011) ‘DNA assembly for synthetic biology: From parts to pathways and beyond’, *Integrative Biology*, 3(2), pp. 109–118. doi: 10.1039/c0ib00070a.

Espah Borujeni, A., Channarasappa, Anirudh S and Salis, H. M. (2014) ‘Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites’, *Nucleic Acids Research*, 42(4), pp. 2646–2659. Available at: <http://dx.doi.org/10.1093/nar/gkt1139>.

Espah Borujeni, A., Channarasappa, Anirudh S. and Salis, H. M. (2014) ‘Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites’, *Nucleic Acids Research*, 42(4), pp. 2646–2659. doi: 10.1093/nar/gkt1139.

Fernandez-Rodriguez, J. *et al.* (2017) ‘Engineering RGB color vision into *Escherichia coli*’, *Nature Chemical Biology*, 13(7), pp. 706–708. doi: 10.1038/nchembio.2390.

Flynn, M. T. (no date) *WATER WALLS ARCHITECTURE: MASSIVELY REDUNDANT*

Bibliography

AND HIGHLY RELIABLE LIFE SUPPORT FOR LONG DURATION EXPLORATION MISSIONS.

Franceschini, A. *et al.* (2013) ‘STRING v9.1: Protein-protein interaction networks, with increased coverage and integration’, *Nucleic Acids Research*, 41(D1), pp. D808–D815. doi: 10.1093/nar/gks1094.

Fric, J. *et al.* (2012) ‘NFAT control of innate immunity’, *Blood*, pp. 1380–1389. doi: 10.1182/blood-2012-02-404475.

Friedland, A. E. *et al.* (2009) ‘Synthetic gene networks that count’, *Science*, 324(5931), pp. 1199–1202. doi: 10.1126/science.1172005.

Fu, J. (2014) ‘Identification of biomarkers in breast cancer by gene expression profiling using human tissues’, *Genom. Data*, 2, pp. 299–301.

Fu, J. *et al.* (2014) ‘Identification of Biomarkers in Breast Cancer by Gene Expression Profiling Using Human Tissues’, *Genom Data*, 2, pp. 299–301. doi: 10.1016/j.gdata.2014.09.004.

G, D.-L. *et al.* (2014) ‘Microflow1, a sheathless fiber-optic flow cytometry biomedical platform: demonstration onboard the international space station’, *Cytometry. Part A: the journal of the International Society for Analytical Cytology*, 85(4), pp. 322–331. doi: 10.1002/CYTO.A.22427.

G, G. *et al.* (1994) ‘Growth and division of Escherichia coli under microgravity conditions’, *Research in microbiology*, 145(2), pp. 111–120. doi: 10.1016/0923-2508(94)90004-3.

Genee, H. J. *et al.* (2016) ‘Functional mining of transporters using synthetic selections’, *Nature Chemical Biology*, 12, p. 1015. Available at: <https://doi.org/10.1038/nchembio.2189>.

Bibliography

Gibson, D. G. *et al.* (2008) ‘Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome’, *Science*, 319(5867), pp. 1215–1220. doi: 10.1126/science.1151721.

Gibson, D. G. *et al.* (2009) ‘Enzymatic assembly of DNA molecules up to several hundred kilobases’, *Nature Methods*, 6(5), pp. 343–345. doi: 10.1038/nmeth.1318.

Gilbert, R. *et al.* (2020) ‘Spaceflight and simulated microgravity conditions increase virulence of *Serratia marcescens* in the *Drosophila melanogaster* infection model’, *npj Microgravity*, 6(1), pp. 1–9. doi: 10.1038/s41526-019-0091-2.

Girardi, C. (2014) ‘Integration analysis of microRNA and mRNA expression profiles in human peripheral blood lymphocytes cultured in modeled microgravity’, *Biomed. Res. Int.*, 2014, p. 296747.

Girardi, C. *et al.* (2014) ‘Integration analysis of MicroRNA and mRNA expression profiles in human peripheral blood lymphocytes cultured in modeled microgravity’, *BioMed Research International*, 2014, p. 296747. doi: 10.1155/2014/296747.

Gòdia, F. *et al.* (2002) ‘MELISSA: A loop of interconnected bioreactors to develop life support in Space’, *Journal of Biotechnology*, 99(3), pp. 319–330. doi: 10.1016/S0168-1656(02)00222-5.

Goñi-Moreno, A. and Nikel, P. I. (2019) ‘High-Performance Biocomputing in Synthetic Biology–Integrated Transcriptional and Metabolic Circuits’, *Frontiers in Bioengineering and Biotechnology*, 7, p. 40. doi: 10.3389/fbioe.2019.00040.

Gottesman, S. and Storz, G. (2011) ‘Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations’, *Cold Spring Harbor Perspectives in Biology*, 3(12), p. a003798. doi: 10.1101/CSHPERSPECT.A003798.

Green, A. A. *et al.* (2017) ‘Complex cellular logic computation using ribocomputing

Bibliography

devices', *Nature*, 548, p. 117. Available at: <https://doi.org/10.1038/nature23271>.

Guet, C. C. *et al.* (2002) 'Combinatorial Synthesis of Genetic Networks', *Science*, 296(5572), pp. 1466 LP – 1470. doi: 10.1126/science.1067407.

Guo, J. *et al.* (2018) 'Light-driven fine chemical production in yeast biohybrids', *Science*, 362(6416), pp. 813–816. doi: 10.1126/science.aat9777.

Gupta, N. *et al.* (2019) 'Cell-based biosensors: Recent trends, challenges and future perspectives', *Biosensors and Bioelectronics*. Elsevier Ltd, p. 111435. doi: 10.1016/j.bios.2019.111435.

H, P. *et al.* (2013) 'Exploring sRNA-mediated gene silencing mechanisms using artificial small RNAs derived from a natural RNA scaffold in *Escherichia coli*', *Nucleic acids research*, 41(6), pp. 3787–3804. doi: 10.1093/NAR/GKT061.

Häder, D.-P. *et al.* (2017) 'Gravireceptors in eukaryotes—a comparison of case studies on the cellular level', *npj Microgravity*, 3(1), p. 13. doi: 10.1038/s41526-017-0018-8.

Hall, E. C. (1996) *Journey to the Moon: The History of the Apollo Guidance Computer*, *Journey to the Moon: The History of the Apollo Guidance Computer*. American Institute of Aeronautics and Astronautics. doi: 10.2514/4.868023.

Hall, L. (2017) 'A Synthetic Biology Architecture to Detoxify and Enrich Mars Soil'. Available at: http://www.nasa.gov/directorates/spacetech/niac/2017_Phase_I_Phase_II/Mars_Soil_Agriculture (Accessed: 27 April 2021).

Hammond, T. G. *et al.* (1999) 'Gene expression in space [3]', *Nature Medicine*. Nature Publishing Group, p. 359. doi: 10.1038/7331.

Hauschild, S. *et al.* (2014) 'T cell regulation in microgravity - The current knowledge from in vitro experiments conducted in space, parabolic flights and ground-based

Bibliography

facilities', *Acta Astronautica*, 104(1), pp. 365–377. doi: 10.1016/j.actaastro.2014.05.019.

Hawkins, W. R. and Ziegelschmid, J. F. (1975) 'Clinical aspects of crew health. Biomedical Results of Apollo', *Nasa Spec. Rep.*, pp. 43–81.

Hensel, W. and Sievers, A. (1980) 'Effects of prolonged omnilateral gravistimulation on the ultrastructure of statocytes and on the graviresponse of roots', *Planta*, 150(4), pp. 338–346. doi: 10.1007/BF00384664.

Hibbs, K. *et al.* (2004) 'Differential gene expression in ovarian carcinoma: Identification of potential biomarkers', *American Journal of Pathology*, 165(2), pp. 397–414. doi: 10.1016/S0002-9440(10)63306-8.

Horneck, G., Klaus, D. M. and Mancinelli, R. L. (2010) 'Space Microbiology', *Microbiology and Molecular Biology Reviews*, 74(1), pp. 121–156. doi: 10.1128/mmbr.00016-09.

Horton, R. M. *et al.* (1989) 'Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension', *Gene*, 77(1), pp. 61–68. doi: 10.1016/0378-1119(89)90359-4.

How Synthetic Biology Will Solve Biological Mysteries and Make Humans Safer in Space by Lucas Hartsough - The Official PLOS Blog (no date). Available at: <https://theplosblog.plos.org/2016/02/how-synthetic-biology-will-solve-biological-mysteries-and-make-humans-safer-in-space-by-lucas-hartsough/> (Accessed: 6 May 2021).

Hroudová, J., Singh, N. and Fišar, Z. (2014) 'Mitochondrial dysfunctions in neurodegenerative diseases: Relevance to alzheimer's disease', *BioMed Research International*. Hindawi Publishing Corporation, p. 175062. doi: 10.1155/2014/175062.

Bibliography

Human Health on the Space Station (no date). Available at: <https://www.issnationallab.org/iss360/human-health-on-the-space-station/> (Accessed: 6 April 2021).

Huntress, W. *et al.* (2006) 'The next steps in exploring deep space - A cosmic study by the IAA', *Acta Astronautica*. Elsevier Ltd, pp. 304–377. doi: 10.1016/j.actaastro.2006.01.004.

Ideker, T., Galitski, T. and Hood, L. (2001) 'A NEW APPROACH TO DECODING LIFE: Systems Biology', *Annual Review of Genomics and Human Genetics*, 2(1), pp. 343–372. doi: 10.1146/annurev.genom.2.1.343.

igem.org (no date). Available at: https://igem.org/Main_Page (Accessed: 26 April 2021).

Influence of Microgravity Environment on Root Growth, Soluble Sugars, and Starch Concentration of Sweetpotato Stem Cuttings in: Journal of the American Society for Horticultural Science Volume 133 Issue 3 (2008) (no date). Available at: <https://journals.ashs.org/jashs/view/journals/jashs/133/3/article-p327.xml> (Accessed: 8 June 2021).

INGBER, D. (1999) 'How cells (might) sense microgravity', *The FASEB Journal*, 13(9001), pp. S3–S15. doi: 10.1096/fasebj.13.9001.s3.

IV Konstantinova, Y. A. V. L. V. Z. (1973) 'Study of reactivity of blood lymphoid cells in crew members of the Soyuz-6, Soyuz-7 and Soyuz-8 spaceships before and after flight', *Space Biol Med*, 7, pp. 48–55.

JS, F. *et al.* (2013) 'Impact of simulated microgravity on the normal developmental time line of an animal-bacteria symbiosis', *Scientific reports*, 3. doi: 10.1038/SREP01340.

Bibliography

- Jusiak, B. *et al.* (2016) 'Engineering Synthetic Gene Circuits in Living Cells with CRISPR Technology', *Trends in Biotechnology*, 34(7), pp. 535–547. doi: 10.1016/j.tibtech.2015.12.014.
- Kampis, G. (1991) *Self-modifying systems in biology and cognitive science : a new framework for dynamics, information, and complexity*. Pergamon Press.
- Kim, H. W., Matin, A. and Rhee, M. S. (2014) 'Microgravity Alters the Physiological Characteristics of Escherichia coli O157:H7 ATCC 35150, ATCC 43889, and ATCC 43895 under Different Nutrient Conditions', *Applied and Environmental Microbiology*, 80(7), p. 2270. doi: 10.1128/AEM.04037-13.
- Kim, M., Steager, E. and Agung, J. (2012) *Microbiorobotics : biologically inspired microscale robotic systems*. Elsevier.
- Kim, P. M. and Tidor, B. (2003) 'Subsystem identification through dimensionality reduction of large-scale gene expression data', *Genome Research*, 13(7), pp. 1706–1718. doi: 10.1101/gr.903503.
- Kinney, M. A. *et al.* (2019) 'Author Correction: A systems biology pipeline identifies regulatory networks for stem cell engineering.', *Nature biotechnology*, 37(8), p. 962. doi: 10.1038/s41587-019-0212-1.
- Kis, Z. *et al.* (2015) 'Mammalian synthetic biology: Emerging medical applications', *Journal of the Royal Society Interface*. Royal Society of London. doi: 10.1098/rsif.2014.1000.
- Kitano, H. (2002) 'Systems biology: A brief overview', *Science*. American Association for the Advancement of Science, pp. 1662–1664. doi: 10.1126/science.1069492.
- Kitto, R. Z. *et al.* (2021) 'Synthetic biological circuit tested in spaceflight', *Life Sciences in Space Research*, 28, pp. 57–65. doi: 10.1016/j.lssr.2020.09.002.

Bibliography

- Klaus, D. M., Todd, P. and Schatz, A. (1998) 'Functional weightlessness during clinorotation of cell suspensions.', *Advances in space research : the official journal of the Committee on Space Research (COSPAR)*, 21(8–9), pp. 1315–1318. doi: 10.1016/s0273-1177(97)00404-3.
- Kliem, C. *et al.* (2012) 'Curcumin suppresses T cell activation by blocking Ca²⁺ mobilization and nuclear factor of activated T cells (NFAT) activation', *Journal of Biological Chemistry*, 287(13), pp. 10200–10209. doi: 10.1074/jbc.M111.318733.
- Kohn, F., Hauslage, J. and Hanke, W. (2017) 'Membrane Fluidity Changes, A Basic Mechanism of Interaction of Gravity with Cells?', *Microgravity Science and Technology*, 29(5), pp. 337–342. doi: 10.1007/s12217-017-9552-y.
- Koonin, E. V. and Wolf, Y. I. (2006) 'Evolutionary systems biology: links between gene evolution and function', *Current Opinion in Biotechnology*. Elsevier Current Trends, pp. 481–487. doi: 10.1016/j.copbio.2006.08.003.
- Kuang, A. *et al.* (2000) 'Influence of microgravity on ultrastructure and storage reserves in seeds of *Brassica rapa* L.', *Annals of Botany*, 85(6), pp. 851–859. doi: 10.1006/anbo.2000.1153.
- Laboratory animals in space life sciences research.* (no date). Available at: <https://agris.fao.org/agris-search/search.do?recordID=US19960141836> (Accessed: 19 May 2021).
- Lander, E. S. *et al.* (2001) 'Initial sequencing and analysis of the human genome', *Nature*, 409(6822), pp. 860–921. doi: 10.1038/35057062.
- Lang, T. F. *et al.* (2006) 'Adaptation of the Proximal Femur to Skeletal Reloading After Long-Duration Spaceflight', *Journal of Bone and Mineral Research*, 21(8), pp. 1224–1230. doi: 10.1359/jbmr.060509.

Bibliography

- Langhoff, S., Cumbers, J. and Worden, S. P. (2011) ‘Workshop Report on What are the Potential Roles for Synthetic Biology in NASA ’ s Mission ?’, (March).
- Lee, D. D. and Seung, H. S. (1999) ‘Learning the parts of objects by non-negative matrix factorization’, *Nature*, 401(6755), pp. 788–791. doi: 10.1038/44565.
- Lehner, B. A. E. *et al.* (2019) ‘End-to-end mission design for microbial ISRU activities as preparation for a moon village’, *Acta Astronautica*, 162, pp. 216–226. doi: 10.1016/j.actaastro.2019.06.001.
- Licorresponding, J. (2013) ‘JAK-STAT and bone metabolism’, *JAKSTAT*, 2, p. e23930.
- Link, B. M., Busse, J. S. and Stankovic, B. (2014) ‘Seed-to-seed-to-seed growth and development of Arabidopsis in microgravity’, *Astrobiology*, 14(10), pp. 866–875. doi: 10.1089/ast.2014.1184.
- Liu, Y. and Wang, E. (2008) ‘Transcriptional analysis of normal human fibroblast responses to microgravity stress.’, *Genomics, proteomics & bioinformatics*, 6(1), pp. 29–41. doi: 10.1016/S1672-0229(08)60018-2.
- Liu, Z. *et al.* (2018) ‘Programming bacteria with light-sensors and applications in synthetic biology’, *Frontiers in Microbiology*, 9(NOV). doi: 10.3389/fmicb.2018.02692.
- Löbrich, M. and Jeggo, P. A. (2019) ‘Hazards of human spaceflight’, *Science*, 364(6436), pp. 127 LP – 128. doi: 10.1126/science.aaw7086.
- Lorenzo, V. *et al.* (2018) ‘The power of synthetic biology for bioproduction, remediation and pollution control’, *EMBO reports*, 19(4), p. e45658. doi: 10.15252/embr.201745658.
- Lutz, R. and Bujard, H. (1997a) ‘Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory

Bibliography

- elements.’, *Nucleic Acids Research*, 25(6), p. 1203. doi: 10.1093/NAR/25.6.1203.
- Lutz, R. and Bujard, H. (1997b) ‘Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements’, *Nucleic acids research*, 25(6), pp. 1203–1210. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/9092630>.
- Maier, Jeanette A M *et al.* (2015) ‘The impact of microgravity and hypergravity on endothelial cells’, *BioMed Research International*. doi: 10.1155/2015/434803.
- Maier, Jeanette A.M. *et al.* (2015) ‘The impact of microgravity and hypergravity on endothelial cells’, *BioMed Research International*. Hindawi Limited, p. 434803. doi: 10.1155/2015/434803.
- Mangala, L. S. *et al.* (2011) ‘Effects of simulated microgravity on expression profile of microRNA in human lymphoblastoid cells’, *Journal of Biological Chemistry*, 286(37), pp. 32483–32490. doi: 10.1074/jbc.M111.267765.
- Mars, K. (2018) ‘5 Hazards of Human Spaceflight’. Available at: <http://www.nasa.gov/hrp/5-hazards-of-human-spaceflight> (Accessed: 5 April 2021).
- Mauclaire, L. and Egli, M. (2010) ‘Effect of simulated microgravity on growth and production of exopolymeric substances of *Micrococcus luteus* space and earth isolates’, *FEMS Immunology and Medical Microbiology*, 59(3), pp. 350–356. doi: 10.1111/j.1574-695X.2010.00683.x.
- McNulty, M. J. *et al.* (2020) ‘Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted’. doi: 10.20944/preprints202009.0086.v1.
- Menezes, A. A. *et al.* (2014) ‘Towards synthetic biological approaches to resource utilization on space missions’, *Journal of The Royal Society Interface*, 12(102), pp. 20140715–20140715. doi: 10.1098/rsif.2014.0715.

Bibliography

- Menezes, A. A., Montague, M. G., *et al.* (2015) 'Grand challenges in space synthetic biology.', *Journal of the Royal Society, Interface / the Royal Society*, 12(113), pp. 20150803-. doi: 10.1098/rsif.2015.0803.
- Menezes, A. A., Cumbers, J., *et al.* (2015) 'Towards synthetic biological approaches to resource utilization on space missions'.
- MERKIS, A. I. and LAURINAVICHYUS, R. S. (1983) 'Complete cycle of individual development in plants of *Arabidopsis thaliana* (L.) Heynh. aboard the salyut-7 orbital station', *Complete cycle of individual development in plants of Arabidopsis thaliana (L.) Heynh. aboard the salyut-7 orbital station*, 271–273.
- Mesland, D. A. *et al.* (1996) 'The Free Fall Machine--a ground-based facility for microgravity research in life sciences.', *Microgravity science and technology*, 9(1), pp. 10–14.
- Monsalve, E. *et al.* (2009) 'Notch1 upregulates LPS-induced macrophage activation by increasing NF- κ B activity', *European Journal of Immunology*, 39(9), pp. 2556–2570. doi: 10.1002/eji.200838722.
- Montague, M., McArthur, George H, *et al.* (2012) 'The role of synthetic biology for in situ resource utilization (ISRU).', *Astrobiology*, 12(12), pp. 1135–42. doi: 10.1089/ast.2012.0829.
- Montague, M., McArthur, George H., *et al.* (2012) 'The role of synthetic biology for in situ resource utilization (ISRU)', *Astrobiology*. *Astrobiology*, pp. 1135–1142. doi: 10.1089/ast.2012.0829.
- Moon, T. S. *et al.* (2012) 'Genetic programs constructed from layered logic gates in single cells', *Nature*, 491, p. 249. Available at: <https://doi.org/10.1038/nature11516>.
- Morabito, C. *et al.* (2020) 'Antioxidant strategy to prevent simulated microgravity-

Bibliography

induced effects on bone osteoblasts', *International Journal of Molecular Sciences*, 21(10), p. 3638. doi: 10.3390/ijms21103638.

Mückl, A. *et al.* (2018) 'Filamentation and restoration of normal growth in *Escherichia coli* using a combined CRISPRi sgRNA/antisense RNA approach', *PLOS ONE*, 13(9), p. e0198058. doi: 10.1371/JOURNAL.PONE.0198058.

Mukhopadhyay, S. *et al.* (2016) 'A systems biology pipeline identifies new immune and disease related molecular signatures and networks in human cells during microgravity exposure', *Scientific Reports*, 6, p. 25975. Available at: <http://dx.doi.org/10.1038/srep25975>.

Mukhopadhyay, S. and Bagh, S. (2020) 'A synthetic biochemical device for sensing microgravity', *bioRxiv*, p. 2020.01.26.920629. doi: 10.1101/2020.01.26.920629.

Musk, E. (2017) 'Making Humans a Multi-Planetary Species', *New Space*. Mary Ann Liebert Inc., pp. 46–61. doi: 10.1089/space.2017.29009.emu.

Na, D. *et al.* (2013) 'Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs.', *Nature biotechnology*, 31(2), pp. 170–4. doi: 10.1038/nbt.2461.

NanoRacks Honored with Space Station Innovation Award | NASA (no date). Available at: <https://www.nasa.gov/content/nanoracks-honored-with-space-station-innovation-award/stationresearch/> (Accessed: 30 August 2021).

'NASA - Synthetic Biology and Microbial Fuel Cells: Towards Self-Sustaining Life Support Systems' (no date).

Nickerson, C. A. *et al.* (2000) 'Microgravity as a novel environmental signal affecting *Salmonella enterica* serovar Typhimurium virulence', *Infection and Immunity*, 68(6), pp. 3147–3152. doi: 10.1128/IAI.68.6.3147-3152.2000.

Nickerson, C. A. *et al.* (2004) 'Microbial responses to microgravity and other low-shear

Bibliography

environments.’, *Microbiology and molecular biology reviews : MMBR*, 68(2), pp. 345–61. doi: 10.1128/MMBR.68.2.345-361.2004.

Nomura, Y. and Yokobayashi, Y. (2015) ‘Aptazyme-Based Riboswitches and Logic Gates in Mammalian Cells BT - RNA Scaffolds: Methods and Protocols’, in Ponchon, L. (ed.). New York, NY: Springer New York, pp. 141–148. doi: 10.1007/978-1-4939-2730-2_12.

Norville, J. E. *et al.* (2010) ‘Introduction of customized inserts for streamlined assembly and optimization of BioBrick synthetic genetic circuits’, *Journal of Biological Engineering*, 4(1), p. 17. doi: 10.1186/1754-1611-4-17.

Nyerges, Á. *et al.* (2016) ‘A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species’, *Proceedings of the National Academy of Sciences*, 113(9), pp. 2502 LP – 2507. doi: 10.1073/pnas.1520040113.

O’Sullivan, L. A. *et al.* (2007) ‘Cytokine receptor signaling through the Jak-Stat-Socs pathway in disease’, *Molecular Immunology*, pp. 2497–2506. doi: 10.1016/j.molimm.2006.11.025.

Olabi, A. A. *et al.* (2002) ‘The effect of microgravity and space flight on the chemical senses’, *Journal of Food Science*. Institute of Food Technologists, pp. 468–478. doi: 10.1111/j.1365-2621.2002.tb10622.x.

Olson, E. J. and Tabor, J. J. (2014) ‘Optogenetic characterization methods overcome key challenges in synthetic and systems biology’, *Nature Chemical Biology*, 10(7), pp. 502–511. doi: 10.1038/nchembio.1559.

P, L. *et al.* (2015) ‘Simulated microgravity disrupts intestinal homeostasis and increases colitis susceptibility’, *FASEB journal : official publication of the Federation of*

Bibliography

- American Societies for Experimental Biology*, 29(8), pp. 3263–3273. doi: 10.1096/FJ.15-271700.
- Pardee, K. *et al.* (2014) ‘Paper-based synthetic gene networks’, *Cell*, 159(4), pp. 940–954. doi: 10.1016/j.cell.2014.10.004.
- Parkinson, H. *et al.* (2007) ‘ArrayExpress - A public database of microarray experiments and gene expression profiles’, *Nucleic Acids Research*, 35(SUPPL. 1), pp. D747–D750. doi: 10.1093/nar/gkl995.
- Patti, M. E. *et al.* (2003) ‘Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1’, *Proceedings of the National Academy of Sciences of the United States of America*, 100(14), pp. 8466–8471. doi: 10.1073/pnas.1032913100.
- Perou, C. M. and Borresen-Dale, A. L. (2011) ‘Systems biology and genomics of breast cancer’, *Cold Spring Harbor Perspectives in Biology*, 3(2), pp. 1–17. doi: 10.1101/cshperspect.a003293.
- Petri Plants-2 Experiment Plate Final Survey | NASA Image and Video Library* (no date). Available at: <https://images.nasa.gov/details-iss054e022372> (Accessed: 8 June 2021).
- Phillips, I. and Silver, P. (2006) ‘A New Biobrick Assembly Strategy Designed for Facile Protein Engineering’. Available at: <https://dspace.mit.edu/handle/1721.1/32535> (Accessed: 21 May 2019).
- Planel, H. (2004) *Space and life : an introduction to space biology and medicine*. CRC Press. Available at: <https://www.crcpress.com/Space-and-Life-An-Introduction-to-Space-Biology-and-Medicine/Planel/p/book/9780415317597> (Accessed: 20 November 2019).

Bibliography

Poon, C. (2020) ‘Factors implicating the validity and interpretation of mechanobiology studies in simulated microgravity environments’, *Engineering Reports*, 2(10), p. e12242. doi: 10.1002/eng2.12242.

Prasad, A. R. S. *et al.* (1993) ‘Flow-related responses of intracellular inositol phosphate levels in cultured aortic endothelial cells’, *Circulation Research*, 72(4), pp. 827–836. doi: 10.1161/01.RES.72.4.827.

Quan, J. and Tian, J. (2009) ‘Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways’, *PLoS ONE*. Edited by P. L. Ho, 4(7), p. e6441. doi: 10.1371/journal.pone.0006441.

Registry - igem.org (no date). Available at: <https://igem.org/Registry> (Accessed: 26 April 2021).

Reich, M. *et al.* (2006) ‘GenePattern 2.0 [2]’, *Nature Genetics*. Nature Publishing Group, pp. 500–501. doi: 10.1038/ng0506-500.

Rhesus Monkey - Miss Sam - Fiberglass Couch - Little Joe (LJ)-1B Flight - Prep | NASA Image and Video Library (no date). Available at: <https://images.nasa.gov/details-b59-00828> (Accessed: 24 May 2021).

Ro, D. K. *et al.* (2006) ‘Production of the antimalarial drug precursor artemisinic acid in engineered yeast’, *Nature*, 440(7086), pp. 940–943. doi: 10.1038/nature04640.

Robertson, G. A. (no date) *Growing Spaceships?* Available at: <http://blip.tv/igem-headquarters/igem-explainer-01-drew-endy-defining-> (Accessed: 9 May 2021).

Roquet, N. and Lu, T. K. (2014) ‘Digital and analog gene circuits for biotechnology’, *Biotechnology Journal*, 9(5), pp. 597–608. doi: 10.1002/biot.201300258.

Ross, J. M., Olson, L. and Coppotelli, G. (2015) ‘Mitochondrial and ubiquitin proteasome system dysfunction in ageing and disease: Two sides of the same coin?’,

Bibliography

- International Journal of Molecular Sciences*. MDPI AG, pp. 19458–19476. doi: 10.3390/ijms160819458.
- Roy, R., Shilpa, P. P. and Bagh, S. (2016) ‘A Systems Biology Analysis Unfolds the Molecular Pathways and Networks of Two Proteobacteria in Spaceflight and Simulated Microgravity Conditions’, *Astrobiology*, 16(9), pp. 677–689. doi: 10.1089/ast.2015.1420.
- Rubens, J. R., Selvaggio, G. and Lu, T. K. (2016) ‘Synthetic mixed-signal computation in living cells’, *Nature Communications*, 7, p. 11658. Available at: <https://doi.org/10.1038/ncomms11658>.
- Ruyters, G. and Braun, M. (2014) ‘Plant biology in space: Recent accomplishments and recommendations for future research’, *Plant Biology*, 16(SUPPL.1), pp. 4–11. doi: 10.1111/plb.12127.
- S, B. *et al.* (2008) ‘Plasmid-borne prokaryotic gene expression: sources of variability and quantitative system characterization’, *Physical review. E, Statistical, nonlinear, and soft matter physics*, 77(2 Pt 1). doi: 10.1103/PHYSREVE.77.021919.
- Sakai, Y. *et al.* (2013) ‘Improving the Gene-Regulation Ability of Small RNAs by Scaffold Engineering in Escherichia coli’, *ACS Synthetic Biology*, 3(3), pp. 152–162. doi: 10.1021/SB4000959.
- Salis, H. M., Mirsky, E. A. and Voigt, C. A. (2009) ‘Automated design of synthetic ribosome binding sites to control protein expression.’, *Nature biotechnology*, 27(10), pp. 946–50. doi: 10.1038/nbt.1568.
- Saltepe, B. *et al.* (2018) ‘Cellular Biosensors with Engineered Genetic Circuits’, *ACS Sensors*, 3(1), pp. 13–26. doi: 10.1021/acssensors.7b00728.
- Sánchez-Gorostiaga, A. *et al.* (2016) ‘Life without Division: Physiology of Escherichia

Bibliography

coli FtsZ-Deprived Filaments.’, *mBio*, 7(5). doi: 10.1128/mBio.01620-16.

Sarkar, K. *et al.* (2019) ‘A frame-shifted gene, which rescued its function by non-natural start codons and its application in constructing synthetic gene circuits’, *Journal of Biological Engineering*, 13(1), p. 20. doi: 10.1186/s13036-019-0151-x.

Schröder, K. P. and Cannon Smith, R. (2008) ‘Distant future of the Sun and Earth revisited’, *Monthly Notices of the Royal Astronomical Society*, 386(1), pp. 155–163. doi: 10.1111/j.1365-2966.2008.13022.x.

Schröder, K. P., Smith, R. C. and Apps, K. (2001) ‘Solar evolution and the distant future of earth’, *Astronomy and Geophysics*, 42(6), pp. 6.26-6.29. doi: 10.1046/j.1468-4004.2001.42626.x.

Sen, A. (2015) *Riding Gravity Away from Doomsday*.

Serrano, L. (2007) ‘Synthetic biology: promises and challenges.’, *Molecular systems biology*, p. 158. doi: 10.1038/msb4100202.

Setlow, R. B. (2003) ‘The hazards of space travel’, *EMBO reports*, 4(11), pp. 1013–1016. doi: 10.1038/sj.embor.7400016.

Shaner, N. C. *et al.* (2004) ‘Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein’, *Nature Biotechnology*, 22(12), pp. 1567–1572. doi: 10.1038/nbt1037.

Shetty, R. P., Endy, D. and Knight, T. F. (2008) ‘Engineering BioBrick vectors from BioBrick parts’, *Journal of Biological Engineering*, 2(1), p. 5. doi: 10.1186/1754-1611-2-5.

da Silveira, W. A. *et al.* (2020) ‘Comprehensive Multi-omics Analysis Reveals Mitochondrial Stress as a Central Biological Hub for Spaceflight Impact’, *Cell*, 183(5), pp. 1185-1201.e20. doi: 10.1016/j.cell.2020.11.002.

Bibliography

- Siuti, P., Yazbek, J. and Lu, T. K. (2013) 'Synthetic circuits integrating logic and memory in living cells', *Nature Biotechnology*, 31, p. 448. Available at: <https://doi.org/10.1038/nbt.2510>.
- Solé, R. V., Montañez, R. and Duran-Nebreda, S. (2015) 'Synthetic circuit designs for earth terraformation', *Biology Direct*, 10(1), p. 37. doi: 10.1186/s13062-015-0064-7.
- Souza, G. R. *et al.* (2010) 'Three-dimensional tissue culture based on magnetic cell levitation', *Nature Nanotechnology*, 5(4), pp. 291–296. doi: 10.1038/nnano.2010.23.
- Souza, K., Etheridge, G. and Callahan, P. X. (1991) *Space Life Sciences Experiments Life into Space*. Available at: <http://www.adobe.com> (Accessed: 30 March 2021).
- Souza, K., Hogan, R. and Ballard, R. (1965) *Life into Space, Life into Space*. Available at: <http://www.adobe.com> (Accessed: 30 March 2021).
- Space Physiology* - Jay C. Buckey - Oxford University Press (no date). Available at: <https://global.oup.com/academic/product/space-physiology-9780195137255?cc=us&lang=en&> (Accessed: 18 June 2021).
- Storz, G., Vogel, J. and Wassarman, K. M. (2011) 'Regulation by Small RNAs in Bacteria: Expanding Frontiers', *Molecular cell*, 43(6), p. 880. doi: 10.1016/J.MOLCEL.2011.08.022.
- Stricker, J. *et al.* (2008) 'A fast, robust and tunable synthetic gene oscillator', *Nature*, 456, p. 516. Available at: <https://doi.org/10.1038/nature07389>.
- Subramanian, A., Subramanian, A., *et al.* (2005) 'Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(43), pp. 15545–50. doi: 10.1073/pnas.0506580102.
- Subramanian, A., Tamayo, P., Mootha, Vamsi K, *et al.* (2005) 'Gene set enrichment

Bibliography

analysis: a knowledge-based approach for interpreting genome-wide expression profiles.’, *Proceedings of the National Academy of Sciences of the United States of America*, 102(43), pp. 15545–50. doi: 10.1073/pnas.0506580102.

Subramanian, A., Tamayo, P., Mootha, Vamsi K., *et al.* (2005) ‘Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles’, *Proceedings of the National Academy of Sciences of the United States of America*, 102(43), pp. 15545–15550. doi: 10.1073/pnas.0506580102.

Swart H.C.M., de and Nederpelt, R. P. (1992) ‘Implication : a survey of the different logical analyses of “if ..., then ...” ’, *Nieuw Archief voor Wiskunde* , pp. 77–104.

Sychev, V. N. *et al.* (2007) ‘Spaceflight effects on consecutive generations of peas grown onboard the Russian segment of the International Space Station’, *Acta Astronautica*, 60(4-7 SPEC. ISS.), pp. 426–432. doi: 10.1016/j.actaastro.2006.09.009.

Synthetic life could make trips to Mars more comfortable (no date). Available at: <https://www.nbcnews.com/id/wbna39344314> (Accessed: 3 May 2021).

Szallasi, Z., Stelling, J. and Periwal, V. (2010) *System Modeling in Cell Biology, From Concepts to Nuts and Bolts*. doi: 10.7551/mitpress/9780262195485.001.0001.

Tamsir, A., Tabor, J. J. and Voigt, C. A. (2011) ‘Robust multicellular computing using genetically encoded NOR gates and chemical “wiresg”’, *Nature*, 469(7329), pp. 212–215. doi: 10.1038/nature09565.

The sun won’t die for 5 billion years, so why do humans have only 1 billion years left on Earth? (no date). Available at: <https://phys.org/news/2015-02-sun-wont-die-billion-years.html> (Accessed: 9 February 2021).

Thiel, C. S. *et al.* (2019) ‘Real-Time 3D High-Resolution Microscopy of Human Cells on the International Space Station’, *International Journal of Molecular Sciences*, 20(8).

Bibliography

doi: 10.3390/IJMS20082033.

Thorn, K. (2016) 'A quick guide to light microscopy in cell biology', *Molecular Biology of the Cell*, 27(2), pp. 219–222. doi: 10.1091/mbc.E15-02-0088.

Tsao, Y. M. D. *et al.* (1994) 'Fluid dynamics within a rotating bioreactor in space and earth environments', *Journal of Spacecraft and Rockets*, 31(6), pp. 937–943. doi: 10.2514/3.26541.

VASSY, J. *et al.* (2001) 'The effect of weightlessness on cytoskeleton architecture and proliferation of human breast cancer cell line MCF-7', *The FASEB Journal*, 15(6), pp. 1104–1106. doi: 10.1096/fj.00-0527fje.

Verhaar, A. P. *et al.* (2014a) 'Dichotomal effect of space flight-associated microgravity on stress-activated protein kinases in innate immunity', *Scientific Reports*, 4, p. 5468. doi: 10.1038/srep05468.

Verhaar, A. P. *et al.* (2014b) 'Dichotomal effect of space flight-associated microgravity on stress-activated protein kinases in innate immunity', *Scientific Reports*, 4(1), pp. 1–5. doi: 10.1038/srep05468.

Vernikos, J. *et al.* (2016) 'Theseus: The European research priorities for human exploration of space', *npj Microgravity*. Nature Publishing Group, pp. 1–3. doi: 10.1038/npjmgrav.2016.34.

Versari, S. *et al.* (2013a) 'The challenging environment on board the International Space Station affects endothelial cell function by triggering oxidative stress through thioredoxin interacting protein overexpression: the ESA-SPHINX experiment', *FASEB Journal*, 27(11), pp. 4466–4475. doi: 10.1096/fj.13-229195.

Versari, S. *et al.* (2013b) 'The challenging environment on board the International Space Station affects endothelial cell function by triggering oxidative stress through

Bibliography

thioredoxin interacting protein overexpression: the ESA-SPHINX experiment', *FASEB Journal*, 27(11), pp. 4466–4475. doi: 10.1096/fj.13-229195.

Verseux, C. N. *et al.* (2015) 'Synthetic biology for space exploration: Promises and societal implications', in *Ambivalences of Creating Life: Societal and Philosophical Dimensions of Synthetic Biology*. Springer International Publishing, pp. 73–100. doi: 10.1007/978-3-319-21088-9_4.

Vidyasekar, P. *et al.* (2015) 'Genome wide expression profiling of cancer cell lines cultured in microgravity reveals significant dysregulation of cell cycle and MicroRNA gene networks', *PLoS ONE*, 10(8), p. e0135958. doi: 10.1371/journal.pone.0135958.

Vitug, E. (2020) 'Space Synthetic Biology (SynBio)'. Available at: http://www.nasa.gov/directorates/spacetech/game_changing_development/projects/SynBio (Accessed: 5 May 2021).

Vogel, J. and Luisi, B. F. (2011) 'Hfq and its constellation of RNA', *Nature Reviews Microbiology*, 9(8), pp. 578–589. doi: 10.1038/nrmicro2615.

Wang, B., Barahona, M. and Buck, M. (2013) 'A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals', *Biosensors and Bioelectronics*, 40(1), pp. 368–376. doi: 10.1016/j.bios.2012.08.011.

Wang, X. *et al.* (2021) 'Reversible thermal regulation for bifunctional dynamic control of gene expression in *Escherichia coli*', *Nature Communications* 2021 12:1, 12(1), pp. 1–13. doi: 10.1038/s41467-021-21654-x.

Ward, N. E. *et al.* (2006) 'Gene expression alterations in activated human T-cells induced by modeled microgravity', *Journal of Cellular Biochemistry*, 99(4), pp. 1187–1202. doi: 10.1002/jcb.20988.

Weinstein, R. and Mermel, L. A. (2013) 'Infection prevention and control during

Bibliography

prolonged human space travel', *Clinical Infectious Diseases*, 56(1), pp. 123–130. doi: 10.1093/cid/cis861.

Weiss, R., Weiss, R. and Basu, S. (2002) 'The Device Physics of Cellular Logic Gates'. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.12.1224> (Accessed: 21 May 2019).

What is Synthetic/Engineering Biology? | EBRC (no date). Available at: <https://ebrc.org/what-is-synbio/> (Accessed: 11 June 2021).

Whitehead, A. N. 1861-1947. (2005) *Principia mathematica*, by Alfred North Whitehead ... and Bertrand Russell. Available at: <http://name.umdl.umich.edu/AAT3201.0001.001> (Accessed: 27 May 2020).

Wiesner, T. F., Berk, B. C. and Nerem, R. M. (1997) 'A mathematical model of the cytosolic-free calcium response in endothelial cells to fluid shear stress', *Proceedings of the National Academy of Sciences of the United States of America*, 94(8), pp. 3726–3731. doi: 10.1073/pnas.94.8.3726.

Wilson, J. W. *et al.* (2002) 'Low-shear modeled microgravity alters the *Salmonella enterica* serovar Typhimurium stress response in an RpoS-independent manner', *Applied and Environmental Microbiology*, 68(11), pp. 5408–5416. doi: 10.1128/AEM.68.11.5408-5416.2002.

Wilson, J. W. *et al.* (2007) 'Space flight alters bacterial gene expression and virulence and reveals a role for global regulator Hfq', *Proceedings of the National Academy of Sciences*, 104(41), pp. 16299–16304. doi: 10.1073/pnas.0707155104.

Xu, T. *et al.* (2016) 'Superwetable Microchips as a Platform toward Microgravity Biosensing', *ACS Nano*, 11(1), pp. 621–626. doi: 10.1021/ACSNANO.6B06896.

Ye, H. *et al.* (2011) 'A synthetic optogenetic transcription device enhances blood-

Bibliography

glucose homeostasis in mice’, *Science*, 332(6037), pp. 1565–1568. doi: 10.1126/science.1203535.

Yokobayashi, Y., Weiss, R. and Arnold, F. H. (2002) ‘Directed evolution of a genetic circuit’, *Proceedings of the National Academy of Sciences*, 99(26), pp. 16587 LP – 16591. doi: 10.1073/pnas.252535999.

ZARM: General Information (no date). Available at: <https://www.zarm.uni-bremen.de/en/drop-tower/general-information.html> (Accessed: 17 June 2021).

ZEISS Microscopy Online Campus | Live-Cell Imaging | Microscopy Techniques (no date). Available at: <http://zeiss-campus.magnet.fsu.edu/articles/livecellimaging/techniques.html> (Accessed: 10 May 2021).

Zhao, Y. *et al.* (2016) ‘Visualized and precise design of artificial small RNAs for regulating T7 RNA polymerase and enhancing recombinant protein folding in *Escherichia coli*’, *Synthetic and Systems Biotechnology*, 1(4), pp. 265–270. doi: 10.1016/J.SYNBIO.2016.08.005.

Zheng, Z. *et al.* (2015) ‘Dynamical and Microrheological Analysis of Amyloplasts in the Plant Root Gravity-Sensing Cells’, *Microgravity Science and Technology*, 27(6), pp. 485–493. doi: 10.1007/s12217-015-9445-x.

APPENDIX

Table A.1 Short description of common gene sets

Gene Set	Description
Oncogenic Signature	
KRAS.300_UP.V1_UP	Genes up-regulated in four lineages of epithelial cell lines over-expressing an oncogenic form of KRAS [Gene ID=3845] gene.
KRAS.600_UP.V1_UP	Genes up-regulated in four lineages of epithelial cell lines over-expressing an oncogenic form of KRAS [Gene ID=3845] gene.
CAHOY_ASTROCYTIC	Genes up-regulated in astrocytes.
IL15_UP.V1_UP	Genes up-regulated in Sez-4 cells (T lymphocyte) that were first starved of IL2 [Gene ID=3558] and then stimulated with IL15 [Gene ID=3600].
IL2_UP.V1_UP	Genes up-regulated in Sez-4 cells (T lymphocyte) that were first starved of IL2 [Gene ID=3558] and then stimulated with IL2 [Gene ID=3558].
Canonical Pathway	
REACTOME_OLFACTORY_SIGNALING_PATHWAY	Genes involved in Olfactory Signaling Pathway
KEGG_OLFACTORY_TRANSDUCTION	Olfactory transduction
REACTOME_SLC_MEDIATED_TRANSMEMBRANE_TRANSPORT	Genes involved in SLC-mediated transmembrane transport
REACTOME_MEIOSIS	Genes involved in Meiosis
REACTOME_MEIOTIC_RECOMBINATION	Genes involved in Meiotic Recombination

Appendix

REACTOME_GENERATION_OF_SECOND_MESSENGER_MOLECULES	Genes involved in Generation of second messenger molecules
KEGG_ALLOGRAFT_REJECTION	Allograft rejection
KEGG_AUTOIMMUNE_THYROID_DISEASE	Autoimmune thyroid disease
KEGG_ASTHMA	Asthma
KEGG_HEMATOPOIETIC_CELL_LINEAGE	Hematopoietic cell lineage
KEGG_TYPE_I_DIABETES_MELLITUS	Type I diabetes mellitus
KEGG_GRAFT_VERSUS_HOST_DISEASE	Graft-versus-host disease
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION	Intestinal immune network for IgA production
PID_NFAT_TFPATHWAY	Calcineurin-regulated NFAT-dependent transcription in lymphocytes
BIOCARTA_CYTOKINE_PATHWAY	Cytokine Network
BIOCARTA_INFLAM_PATHWAY	Cytokines and Inflammatory Response
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	Cytokine-cytokine receptor interaction
BIOCARTA_NKT_PATHWAY	Selective expression of chemokine receptors during T-cell polarization

Appendix

REACTOME_RESPIRATORY_ELECTRON_TRANSPORT_ATP_SYNTHESIS_BY_CHEMIOSMOTIC_COUPLING_AND_HEAT_PRODUCTION_BY_UNCOUPLING_PROTEINS_	Genes involved in Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins.
REACTOME_TCA_CYCLE_AND_RESPIRATORY_ELECTRON_TRANSPORT	Genes involved in The citric acid (TCA) cycle and respiratory electron transport
KEGG_LEISHMANIA_INFECTION	Leishmania infection
REACTOME_PD1_SIGNALING	Genes involved in PD-1 signaling
Cancer modules	
MODULE_47	Genes in the cancer module 47
MODULE_75	Genes in the cancer module 75
MODULE_46	Genes in the cancer module 75
MODULE_6	Genes in the cancer module 6
MODULE_123	Genes in the cancer module 123
MODULE_114	Genes in the cancer module 114
MODULE_151	Genes in the cancer module 151
MODULE_22	Genes in the cancer module 22
Chemical and Genetic perturbation	
RICKMAN_HEAD_AND_NECK_CANCER_C	Cluster c: genes identifying an intrinsic group in head and neck squamous cell carcinoma (HNSCC).

Appendix

DOANE_BREAST_CANCER_ESR1_DN	Genes down-regulated in breast cancer samples positive for ESR1 [GeneID=2099] compared to the ESR1 negative tumors.
BOSCO_EPITHELIAL_DIFFERENTIATION_MODULE	Genes representing epithelial differentiation module in sputum during asthma exacerbations.
SENGUPTA_NASOPHARYNGEAL_CARCI-NOMA_WITH_LMP1_DN	Genes down-regulated in nasopharyngeal carcinoma (NPC) positive for LMP1 [GeneID=9260], a latent gene of Epstein-Barr virus (EBV).
ROLEF_GLIS3_TARGETS	Genes down-regulated in nasopharyngeal carcinoma (NPC) positive for LMP1 [GeneID=9260], a latent gene of Epstein-Barr virus (EBV).
LU_EZH2_TARGETS_UP	Genes up-regulated in SKOV3ip1 cells (ovarian cancer) upon knockdown of EZH2 [GeneID=2146] by RNAi.
ALTEMEIER_RESPONSE_TO_LPS_WITH_MECHANICAL_VENTILATION	Genes up-regulated in lung tissue upon LPS aspiration with mechanical ventilation (MV) compared to control (PBS aspiration without MV).
MARTENS_BOUND_BY_PML_RARA_FUSION	Genes with promoters occupied by PML-RARA fusion [GeneID=5371,5914] protein in acute promyelocytic leukemia (APL) cells NB4 and two APL primary blasts, based on Chip-seq data.
GALINDO_IMMUNE_RESPONSE_TO_ENTEROTOXIN	Genes up-regulated in macrophages by aerolysin-related cytotoxic enterotoxin (Act) from <i>Aeromonas hydrophila</i> .
KUROZUMI_RESPONSE_TO_ONCOCYTIC_VIRUS	Inflammatory cytokines and their receptors modulated in brain tumors after treatment with an oncocytic virus, a potential anticancer therapy.
VILIMAS_NOTCH1_TARGETS_UP	Genes up-regulated in bone marrow progenitors by constitutively active NOTCH1 [GeneID=4851].
DAIRKEE_TERT_TARGETS_UP	Genes up-regulated in non-spontaneously immortalizing (NSI) primary breast cancer tumor cultures upon expression of TERT [GeneID=7015] off a retroviral vector.
ZHAN_V1_LATE_DIFFERENTIATION_GENES_DN	The v1LDG down-regulated set: most variable late differentiation genes (LDG) with similar expression patterns in tonsil plasma cells (TPC) and multiple myeloma (MM) samples.
LI_DCP2_BOUND_MRNA	Genes encoding mRNA transcripts specifically bound by DCP2 [GeneID=167227].
QI_PLASMACYTOMA_UP	Up-regulated genes that best discriminate plasmablastic plasmacytoma from plasmacytic plasmacytoma tumors.

Appendix

GESERICK_TERT_TARGETS_DN	Genes down-regulated in MEF cells (embryonic fibroblasts) with TERT [GeneID=7015] knock-out, after expression of the gene off a retroviral vector.
AMIT_SERUM_RESPONSE_120_MCF10A	Genes whose expression peaked at 120 min after stimulation of MCF10A cells with serum.
SMIRNOV_CIRCULATING_ENDOTHELIO-CYTES_IN_CANCER_UP	Genes up-regulated in circulating endothelial cells (CEC) from cancer patients compared to those from healthy donors.
KIM_ALL_DISORDERS_DURATION_CORR_DN	Genes whose expression in brain significantly and negatively correlated with the duration of all psychiatric disorders studied.
PHONG_TNF_TARGETS_UP	Genes up-regulated in Calu-6 cells (lung cancer) at 1 h time point after TNF [GeneID=7124] treatment.
BROCKE_APOPTOSIS_REVERSED_BY_IL6	Genes changed in INA-6 cells (multiple myeloma, MM) by re-addition of IL6 [GeneID=3569] after its initial withdrawal for 12h.
BASSO_CD40_SIGNALING_UP	Gene up-regulated by CD40 [GeneID=958] signaling in Ramos cells (EBV negative Burkitt lymphoma).
TIAN_TNF_SIGNALING_NOT_VIA_NFKB	Genes modulated in HeLa cells (cervical carcinoma) by TNF [GeneID=7124] not via NFKB pathway.
SEKI_INFLAMMATORY_RESPONSE_LPS_UP	Genes up-regulated in hepatic stellate cells after stimulation with bacterial lipopolysaccharide (LPS).
DIRMEIER_LMP1_RESPONSE_EARLY	Clusters 1 and 2: genes up-regulated in B2264-19/3 cells (primary B lymphocytes) within 30-60 min after activation of LMP1 (an oncogene encoded by Epstein-Barr virus, EBV).
AMIT_DELAYED_EARLY_GENES	Delayed early genes (DEG) which are coordinately down-regulated in multiple epithelial tumor types.
OSWALD_HEMATOPOIETIC_STEM_CELL_IN_COLLAGEN_GEL_UP	Genes up-regulated in hematopoietic stem cells (HSC, CD34+ [GeneID=947]) cultured in a three-dimensional collagen gel compared to the cells grown in suspension.
KOBAYASHI_EGFR_SIGNALING_6HR_DN	Genes down-regulated in H1975 cells (non-small cell lung cancer, NSCLC) resistant to gefitinib [PubChem=123631] after treatment with EGFR inhibitor CL-387785 [PubChem=2776] for 6h.
RODWELL_AGING_KIDNEY_UP	Genes whose expression increases with age in normal kidney.

Appendix

STARK_PREFRONTAL_CORTEX_22Q11_DELETION_DN	Genes down-regulated in prefrontal cortex (PFC) of mice carrying a hemizygotic microdeletion in the 22q11.2 region.
MOOTHA_VOXPPOS	Genes involved in oxidative phosphorylation; based on literature and sequence annotation resources and converted to Affymetrix HG-U133A probe sets.
YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_17	Genes co-regulated in uterus during a time course response to progesterone [PubChem=5994]: SOM cluster 17.
GOLDRATH_ANTIGEN_RESPONSE	Genes up-regulated at the peak of an antigen response of naive CD8+ [GeneID=925;926] T-cells.
CHUNG_BLISTER_CYTOTOXICITY_UP	Genes up-regulated in blister cells from patients with adverse drug reactions (ADR).
MORI_MATURE_B_LYMPHOCYTE_UP	Up-regulated genes in the B lymphocyte developmental signature, based on expression profiling of lymphomas from the Emu-myc transgenic mice: the mature B
HSIAO_HOUSEKEEPING_GENES	Housekeeping genes identified as expressed across 19 normal tissues.
WONG_MITOCHONDRIA_GENE_MODULE	Genes that comprise the mitochondria gene module
YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_13	Genes co-regulated in uterus during a time course response to progesterone [PubChem=5994]: SOM cluster 13.
GAVIN_FOXP3_TARGETS_CLUSTER_P4	Cluster P4 of genes with similar expression profiles in peripheral T lymphocytes after FOXP3 [GeneID=50943] loss of function (LOF).
BECKER_TAMOXIFEN_RESISTANCE_UP	Genes up-regulated in a breast cancer cell line resistant to tamoxifen [PubChem=5376] compared to the parental line sensitive to the drug.
WIELAND_UP_BY_HBV_INFECTION	Genes induced in the liver during hepatitis B (HBV) viral clearance in chimpanzees.
RUTELLA_RESPONSE_TO_HGF_DN	Genes down-regulated in peripheral blood monocytes by HGF [GeneID=3082]
Cancer Gene neighbourhood	
GNF2_IL2RB	Neighborhood of IL2RB

Appendix

GSE9988_LPS_VS_LPS_AND_ANTI_TREM1_MONO-CYTE_DN	Genes down-regulated in comparison of monocytes treated with 5000 ng/ml LPS (TLR4 agonist) versus monocytes treated with anti-TREM1 [GeneID=54210].
GSE9006_HEALTHY_VS_TYPE_2_DIABETES_PBMAT_DX_UP	Genes up-regulated in comparison of peripheral blood mononuclear cells (PBMC) from healthy donors versus PBMCs from patients with type 2 diabetes at the time of diagnosis.
GSE9988_LPS_VS_VEHICLE_TREATED_MONO-CYTE_UP	Genes up-regulated in comparison of monocytes treated with 1 ng/ml LPS (TLR4 agonist) versus monocytes treated with vehicle.
GSE2706_UNSTIM_VS_2H_LPS_DC_DN	Genes down-regulated in comparison of unstimulated dendritic cells (DC) at 0 h versus DCs stimulated with LPS (TLR4 agonist) for 2 h.
GSE9988_LOW_LPS_VS_VEHICLE_TREATED_MONO-CYTE_UP	Genes up-regulated in comparison of monocytes treated with 1 ng/ml LPS (TLR4 agonist) versus monocytes treated with control IgG.
GSE37416_CTRL_VS_12H_F_TULARENSIS_LVS_NEUTROPHIL_DN	Genes down-regulated in comparison of control polymorphonuclear leukocytes (PMN) at 12 h versus PMN treated with F. tularensis vaccine at 12 h.
GSE9988_LOW_LPS_VS_CTRL_TREATED_MONO-CYTE_UP	Genes up-regulated in comparison of monocytes treated with 1 ng/ml LPS (TLR4 agonist) versus monocytes treated with control IgG.
GSE2706_UNSTIM_VS_2H_R848_DC_DN	Genes down-regulated in comparison of unstimulated dendritic cells (DC) at 0 h versus DCs stimulated with R848 for 2 h.
GSE14769_UNSTIM_VS_40MIN_LPS_BMDM_DN	Genes down-regulated in comparison of unstimulated macrophage cells versus macrophage cells stimulated with LPS (TLR4 agonist) for 40 min.
GSE9988_LPS_VS_CTRL_TREATED_MONOCYTE_UP	Genes up-regulated in comparison of monocytes treated with 5000 ng/ml LPS (TLR4 agonist) versus monocytes treated with control IgG.
GSE22886_CD4_TCELL_VS_BCELL_NAIVE_UP	Genes up-regulated in comparison of naïve CD4 [GeneID=920] T cells versus naïve B cells.
GSE2706_UNSTIM_VS_2H_LPS_AND_R848_DC_DN	Genes down-regulated in comparison of unstimulated dendritic cells (DC) at 0 h versus DCs stimulated with LPS (TLR4 agonist) and R848 for 2 h.

Appendix

GSE9988_ANTI_TREM1_VS_CTRL_TREATED_MONOCYTES_UP	Genes up-regulated in comparison of monocytes treated with anti-TREM1 [GeneID=54210] versus monocytes treated with control IgG.
GSE22886_NEUTROPHIL_VS_MONOCYTE_DN	Genes down-regulated in comparison of neutrophils versus monocytes.
GSE360_L_DONOVANI_VS_B_MALAYI_HIGH_DOSE_MAC_DN	Genes down-regulated in comparison of macrophages exposed to <i>L. donovani</i> versus macrophages exposed to 50 worms/well <i>B. malayi</i> .
GSE17721_0.5H_VS_4H_CPG_BMDM_UP	Genes up-regulated in comparison of dendritic cells (DC) stimulated with CpG DNA (TLR9 agonist) at 0.5 h versus those stimulated with CpG DNA (TLR9 agonist) at 4 h.
GSE28237_FOLLICULAR_VS_LATE_GC_BCELL_DN	Genes down-regulated in comparison of follicular B cells versus late germinal center (GC) B cells.
GSE29618_BCELL_VS_MDC_DN	Genes down-regulated in comparison of B cells versus myeloid dendritic cells (mDC).
GSE29618_PDC_VS_MDC_DAY7_FLU_VACCINE_DN	Genes down-regulated in comparison of plasmacytoid dendritic cells (DC) from influenza vaccinee at day 7 post-vaccination versus myeloid DCs at day 7 post-vaccination.
GSE17580_TREG_VS_TEFF_S_MANSONI_INF_UP	Genes up-regulated in comparison of regulatory T cell (Treg) from mice infected with <i>S. mansoni</i> versus T effector cells from the infected mice.
GSE1460_DP_THYMOCYTE_VS_NAIVE_CD4_TCELL_ADULT_BLOOD_UP	Genes up-regulated in comparison of CD4 [GeneID=920] CD8 thymocytes versus naive CD4 [GeneID=920] T cells from adult blood.
GSE11057_NAIVE_CD4_VS_PBMC_CD4_TCELL_DN	Genes down-regulated in comparison of naive T cells versus peripheral blood mononuclear cells (PBMC).
Hallmark gene set	
HALLMARK_TNFA_SIGNALLING_VIA_NFKB	Genes regulated by NF- κ B in response to TNF [GeneID=7124].
HALLMARK_OXIDATIVE_PHOSPHORYLATION	Genes encoding proteins involved in oxidative phosphorylation.
Motif gene sets	
Positional gene sets	

Appendix

[illegible]

Appendix

		Upregulated Pathways									
4209, 43582, 57418, 38836	1	RICKMAN_HEAD_AND_NECK_CANCER_C	0.01847826	0.75015676	<0.001	0.017004	<0.001	0.03	0.038526	0.19121702	This study only
38836, 57418, 4209	4	DOANE_BREAST_CANCER_ESR1_DN	0.0019084	0.00661987	0.008529	0.064161	0.00191	0.00662			This study only
		BOSCO_EPITHELIAL_DIFFERENTIATION_MODULE	<0.001	0.00244102	0.015945	0.102752	<0.001	0.00244			This study only
		SENGUPTA_NASOPHARYNGEAL_CARCI-NOMA_WITH_LMP1_DN	<0.001	0.06067855	0.002375	0.111207	<0.001	0.06068			This study only
		ROLEF_GLIS3_TARGETS	0.02509653	0.10299574	0.027586	0.142122	0.0251	0.103			This study only
		Downregulated Pathways									
38836, 43582, 4209	23	LU_EZH2_TARGETS_UP	<0.001	0.01079626			<0.001	0.0108	0.015625	0.34704456	This study only
		ALTEMEIER_RESPONSE_TO_LPS_WITH_MECHANICAL_VENTILATION	<0.001	0.00809729			<0.001	0.0081	0.00266	0.1373343	This study only
		MARTENS_BOUND_BY_PML_RARA_FUSION	<0.001	0.07663617			<0.001	0.07664	<0.001	0.30797347	This study only
		GALINDO_IMMUNE_RESPONSE_TO_ENTEROTOXIN	0.01642711	0.07373151			0.01643	0.07373	0.002695	0.08092894	This study only
		KUROZUMI_RESPONSE_TO_ONCOCYTIC_VIRUS	<0.001	0.0019607			<0.001	0.00196	0.018433	0.17784968	This study only
		VILIMAS_NOTCH1_TARGETS_UP	<0.001	<0.001			<0.001	<0.001	0.011521	0.12423761	This study only
		DAIRKEE_TERT_TARGETS_UP	0.00767754	0.11923367			0.00768	0.11923	0.006711	0.31035018	This study only
		ZHAN_V1_LATE_DIFFERENTIATION_GENES_DN	0.02625821	0.0233809			0.02626	0.02338	0.046185	0.1519671	This study only
		LI_DCP2_BOUND_MRNA	<0.001	<0.001			<0.001	<0.001	0.042714	0.2685578	This study only
		QI_PLASMACYTOMA_UP	<0.001	<0.001			<0.001	<0.001	0.012308	0.29097813	This study only
		GESERICK_TERT_TARGETS_DN	0.00662252	0.01780009			0.00662	0.0178	<0.001	0.01704472	This study only

Appendix

		AMIT_SERUM_RESPONSE_120_MCF10A	0.00803 213	0.047 35798			0.00803	0.04736	0.00485 4	0.13040 653	This study only
		SMIRNOV_CIRCULATING_ENDOTHELIO- CYTES_IN_CANCER_UP	0.01750 973	0.118 82599			0.01751	0.11883	0.01169 6	0.20567 222	This study only
		KIM_ALL_DISORDERS_DURATION_CORR_DN	0.00204 082	0.026 47004			0.00204	0.02647	<0.001	0.13470 972	This study only
		PHONG_TNF_TARGETS_UP	0.00414 938	0.016 29646			0.00415	0.0163	<0.001	<0.001	
		BROCKE_APOPTOSIS_REVERSED_BY_IL6	0.00804 829	0.060 23157			0.00805	0.06023	0.02216 1	0.26465 12	E_GEOD_38836
											E-GEOD-43582
											EGEOD_4209
		BASSO_CD40_SIGNALING_UP	<0.001	<0.00 1			<0.001	<0.001	0.00777 2	0.12496 531	This study only
		TIAN_TNF_SIGNALING_NOT_VIA_NFKB	0.01476 793	0.023 17965			0.01477	0.02318	0.01307 2	0.05104 666	This study only
		SEKI_INFLAMMATORY_RESPONSE_LPS_UP	<0.001	<0.00 1			<0.001	<0.001	0.01658 8	0.18592 924	This study only
		DIRMEIER_LMP1_RESPONSE_EARLY	<0.001	0.001 00827			<0.001	0.00101	<0.001	<0.001	This study only
		AMIT_DELAYED_EARLY_GENES	0.02380 952	0.055 82707			0.02381	0.05583	<0.001	0.02214 7698	This study only
		OSWALD_HEMATOPOI- ETIC_STEM_CELL_IN_COLLAGEN_GEL_UP	0.03625 954	0.165 77902			0.03626	0.16578	0.02013 4	0.32007 053	This study only
		KOBAYASHI_EGFR_SIGNALING_6HR_DN	0.02844 639	0.036 91696			0.02845	0.03692	0.02136 8	0.08795 31	This study only
38836, 57418, 4209	14	RODWELL_AGING_KIDNEY_UP	0.00187 266	0.100 17705	<0.00 1	0.126 329	0.00187	0.10018			This study only
		STARK_PREFRONTAL_CORTEX_22Q11_DELE- TION_DN	<0.001	0.001 14437	0.002 907	0.176 888	<0.001	0.00114			This study only
		MOOTHA_VOXPPOS	<0.001	<0.00 1	<0.00 1	0.024 884	<0.001	<0.001			This study only
		YAO_TEMPORAL_RESPONSE_TO_PROGESTER- ONE_CLUSTER_17	<0.001	<0.00 1	0.004 808	0.192 914	<0.001	<0.001			This study only

Appendix

		GOLDRATH_ANTIGEN_RESPONSE	<0.001	<0.001	<0.001	0.163859	<0.001	<0.001			This study only
		CHUNG_BLISTER_CYTOTOXICITY_UP	<0.001	<0.001	0.018663	0.211501	<0.001	<0.001			This study only
		MORI_MATURE_B_LYMPHOCYTE_UP	<0.001	0.00187997	<0.001	0.040834	<0.001	0.00188			This study only
		HSIAO_HOUSEKEEPING_GENES	<0.001	0.05013751	0.034483	0.353601	<0.001	0.05014			This study only
		WONG_MITOCHONDRIA_GENE_MODULE	<0.001	<0.001	0.001587	0.108873	<0.001	<0.001			This study only
		YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_13	<0.001	<0.001	0.031397	0.264109	<0.001	<0.001			This study only
		GAVIN_FOXP3_TARGETS_CLUSTER_P4	0.00406504	0.02379203	0.028099	0.223921	0.00407	0.02379			This study only
		BECKER_TAMOXIFEN_RESISTANCE_UP	0.02240326	0.06684875	0.017825	0.133775	0.0224	0.06685			This study only
		WIELAND_UP_BY_HBV_INFECTION	<0.001	<0.001	0.001701	0.038188	<0.001	<0.001			This study only
		RUTELLA_RESPONSE_TO_HGF_DN	<0.001	0.00648465	0.007692	0.205643	<0.001	0.00648			This study only
		Cancer Gene neighbourhood									
		Downregulated Pathways									
38836, 43582, 4209	1	GNF2_IL2RB	0.00540541	0.05146252			<0.001	<0.001	0.03125	0.10910398	This study only
		Immunologic Signature									
		Downregulated Pathways									
57418, 43582, 38836	1	GSE9988_LPS_VS_LPS_AND_ANTI_TREM1_MONOCYTE_DN	<0.001	0.01122794	<0.001	<0.001			0.034483	0.2374903	This study only
38836, 43582, 4209	12	GSE9006_HEALTHY_VS_TYPE_2_DIABETES_PBMC_AT_DX_UP	<0.001	0.01405743			<0.001	<0.001	<0.001	0.048687838	This study only

Appendix

		GSE9988_LPS_VS_VEHICLE_TREATED_MONO-CYTE_UP	<0.001	<0.001			0.00204	<0.001	0.002778	0.048585434	This study only
		GSE2706_UNSTIM_VS_2H_LPS_DC_DN	<0.001	<0.001			<0.001	<0.001	0.003155	0.011890802	This study only
		GSE9988_LOW_LPS_VS_VEHICLE_TREATED_MONOCYTE_UP	<0.001	<0.001			<0.001	<0.001	<0.001	0.056588795	This study only
		GSE37416_CTRL_VS_12H_F_TULARENSIS_LVS_NEUTROPHIL_DN	<0.001	0.00401273			<0.001	0.00332	<0.001	0.004799189	This study only
		GSE9988_LOW_LPS_VS_CTRL_TREATED_MONOCYTE_UP	<0.001	<0.001			<0.001	<0.001	0.014493	0.08334377	This study only
		GSE2706_UNSTIM_VS_2H_R848_DC_DN	<0.001	<0.001			<0.001	<0.001	0.011204	0.08299657	This study only
		GSE14769_UNSTIM_VS_40MIN_LPS_BMDM_DN	<0.001	<0.001			0.01247	0.02814	<0.001	0.047810346	This study only
		GSE9988_LPS_VS_CTRL_TREATED_MONOCYTE_UP	<0.001	<0.001			<0.001	<0.001	0.008451	0.06785608	This study only
		GSE22886_CD4_TCELL_VS_BCELL_NAIVE_UP	<0.001	<0.001			0.00185	0.00119	0.013587	0.09954387	This study only
		GSE2706_UNSTIM_VS_2H_LPS_AND_R848_DC_DN	<0.001	<0.001			<0.001	<0.001	<0.001	0.048592165	This study only
		GSE9988_ANTI_TREM1_VS_CTRL_TREATED_MONOCYTES_UP	<0.001	<0.001			0.01186	0.0202	0.020833	0.110544026	This study only
38836, 57418, 4209	9	GSE22886_NEUTROPHIL_VS_MONOCYTE_DN	<0.001	0.00189392	<0.001	<0.001	<0.001	<0.001			This study only
		GSE360_L_DONOVANI_VS_B_MALAYI_HIGH_DOSE_MAC_DN	<0.001	0.05330102	<0.001	<0.001	<0.001	<0.001			This study only
		GSE17721_0.5H_VS_4H_CPG_BMDM_UP	0.03125	0.14473116	<0.001	<0.001	<0.001	<0.001			This study only
		GSE28237_FOLLICULAR_VS_LATE_GC_BCELL_DN	<0.001	0.05972375	0.024961	0.102298	<0.001	<0.001			This study only
		GSE29618_BCELL_VS_MDC_DN	<0.001	0.02312637	<0.001	<0.001	<0.001	<0.001			This study only

Appendix

Table A.3. Distribution of leading edge genes in NMFC clusters

E-GEOD-38836 (Upregulated)		E-GEOD-38836 (Downregulated)		E-GEOD-43582 (Upregulated)		E-GEOD-43582 (downregulated)		E-GEOD-57418 (Upregulated)		E-GEOD-57418 (Down-regulated)		E-GEOD-4209 (Up-regulated)		E-GEOD-4209 (Down-regulated)	
Cluster	Member genes	Cluster	Member genes	Cluster	Member genes	Cluster	Member genes	Cluster	Member genes	Cluster	Member genes	Cluster	Member genes	Cluster	Member genes
1	DHRS9	1	EEF1A1	1	OR11H1	1	DUSP2	1	PTPRC	1	CD9	1	SLC4A4	1	CCNC
1	C19orf40	1	SOCS3	1	CDK2	1	TNNI3	1	GADD45A	1	FOS	1	APP	1	PRPF3
1	UGT2B17	1	CDKN2C	1	SLC22A4	1	HCRT	1	TNFSF4	1	LGALS3	1	ITPR2	1	ADAR
1	OPN3	1	NFATC1	1	KIF2A	1	VEGFB	1	RNPS1	1	PSAP	1	LYZ	1	U2AF1
1	VEGFA	1	TAP2	1	PDE3A	1	TNFRSF14	1	EXOC2	1	SDC3	1	SPTBN1	1	INPP5D
1	TAAR9	1	FCGR1A	1	ZNF610	1	GRM7	1	IRS1	1	LIPA	1	ALDH5A1	1	RABGGTA
1	PIPOX	1	IFNA2	1	NUP153	1	MAPK8IP3	1	PSMB2	1	APOE	1	CCR9	1	ICAM1
1	ADRA2B	1	CCL4	1	OR2H1	1	MADCAM1	1	PSMA5	1	IL1R2	1	ABLIM1	1	UBE4A
1	CNDP1	1	TNFSF9	1	AGPS	1	MAPT	1	SRSF10	1	RPS6KA2	1	LEF1	1	POLE
1	ABCD1	1	IL5RA	1	SLC12A6	1	SGCA	1	ZNF566	1	FN1	1	ACTN1	1	SEC24B
1	COL3A1	1	PDGFRB	1	KLHL20	1	STX1A	1	POLR3A	1	PPARG	1	COL6A2	1	RFC4
1	OR5L2	1	AKAP5	1	RAD54L	1	CXCL12	1	BMPR1A	1	ARHGDIG	1	HMOX1	1	CENPM
1	ATP1A2	1	KIF20A	1	TDP1	1	EGR4	1	ZNF596	1	SNCA	1	HIST1H2A C	1	PFDN1
1	AKR1D1	1	TAF1	1	CDC14B	1	NMB	1	GADD45B	1	APOC2	1	HIST1H2B D	1	PIGA
1	CYSLTR2	1	KIR2DS3	1	ZNF559	1	RAMP1	1	KLC1	1	VWF	1	S100A9	1	PARP3
1	ADM	1	IL17A	1	OIP5	1	CACNG8	1	TYK2	1	HLA.DRB4	1	CLDN9	1	BCL2L1
1	ATP8A2	1	CD72	1	DLAT	1	MMP15	1	ADIPOR1	1	SMO	1	COL4A5	1	PELI1
1	SNCA	1	CDC20	1	NARS	1	COX6A2	1	E2F4	1	A2M	1	MBL2	1	IDH3B
1	SULT1E1	1	HIST1H4G	1	SMC1B	1	GABRE	1	PSMD2	1	LPL	1	RXFP3	1	TDRD7
1	SLC26A7	1	E2F5	1	SLC35C1	1	CNR1	1	RAD9A	1	CD1C	1	CXCL6	1	SRM
1	GABRB3	1	PDCD1LG2	1	KIAA1598	1	PLA2G6	1	ZNF19	1	CHIT1	1	FOXN1	1	CYB5R3
1	PTK2B	1	ACTN4	1	LDHB	1	MUC6	1	IL7R	1	IL1R1	1	PLA2G6	1	PSMD7

Appendix

1	CHRNE	1	IFNA4	1	FANCF	1	LPAR2	1	ZNF420	1	CD68	1	GLRB	1	ANAPC5
1	S1PR3	1	NPPA	1	ZNF221	1	PROC	1	DVL3	1	HLA.DMB	1	PHKG1	1	CD81
1	ABAT	1	VIP	1	UGDH	1	PPP3R2	1	MAP4K4	2	CD40LG	1	OMD	1	PSMC2
1	KCNMB2	1	SPRY2	1	NEK2	1	PTPRN	1	TGS1	2	MADCAM1	1	CLEC4E	1	RCAN2
1	FAN1	1	FASLG	1	PTPN11	1	CD70	1	EIF2AK2	2	ARPC4	1	CRISPLD2	1	CCND3
1	ABCC9	1	HIST3H3	1	DDB1	1	GP1BB	1	PHLPP1	2	ARHGEF11	1	SLC7A2	1	CS
1	AGTR1	1	CTF1	1	PTPN21	1	GRIN2D	1	AKT1	2	SUCLG2	1	MLN	1	ADD1
1	SLC30A2	1	ELANE	1	AMACR	1	THPO	1	ABL1	2	PCK2	1	CHRD1	1	UBA3
1	ATP11B	1	CD33	1	ALDH3A2	1	CX3CL1	1	ZNF584	2	ACAT1	1	MUC4	1	FPGS
1	ATP2B2	1	SPRED1	1	GIT2	1	ATP1A2	1	CENPL	2	SMPD3	1	LAMC2	1	TUBB4B
1	GAD2	1	S100A2	1	CASP2	1	GRIK4	1	NRBP1	2	MEF2A	1	MUC6	1	NT5C
1	KNG1	1	LIN9	1	DOCK1	1	EPOR	1	CCNT2	2	PIK3R1	1	PKP2	1	IRF1
1	CCKAR	1	IFNG	1	MCM3	1	NR0B1	1	HMGCS1	2	P2RY1	1	SLC7A10	1	NDUFA9
1	SLC35D2	1	IFNA1	1	UBE2D1	1	ALOX5	1	CYP51A1	2	UNC93B1	1	ADAM11	1	PTTG1
1	CAMK2B	1	ATP5A1	1	E2F7	1	ARR3	1	RANGAP1	2	VAMP8	1	PARD3	1	RALA
1	SSTR1	1	CDK5	1	OR2M4	1	KIR3DL1	1	CYP1A1	2	PCCB	1	SLC8A2	1	POLD1
1	SLC28A2	1	CCND1	1	CBX5	1	GRM4	1	ZNF189	2	MBOAT1	1	OR5I1	1	PSMD14
1	AKR1C1	1	S100B	1	TRIP12	1	INHBB	1	HIF1AN	2	MCEE	1	COL4A6	1	TOLLIP
1	AKR1C2	1	ACTA1	1	UBE2G2	1	TNFRSF18	1	PARP14	2	PLA2G12A	1	OR10C1	1	NME7
1	EPB41L1	1	HIST1H4J	1	GHR	1	PRKACG	1	ARAF	2	ABAT	1	HIST1H2BJ	1	CASP8
1	SLC1A3	1	NDUFA3	1	ADCY3	1	GUCY2D	1	BCL2L11	2	GRIN2C	1	CFH	1	SF3B2
1	SYK	1	BDKRB2	1	PSMD12	1	PRLHR	1	ZNF141	2	COX8A	1	PVRL1	1	CDK9
1	CALCR	1	MME	1	PRPF8	1	DNAJC5	1	ZNF230	2	COX6C	1	NR1H3	1	NCAPD2
1	CHDH	2	RPL31	1	AKT1	1	IL9	1	HNRNPA2B1	2	NET1	1	CSTA	1	NDUFV1
1	MAOB	2	TCEB1	1	OR6T1	1	PIK3CD	1	POLE	2	PRR5.ARHGAP8	1	ANGPT2	1	TSG101
1	PLA2G5	2	ARPC1B	1	GCH1	1	GNAO1	1	UNG	2	PARK7	1	CCR8	1	CYC1
1	NPFF	2	S100A12	1	PGM2	1	RAMP2	1	ZNF468	2	SDHD	1	GH2	1	PPP2R5E
1	LAMB1	2	CHP2	1	OR11H6	1	GRIK3	1	NUPL1	2	NDUFA8	1	SCN2B	1	TRIP12

Appendix

1	CTNNA1	2	C7	1	HGSNAT	1	ELANE	1	ZNF614	2	MDH1	1	CHRM3	1	RHOA
1	SYN1	2	CTLA4	1	TUBA1C	1	CCKBR	1	PSMB3	2	SDHB	1	HIST1H2AE	1	SNRNP40
1	OR2L2	2	NUP35	1	TAF6	1	DUSP4	1	SRSF9	2	ATP5I	1	UGT1A3	1	CD3D
1	COL2A1	2	ZNRD1	1	AP1M1	1	NPFFR1	1	RASA1	2	STX10	1	ADAMTS5	1	SPCS1
1	GRIN3B	2	TAF12	1	SMARCA2	1	TNNI1	1	ATF1	2	DGKH	1	CHRM4	1	DHX16
1	HMMR	2	MAFF	1	PDK2	1	BLK	1	PCBP1	2	TPM1	1	PLXNB3	1	PMPCA
2	HNMT	2	HRAS	1	SLC23A2	1	TACR2	1	MAGI2	2	NDUFA2	1	SERPINB7	1	CCR10
2	SLC5A7	2	SRP19	1	SLCO3A1	1	ADRA2C	1	CDK9	2	PDHA1	1	ITGA2	1	GPAA1
2	G6PC2	2	ARPC3	1	OR13C9	1	CHRNA2	1	ZNF140	2	ABCG1	1	COL13A1	1	PTP4A1
2	KCNA6	2	CD3D	1	PIGG	1	KCNK17	1	MED17	2	ACADM	1	COL9A3	1	TAF10
2	SLC4A3	2	RBL2	1	ZNF773	1	TNFRSF4	1	RAC3	2	ESRRA	1	HCN4	1	GNAI3
2	GAST	2	EIF3D	1	PRPF4	1	ADRA2A	1	NUDC	2	NDUFS3	1	MMP3	1	LDHA
2	SLC30A8	2	TAF10	1	CKAP5	1	GABRB3	1	SLC30A5	2	AP1B1	1	SCUBE2	1	HARS
2	ADCYAP1	2	ATP5J	1	SNRNP200	1	GRIP2	1	EIF4G1	2	P4HB	1	CLDN16	1	RPL23
2	PPYR1	2	SF3B5	1	DARS	1	CD80	1	RNMT	2	PDK3	1	PLBD1	1	GCNT1
2	CTH	2	HNRNPA3	1	MAPRE1	1	KCNS2	1	CCT5	2	ARSG	1	GABRR1	1	ID2
2	KCNJ5	2	COX6A1	1	NF2	1	IL13	1	CSTF3	2	AP2A2	1	ALDH2	1	ERCC1
2	GRIP2	2	ATP5L	1	RPL23	1	MUC16	1	ILK	2	CTSA	1	CALML5	1	TUBA1C
2	HES1	2	PTPN7	1	MNAT1	1	ATP2A1	1	DCTN1	2	PLCB1	1	MET	1	CNOT1
2	FGB	2	IL22	1	CCT4	1	CYP2C8	1	EXOC4	2	RALBP1	1	KCND2	1	DDX23
2	SLC44A4	2	OGDH	1	NUP214	1	TACR3	1	NTRK1	2	HLA.DOA	1	OR2H1	1	DHX15
2	KCNMB1	2	NCR3	1	GAN	1	CYP2D6	1	AGRN	2	LOC643454	1	TGM4	1	OXA1L
2	ADCYAP1 R1	2	PQBP1	1	MTR	1	P2RX3	1	CHUK	2	NDUFB10	1	SLC22A2	1	CDK4
2	CDH24	2	RIT2	1	EIF2S3	1	CCL4	1	DDX39B	2	CASP1	1	PAK3	1	RAF1
2	GLRA3	2	TCEB3	1	LIN54	1	ANPEP	1	GTF2H3	2	SUCLG1	1	RAB3B	1	RNGTT
2	SLC15A1	2	POLR2A	1	MTMR4	1	CCL25	2	PSMD6	2	ATP5J	1	SPTA1	1	WDR77
2	KCNJ9	2	PRPF6	1	PCM1	1	GRM3	2	HNRNPU	2	MAPK13	1	MMP28	1	CDC27
2	CLCA1	2	RAET1E	1	KIF2C	1	DGKA	2	POLR2D	2	PPP3CA	1	PDE1C	1	POLR3B

Appendix

2	PNOC	2	POLE	1	CLASP1	1	GABRR1	2	POLR1C	2	GGA2	1	CYP7A1	1	RBBP7
2	SLC6A14	2	FOXO4	1	FCGR1B	1	MC5R	2	SQSTM1	2	PNPLA8	1	GLI3	1	PSMC1
2	ADH7	2	RBM8A	1	PXMP4	1	SPN	2	PPIL1	2	STX8	1	KCNC4	1	PSME1
2	ITGA7	2	TNPO1	1	CSE1L	1	IFNA16	2	RBM8A	2	C7	1	RASGRF1	1	TUBA1B
2	RDH16	2	VAV2	1	DDX20	1	JUN	2	ERCC6L	2	CAPN2	1	PTCH2	1	ARIH2
2	OR2B6	2	PRKCZ	1	MTMR3	1	CD247	2	DHX8	2	GNAI3	1	HIST1H3A	1	PARK7
2	PSAT1	2	RBM17	1	TXNRD1	1	KCNG1	2	ZNF471	2	OR7E24	1	OPRM1	1	CTLA4
2	CNR2	2	CALR	1	ZNF614	1	COL6A2	2	RETNLB	2	PMM2	1	SYT2	1	EIF2B1
2	FPR3	2	IFNA16	1	CENPA	1	P2RX2	2	PPARD	2	RAB5C	1	LHCGR	1	MAP2K4
2	SLC5A3	2	CD28	1	UBA6	1	MAPK11	2	SRF	2	ARHGEF12	1	GNG3	1	UBA2
2	SLC26A4	2	TGFB1	1	STT3A	1	WNT9A	2	CD70	2	TXNDC5	1	LAMA4	1	ATIC
2	SLC6A7	2	RPS21	1	CFLAR	1	CNTFR	2	PSENEN	2	CACNG5	1	CST6	1	TLR7
2	TNN	2	POLD3	1	SEPHS1	1	GRIN3B	2	ZNF559	2	CHRNA7	1	ITGA7	1	HLA.DPA1
2	NPB	2	FAU	1	CLCA4	1	GPR132	2	MALT1	2	DCT	1	RYSR2	1	CYCS
2	RPS6KA6	2	RPS20	1	OR10J5	1	GRIK5	2	ZNF473	2	AP1G1	1	CFB	1	MINA
2	CCR9	2	RPL4	1	ITGA5	1	SYN3	2	FDFT1	2	PRSS2	1	CHRNA4	1	EIF3A
2	CACNB2	2	SF3B4	1	YARS2	1	F2RL3	2	HIST1H2B O	2	NDUFV2	1	OPN1SW	1	LSM6
2	SLC3A2	2	RPS2	1	ALG10B	1	CACNG4	2	PFN2	2	NDUFS6	1	SLC7A11	1	GRB2
2	OR7E24	2	RPL28	1	CCNG1	1	HRH3	2	HIST1H2B B	2	NDUFS1	1	LAMA3	1	NCAPH
2	GNB3	2	RPL39	1	PRKAR2B	1	MLNR	2	HIST1H2B E	2	NDUFA9	1	KCNS1	1	TRIM37
2	AGTR2	2	HIST1H4H	1	UBA3	1	DEFB123	2	CX3CL1	2	NDUFA7	1	CREB1	1	SEC23A
2	SLC5A5	2	NUP50	1	PRIM1	1	LIF	2	IRF3	2	NDUFA11	1	GUCY1A2	1	UBE2G2
2	DARC	2	XCR1	1	SLC11A2	1	CTRB1	2	MCM8	2	HIBCH	1	C1S	1	ADAM10
2	CNGB1	2	MEF2C	1	CCNB3	1	DRD4	2	TRIM33	2	GALK2	1	REN	1	MARCKSL1
2	SLC4A8	2	CYR61	1	B2M	1	ADH7	2	ARIH1	2	FZD9	1	LPAR3	1	PUF60
2	SLCO1C1	2	TNFSF8	1	MAPK9	1	UTS2R	2	IL1RAP	2	ATP1B3	1	SRPX2	1	ADSS
2	DAB2	2	TNFSF11	1	PSMA6	1	CHAT	2	PPP3CB	2	ATP5G1	1	PTH2R	1	SMC4
2	LPAR1	2	SDC1	1	FGFR1OP	1	GALR3	2	SNRNP27	2	IDUA	1	ADAMTS13	1	PPM1B

Appendix

2	MIP	2	IDUA	1	RPA2	1	GPR50	2	ZNF530	2	VDAC2	1	PI3	1	GNB2L1
2	LAMA1	2	EIF2B5	1	SYNE1	1	SLC22A2	2	TRIM28	2	UQCRC2	1	KCNK2	1	IL17A
2	NTSR2	2	GATA4	1	USP1	1	CCR5	2	PRPF31	2	TBC1D8B	1	CCL1	1	ARPC5
2	GATM	2	RPL36	1	CYCS	1	KLK2	2	CCDC99	2	NDUFB2	1	TSHR	1	EDEM3
2	MYL9	2	RPS12	1	DIDO1	1	ENO3	2	GRK5	2	MAPK7	1	UGDH	1	ELAVL1
2	CCL25	2	RPL19	1	TRIM21	1	ZAP70	2	PRKCZ	2	KCNC4	1	IFNA8	1	MGAT5
2	GUCA1B	2	PAK7	1	ZNF555	1	STX4	2	CDC23	2	GPX2	1	VCAN	1	TDG
2	GHRH	2	EHD4	1	HERC2	1	ADRBK1	2	SRSF11	2	D2HGDH	1	OR3A3	1	TNFAIP3
2	INS	2	IKBKAP	1	ZNF30	1	ALDH1A2	2	ZNF226	2	CLTA	1	TRPC5	1	NRAS
2	GABRE	2	TAF4	1	ARSB	1	KCNC3	2	PAG1	2	ABCG2	1	PPARA	1	PRKCQ
2	EGFR	2	RPL24	1	NUP54	1	FGR	2	PDE3B	2	CHRM3	1	KCNB2	1	CD44
2	SLC9A7	2	RPL23	1	GART	1	GGT5	2	CDC26	2	NDUFA5	1	LIMK1	1	TIAM1
2	LAMB4	2	MAP3K2	1	PSIP1	1	CCR2	2	SATB1	2	ADCY3	1	RGS4	1	IL6ST
2	ABCA12	2	EEF2	1	TPR	1	FGF8	2	ZNF544	2	APBB1	1	ADAM18	1	PRPF4
2	ACTN2	2	AGT	1	HDAC1	1	TRAF2	2	TUBA4A	2	ATP5G2	1	GATA4	1	KARS
2	OR1A2	2	CXCL2	1	EIF2S2	1	CACNA1A	2	NR3C1	2	COX5B	1	KDR	1	VARS
2	SST	2	MAP3K6	1	NUP85	1	GALR2	2	LAT	2	COX7A2	1	MC4R	1	TAF4B
2	GABRB1	2	CAV1	1	SNRPB	1	KCNC1	2	CCRL1	2	NDUFB9	1	FRAS1	1	CDC40
2	PTAFR	2	HLA.DQA1	1	HUWE1	1	TAS1R3	2	NUF2	2	VDAC3	1	SPAM1	1	SF3B1
2	PTH2	2	RPL22	1	CDC14A	1	DRD5	2	UBE2D2	2	ATP5L	1	OR6A2	1	DLAT
2	CSF2RA	2	RPL23A	1	MTMR1	1	RELA	2	MED31	2	NDUFS5	1	TNNT1	1	FCGR3B
2	FIGF	2	IFNA21	1	MCFD2	1	OPN1MW2	2	PRKAR1B	2	UQCRRS1	1	EFEMP1	1	KIF4A
2	THBS2	2	KIF11	1	SLC35D2	1	ESRRA	2	ZNF337	2	SLC25A5	1	NR0B1	1	TTC37
2	BCMO1	2	DKK1	1	OR1L4	1	MKNK2	2	IKBK6	2	POLR2L	1	GLP1R	1	NCF2
2	ITGB5	2	CD19	1	OR52E6	1	ST6GAL-NAC4	2	WWOX	2	GGCT	1	GRPR	1	PRKCD
2	DEFB129	2	CXCL3	1	MCL1	1	CACNA1H	2	CEP70	2	NDUFA4	1	CA9	1	HADHA
2	JAM2	2	ACTB	1	TXNL4A	1	HTR1B	2	ATF4	2	ATP6V1G1	1	SEMA5A	1	SSRP1
2	ABCC4	2	NFATC3	1	AARS	1	TNNI2	2	CASP9	2	GSTP1	1	EMID1	1	IRS1

Appendix

2	SLC17A6	2	SF3A2	1	ANAPC13	1	FOXO3	2	WNT11	2	HADHB	1	PLCB4	1	PMEP1
2	WNT6	2	FBXW7	1	IL13RA1	1	CNGB1	2	DHX16	2	KCNG4	1	HIST1H4A	1	HK1
2	UROCI	2	POLR2B	1	ASB13	1	MUC7	2	ZNF514	2	GUSB	1	LIMS3	1	LZTS1
2	UGT2B4	2	RBX1	1	CLIP1	1	SLC6A7	2	TNFRSF13B	2	IFI30	1	RIMS1	1	ZNF143
2	SLC2A2	2	RIPK3	1	FUCA1	1	SLC6A2	2	ICOS	2	ATP5O	1	CTNNA2	1	GAPDH
2	SLC11A1	2	PFDN2	1	HNRNPA3	1	FGF19	2	PSMB4	2	NDUFA13	1	NODAL	1	MKRN1
2	SERPING1	2	CSNK2B	1	DTX3L	1	IL20	2	FOXO4	2	APAF1	1	TRHR	1	VAV1
2	PTH	2	THOC4	1	SLC6A9	1	CETP	2	TGIF2	2	ARPC3	1	KCNJ10	1	ARHGAP10
2	P2RX3	2	CD55	1	YWHAH	1	KNG1	2	AIM2	2	ATP5H	1	UGT2B17	1	CUX1
2	OR51I1	2	MAL	1	INPP5K	1	PLA2G2C	2	RELA	2	UQCRB	1	LMAN1L	1	NCBP2
2	MTNR1B	2	HNRNPH2	1	OR2M2	1	GRIN2C	2	LSS	2	RAC1	1	SLC24A1	1	RPL12
2	KCNMB3	2	ETF1	1	RAB1A	1	GPR44	2	PUF60	2	ATP5C1	1	APCS	1	POU2F2
2	HSD11B2	2	HIST1H2BK	1	TAF4B	1	IL21R	2	SF3B14	2	NDUFC1	1	GABRA2	1	ARFGEF2
2	HSD11B1	2	LCP2	1	DSN1	1	PRSS3	2	CCT3	2	UBE2L3	1	ESRRB	1	DPM2
2	GRM7	2	CCT5	1	DPM1	1	KCNK13	2	FBXW2	2	NDUFB1	1	P2RY13	1	CENPI
2	GPER	2	C12orf5	1	OR5AR1	1	A2M	2	RAE1	2	COX6B1	1	PTH1R	1	ELL
2	FZD9	2	RPA3	1	PPP2CB	1	IKBKG	2	TRAF6	2	AUH	1	SEMA3B	1	TXNL4A
2	F11	2	POLR3GL	1	RRM2	1	LTB4R	2	SNW1	2	RRM1	1	KCNQ2	1	CSF2
2	EDNRA	2	SEC11A	1	TIPIN	1	FCER2	2	KLC4	2	ATP5F1	1	OR10J1	1	PAK2
2	DEFB114	2	EIF3I	1	ZNF583	1	NRTN	2	CCDC12	2	POLR2E	1	AXIN1	1	TMEM189.UBE2V1
2	DEFA6	2	SHC1	1	OR4C46	1	CHRNA1	2	MED20	2	DLAT	1	FMO2	1	RPS6KA1
2	CYP11B2	2	HLA.A	1	CTNS	1	OPRD1	2	NCOA3	2	ETFA	1	LOXL3	1	SLC19A1
2	CALML3	2	EIF4G2	1	HNRN-PUL1	1	SCTR	2	ZFP1	2	SNX5	1	SYN1	1	PHLPP1
2	ADCY4	2	RPL8	1	OR2G3	1	IL28B	2	ADD1	2	NDUFB5	1	AQP3	1	SRF
2	APOA1	2	SF3A3	1	TRIM28	1	RAC3	2	NXF1	2	PLDN	1	BPNT1	1	IFIH1
2	WNT2B	2	LSM7	1	PPIA	1	FZD10	2	ITPR3	2	AP3S1	1	ATP12A	1	ACTB
2	SLC1A6	2	RPL29	1	USP18	1	CTRB2	2	SLC7A9	2	MCCC1	1	CCL13	1	ABAT

Appendix

2	RXFP2	2	MAP-KAPK3	1	ABCG2	1	TIMP2	2	PTPN13	2	RHOC	1	HIST1H2BF	1	IKZF1
2	ABP1	2	FOXO3	1	PSMD3	1	CXCL6	2	ITGA6	2	ARPC1B	1	CSNK2A2	1	RPS2
2	SLC5A11	2	EEF1B2	1	NRAS	1	XCL2	2	PABPN1	2	FZD4	1	ADORA1	1	NUP54
2	SLCO1B3	2	RPL7A	1	TBL1XR1	1	CAMKK1	2	PPP2R5B	2	VEGFB	1	HEPH	1	RPS6KA2
2	SLC8A2	2	BCL3	1	EIF4G2	1	GNB2	2	ZNF773	2	BLOC1S1	1	DYRK1A	1	E2F3
2	SLC22A1	2	MARCKSL1	1	GK2	1	LTB4R2	2	ZNF12	2	CALM3	1	KCNG2	1	NDUFA2
2	SLC14A2	2	PRDX1	1	OR8A1	1	COL9A1	2	ZKSCAN1	2	NDUFA1	1	PYY	1	MAP2K3
2	RDH12	2	EIF4H	1	TBL1X	1	CALML3	2	SUGP1	2	UQCRC1	1	ANGPT1	1	NCK2
2	RAMP2	2	CEBPZ	1	RNF41	1	FGFR4	2	PKD2	2	ITPR1	1	AQP1	1	UBE3B
2	LRAT	2	ITPKB	1	CTSB	1	MAPK7	2	NUP88	2	NDUFS7	1	FGF4	1	DHX8
2	LAMA3	2	RPL3L	1	HEMK1	1	PPY	2	NFATC1	2	ATP5A1	1	FGF13	1	SNF8
2	KDR	2	SLC25A5	1	CASP6	1	ADRB1	2	GTF3A	2	CAMK2D	1	GPX5	1	BCAT1
2	KCNA2	2	SMARCA4	1	MTOR	1	KCNC4	2	CDC7	2	AP3B1	1	SCN5A	1	REL
2	HSD3B2	2	NCR2	1	ATP6V0D2	1	FGF17	2	GTF2F1	2	PRKCA	1	SLC22A4	1	DLG4
2	HMOX1	2	ATM	1	GTF2H5	1	ADORA1	2	SNRPA	2	SLC2A4	1	FCN3	1	GALNT10
2	HCN1	2	CHERP	1	RNASEH1	1	LCK	2	CYLD	2	KCNJ12	1	F2RL2	1	PDE3B
2	GRP	2	PTPRC	1	MED6	1	GNG4	2	NUP155	2	OR4A15	1	TEX15	1	PIGL
2	GPR77	2	HNRNPA2B1	1	OR10A3	1	TBXA2R	2	ZNF202	2	GNGT2	1	ATP6V1B1	1	MAP3K14
2	F2RL2	2	EXOSC2	1	OR56A4	1	DUSP5	2	ZNF418	2	OR5V1	1	PDGFRA	1	DDX39B
2	DEFA5	2	SRSF2	1	OR9Q2	1	MYL3	2	MIS12	2	PRKCG	1	HYAL1	1	ST8SIA3
2	APLNR	2	SMAD3	1	MTHFD2	1	CD2	2	TRIB3	2	NEFL	1	CTSE	1	ATF2
2	BCAR1	2	EIF4A2	1	ZNF286A	1	MUC3B	2	ZNF564	2	WNT1	1	KCNH1	1	GTF2F1
2	NPY5R	2	IL4	1	KIFC1	1	SSTR2	2	ZNF2	2	COMT	1	SLC35D1	1	PAPOLG
2	RASGRF1	2	ELF1	1	GIN54	1	COL27A1	2	TBXA2R	2	HRH3	1	BRAF	1	SRSF1
2	SCT	2	TNFRSF10A	1	PNLIP	1	RGS4	2	SMC3	2	ECHS1	1	DRD2	1	DHODH
2	HTR2A	2	ATP5H	1	PARS2	1	ITPKB	2	SMAD3	2	HADHA	1	RGS20	1	AKAP9
2	PLA2G2D	2	MAFK	1	UGT2B7	1	KCNK18	2	RPA2	2	TH	1	TAS2R16	1	PRKCA

Appendix

2	VIPR2	2	U2AF1	1	RNF123	1	ADAMTS14	2	RNGTT	2	OR1E1	1	ANGPTL7	1	TXNRD2
2	NPY2R	2	IL6R	1	CBS	1	GALR1	2	PRPF6	2	NDUFV3	1	PLA2G10	1	PANK3
2	COMP	2	MAP3K14	1	ERBB2	1	COMT	2	POLR2G	2	DHH	1	CXCL13	1	CHERP
2	LAMC1	2	EIF5B	1	H3F3A	1	GNAT1	2	OTUD5	2	CCR6	1	CETP	1	MEF2C
2	TNC	2	IL27RA	1	ARF3	1	ADRA1B	2	ORC2	2	TF	1	GUCY2D	1	DIS3
2	SLC6A5	2	TGIF2	1	SHC1	1	RXFP3	2	NUP210	3	KCNJ10	1	RGS12	1	IVD
2	KCNK1	2	RPL3	1	SLC1A5	1	MAFG	2	NHP2L1	3	CALM2	1	SLCO2B1	1	ATP5G2
2	PDE1C	2	OAZ1	1	MYBL2	1	GAL	2	DFFB	3	HDAC5	1	TGM2	1	VAV3
2	SLC4A1	2	TUBA1B	1	GPAA1	1	IL11	2	DCLRE1C	3	GNAI1	1	CFI	1	EIF2AK2
2	RHAG	2	TUBA3D	1	CASP9	1	GNG2	2	CLCF1	3	PCCA	1	GP6	1	NFATC1
2	GJA1	2	STRAP	1	CTSL2	1	SSTR3	2	CD79A	3	GSTZ1	1	INSL5	1	TRA2B
2	RHCG	2	HMGA1	1	RARG	1	CSF1	2	CD19	3	DAGLB	1	FGD1	1	TANK
2	TAAR8	2	ADAM17	1	PRKACA	1	RELB	2	CBLB	3	IFNGR1	1	MMP13	1	REST
2	OR1E1	2	BAX	1	RFT1	1	CYP2F1	2	ACD	3	BTK	1	SLC22A8	1	TAF4
2	CHRNA1	2	RNMT	1	OR4S1	1	SLC17A7	2	BLM	3	AP1S2	1	CPB2	1	HMGCS1
2	KCND2	2	DDX5	1	OR6F1	1	WNT4	2	RFXANK	3	POLR2J2	1	FZD7	1	RNF41
2	TH	2	NFAT5	1	IFIT1	1	KCNN1	2	BAD	3	ABCC4	1	PRKCE	1	OAZ2
2	GPR83	2	TCP1	1	SPC24	1	COL7A1	2	POLR2H	3	IGF1R	1	EDNRB	1	PIK3CG
2	F13A1	2	PSMC1	1	FZD1	1	RGS11	2	CCR7	3	CHP	1	RARG	1	NUP153
2	GNMT	2	STAM	1	RASGRF2	1	COL9A3	2	RRM2	3	PPP3R1	1	SERPINI1	1	UPF3B
2	CYBRD1	2	YWHAB	1	ABCA8	1	ENO2	2	ZNF287	3	MGLL	1	CRHR2	1	NEDD4
2	KCNJ14	2	ACTN1	1	MCM7	1	NR4A1	2	WASL	3	ATP6V0B	1	NTN1	1	XRCC5
2	FMO4	2	SNRPG	1	MCM4	1	ADCYAP1R1	2	USP11	3	ARHGDIB	1	CYP46A1	1	RANBP2
2	SLC28A3	2	RASSF5	1	HEPH	1	BAG1	2	LCK	3	ATP6V1A	1	GNG13	1	PIGV
2	HTR1F	2	PIK3CA	1	OR2D3	1	GPER	2	UBA7	3	ARRB1	1	GNA11	1	ARHGEF7
2	SLC7A9	2	ITK	1	OR51A4	1	MAP2K7	2	TBC1D4	3	FH	1	TGM1	1	PRPF18
2	KCNA5	2	SNRPA	1	PDPK1	1	WNT10B	2	RFXAP	3	SNX2	1	CDH2	1	BRCA2
2	TYRP1	2	KLC1	1	UGT2B10	1	AVPR1A	2	ZNF235	3	ARHGAP12	1	HLA.DOA	1	PPP2R5C

Appendix

2	ADRA2A	2	RIT1	1	GSTA1	1	NPW	2	TUBGCP6	3	UQCRQ	1	NELL1	1	SLC16A1
2	OR5V1	2	BRAF	1	OR10A2	1	FGFR3	2	TAF1C	3	NDUFA12	1	HIST1H2BK	1	NPM1
3	GABBR1	2	PUF60	1	DNASE2B	1	NR2F1	2	SUPT5H	3	AP3B2	1	DRD5	1	GIT2
3	PRLHR	2	COX6C	1	LOC391764	1	SLC6A20	2	RBM5	3	COX7C	1	GNGT1	1	SH3BP2
3	NUP155	2	PRPF8	1	OR5D14	1	CBR3	2	METTL3	3	FTL	1	PPP1R12B	1	DAPK3
3	PLA2G1B	2	EIF1	1	PPM1B	1	RYR1	2	LIPE	3	CLTC	1	ABCG4	1	ATF6
3	OR2M4	2	PSMB5	1	PRDM9	1	KCNA10	2	ATR	3	FPGT	1	ANXA8L1	1	PDE8A
3	SLCO2A1	2	CAMK4	1	APIB1	1	ATP1B4	2	HNRNPK	3	MCOLN1	1	DSC2	1	DOK1
3	PRSS3	2	HSP90AA1	1	NCAPD2	1	PGAM2	2	TRADD	3	UQCR10	1	CHRNA1	1	AQR
3	TOP3A	2	ITCH	1	SEPHS2	1	HSPA1B	2	TGFBR2	3	VAMP7	1	COL21A1	1	SRSF3
3	RAD1	2	CYFIP2	1	PKMYT1	1	NR2C2AP	2	STK11	3	SMS	1	SLC24A3	1	POLR2L
3	RMI1	2	UQCRH	1	SLC6A15	1	ECSIT	2	RBM17	3	AGA	1	FZD8	1	COX6C
3	C17orf70	2	EIF3H	1	OR4N5	1	EMR2	2	NFAT5	3	CREB3L2	1	KCNK3	1	STAG2
3	PTPN1	2	RPN1	1	OR52B4	1	NR2F6	2	CD3D	3	PGM2	1	WNT16	1	VDAC3
3	KCNJ3	3	IRF7	1	ELOVL2	1	TGFB1	2	IGLL1	3	AP1S1	1	CDH5	1	GTF3C2
3	SLC12A1	3	UCP2	1	OR52B6	1	RAB3A	2	ZNF436	3	CD58	1	EMR3	1	MPI
3	GHRL	3	HIST1H3H	1	OR13C8	1	STX10	2	ZNF485	3	STX7	1	SEMA6B	1	DCTD
3	OR3A3	3	IRF1	1	OR2K2	1	IHH	2	ZNF585B	3	ECT2	1	FTL	1	MAP2K1
3	CYP3A7	3	ARF6	1	FYN	1	KCNJ14	2	ZNF433	3	SUMF1	1	LOX	1	MNAT1
3	ITGA2	3	ICOSLG	1	A1CF	1	NPB	2	ZNF529	3	ATP6V1D	1	MMP14	1	PSMD6
3	PTPRA	3	KIR2DL4	1	ESPL1	1	PRKAG3	2	CTBP1	3	SNAP23	1	ABCG2	1	IFIT3
3	SSTR4	3	SRSF6	1	RASA4	1	CXCR3	2	ZNF248	3	HEXB	1	CTTN	1	PTPN11
3	TCIRG1	3	PAK2	1	HNRNPA0	1	PFKM	2	FCAMR	3	LTBR	1	ARR3	1	NFYB
3	ATP6V1F	3	ACVR1	1	RHOBTB2	1	PTGS1	2	NOTCH3	3	HSP90B1	1	CYP2A6	1	PCNT
3	GP1BA	3	PSMB3	1	OR4K13	1	COX6B2	2	EHD1	3	MAPKAPK3	1	KCNJ8	1	FBXW2
3	NUPL2	3	NDUFA8	1	ADCY5	1	ST6GAL-NAC2	2	CEP250	3	IFNGR2	1	CYP27B1	1	RAC1
3	ALDH7A1	3	GPC6	1	PIK3R6	1	PDYN	2	CRADD	3	MEF2BNB.MEF2B	1	FERMT2	1	TUBA3C

Appendix

3	FZD7	3	PDHB	1	HNRNPK	1	EPO	2	DHCR7	3	ARPC5	1	CHAT	1	EIF3J
3	ATP10A	3	SRP14	1	ZNF514	1	GADD45A	2	HMGCR	3	NPC2	1	OR7C1	1	NEIL3
3	KCNK6	3	ARID3A	1	ZNF419	1	GAPDHS	2	ZNF37A	3	NDUFB6	1	S1PR4	1	LG MN
3	P2RX1	3	IL23A	1	UNG	1	IL8	2	ERCC3	3	MITF	1	COL19A1	1	GMPS
3	GLRA1	3	SP1	1	FANCA	1	CRHR1	2	SF3A1	3	MGST3	1	TNN	1	UBE2N
3	SLC30A1	3	GTF2B	1	H2AFZ	1	GNA15	2	COL5A1	3	MBOAT2	1	ATP6V1G2	1	C1GALT1C1
3	GSTM5	3	STIM1	1	SATB1	1	OPRK1	2	BAX	3	HYAL2	1	KCNMB4	1	COX15
3	SLC5A8	3	EIF4A3	1	MLF1IP	1	ICAM3	2	ZNF212	3	HTR7	1	ACSM2B	1	ASB13
3	CYP11A1	3	AFP	1	ZKSCAN5	1	SLC18A3	2	B3GAT3	3	DUSP6	1	HCRT	1	CCND2
3	EGF	3	PHF5A	1	LIN9	1	GPR55	2	EBP	3	CD63	1	FBLN1	1	BMI1
3	ALDH3B2	3	IFNGR2	1	PIGK	1	IL28A	2	CD8B	3	CALML3	1	SERPINH1	1	GEMIN6
3	KCNB1	3	GPC2	1	ARHGDIB	1	CSF2	2	MAP4K1	3	ARHGAP18	1	HIST1H4I	1	CD28
3	SLC47A1	3	MCL1	1	PTPN13	1	MUC13	2	HNRN- PUL1	3	CACNA2D4	1	OPN1MW2	1	ACADSB
3	OR1G1	3	SRSF3	1	GIN52	1	INSL3	2	POLR2B	3	MYL3	1	ISL1	1	SNAPC1
3	GHR	3	HLX	1	RPA4	1	OXT	2	HIST1H1E	3	FZD5	1	SERPINB10	1	DCPS
3	GUCY2D	3	HCK	1	C19orf40	1	SCT	2	ZNF324	3	ATP6V1C1	1	SAG	1	ITGAV
3	OR6A2	3	CDC25A	1	MAN1A2	1	APOE	2	DDX42	3	CTSB	1	STEAP3	1	CDC42
3	SLC7A1	3	FOXP3	1	GTF3C2	1	GNG13	2	HDAC1	3	NANS	1	KCNV2	1	MALT1
3	OR2W1	3	PPP2CB	1	NT5C2	1	COL15A1	2	FDPS	3	DEPDC1B	1	CILP	1	SLK
3	SARDH	3	TNFRSF12 A	1	CSTF1	1	MAOB	2	CEP164	3	RHOA	1	DAB2	1	TRA2A
3	LEPR	3	RPL13	1	CAD	1	ADRA1D	2	SNRNP200	3	NLRP3	1	KCNA3	1	RPA2
3	SORBS3	3	TAB2	1	FANCG	1	FLT3	2	CDK11B	3	ACSS2	1	TLE2	1	DNAJC8
3	KCNJ2	3	RAB5A	1	MAT2B	1	CX3CR1	2	FLOT2	3	SDHC	1	SERPINA3	1	POLR2H
3	RGR	3	ITGA2B	1	MARS	1	KCNG4	2	CYP2J2	3	UQCRH	1	UNC5C	1	PPIA
3	FANCM	3	NFKBIB	1	SLC1A4	1	CD4	2	SRSF1	3	RAP1A	1	CLDN18	1	ABI1
3	SLC6A3	3	TNFAIP3	1	RNF6	1	TPSAB1	2	RAG1	3	SRGN	1	CLTCL1	1	STAG3
3	AQP9	3	CREB1	1	SMC1A	1	GNG8	2	ZNF92	3	STARD8	1	TAS2R14	1	SRSF9
3	CNGA4	3	SUCLG1	1	OR6N2	1	HSPA1A	2	HRAS	3	CDK5R1	1	CA5B	1	EIF4E

Appendix

3	ATRIP	3	HNRNPR	1	SLC27A4	1	SSTR4	2	DIDO1	3	KIF15	1	ADH7	1	POLR2D
3	PHGDH	3	KLRK1	1	TBPL1	1	COL17A1	2	ZNF606	3	ADRBK2	1	TBXA2R	1	ORC3
3	UCHL1	3	SOS2	1	CCNA2	1	SLC25A5	2	IP6K2	3	SORT1	1	CP	1	TBPL1
3	SLC2A6	3	HMGB1	1	MLEC	1	PLA2G2F	2	ZNF425	4	ADCY7	1	MARCKS	1	NCR3
3	ICAM1	3	ANAPC7	1	PDIA3	1	GFRA4	2	ZNF510	4	NAGLU	1	MAPK8	1	ATP5A1
3	INSL5	3	PCNA	1	CALM1	1	KCNV2	2	DEDD	4	FZD2	1	ADORA3	1	POLR1E
3	GNGT1	3	ATP5B	1	VHL	1	KCNG2	2	E2F5	4	NCF4	1	BRINP2	1	PSMC5
3	LAMC2	3	CDK7	1	ZNF558	1	PTGER4	2	LTB	4	IDH2	1	PARD6B	1	ATP5C1
3	ITGA9	3	RPN2	1	HNRNPD	1	PLTP	2	IPCEF1	4	DLG4	1	HSD11B1	1	CD8B
3	SLC8A1	3	EXOSC9	1	OR2T1	1	SLC6A1	2	DOCK9	4	OR10J5	1	SLC4A3	1	FANCF
3	KCNH3	3	TRAF2	1	TAF4	1	PTPN5	2	FYN	4	RHOBTB2	1	SNCA	1	STRAP
3	KCNK3	3	NFKBIL1	1	UBA7	1	ADRA1A	2	ZNF211	4	OR10A4	1	GRIN2B	1	SMNDC1
3	GRM4	3	H3F3B	1	ZNF445	1	KCNQ4	2	MYBL1	4	BACE1	1	DRD1	1	SNAPC3
3	PVRL1	3	NME1	1	ZNF430	1	CYP1A2	2	PARP1	4	SYK	1	SERPINA7	1	CASP6
3	GABRA5	3	CASP10	1	CENPH	1	OPN1LW	2	PLRG1	4	C4BPA	1	FGF21	1	CCT6A
3	LTB4R	3	IL2RA	1	PIGY	1	COL11A2	2	KIF2A	4	HAP1	1	RELN	1	MAPK13
3	SCTR	3	PSMD14	1	KIF23	1	NMS	2	ZNF484	4	KCNK9	1	CYP39A1	1	POLR3F
3	DEFB4A	3	JMJD1C	1	OR52N2	1	IL29	2	ITK	4	SLC17A7	1	CREB3L1	1	MAN1A2
3	CHRNA5	3	SNRPB	1	LMNB1	1	CCR6	2	WBP11	4	IVD	1	OR2W1	1	CDCA8
3	ALDH3A1	3	CSNK1A1	1	ZNF267	1	CNGA1	2	PRPF3	4	ADHFE1	1	CCK	1	PSIP1
3	SLC18A1	3	SRF	1	ABCA10	1	LAT	2	SPTAN1	4	BCL2	1	ABCA8	1	CETN2
3	GNRH1	3	SMPD1	1	PSMA5	1	PLAU	2	MED25	4	GIPR	1	PTGER3	1	DKC1
3	SLC3A1	3	UBE2N	1	ABCB1	1	FGF5	2	RAC2	4	GLG1	1	NR0B2	1	SLC25A13
3	ASPA	3	CSF3R	1	RPA3	1	GNGT1	2	MED14	4	KCNAB2	1	UGT1A7	1	EIF4A1
3	ATP6V0A4	3	EIF3E	1	HIST1H4L	1	IL12A	2	NCOA1	4	NLRX1	1	P2RY4	1	IFITM1
3	CYP26A1	3	HDAC1	1	MCM5	1	GPR35	2	PHF5A	4	GSTA4	1	CYP4F3	1	RFC1
3	MDM2	3	EXOSC1	1	LIG4	1	OPN1MW	2	PSMD4	4	BCKDHA	1	BMP10	1	UQCR10
3	GCG	3	SF3B2	1	OR56A1	1	HNF4A	2	CASP3	4	HADH	1	ERBB2	1	RNMT
3	OR11A1	3	UBE2A	1	AURKA	1	JAM3	2	GTF3C3	4	CTSS	1	SERPINI2	1	YWHAQ

Appendix

3	FSHB	3	POLR3G	1	PANK2	1	RHO	2	NSMAF	4	KCNK13	1	ALOX15B	1	DDX46
3	GRIA4	3	RBM25	1	HIST1H4F	1	ALOX15	2	MST4	4	DRD2	1	TCHH	1	CAD
3	EMR3	3	PRKCH	1	TPK1	1	DDIT3	2	ZNF23	4	DUSP3	1	SEMA4G	1	ATP5B
3	GRIA3	3	TAB3	1	TUBA3C	1	EVL	2	PPP2R2A	4	PAK1	1	GJB4	1	ATP5G1
3	SLC17A8	3	RFWD2	1	UBA1	1	FGF21	2	ELL	4	TNFSF13	1	LTBP2	1	POLD3
3	F10	3	PSMD7	1	SKIL	1	KCNK3	2	FGFR1	4	HLA.DRB1	1	CTBP2	1	RALB
3	NPFFR2	3	NFKB2	1	IL4	1	HCN2	2	GNB5	4	HLA.DRB5	1	SLC9A7	1	ATG7
3	SLC6A9	3	DCP1A	1	ZWINT	1	CHRNA6	2	ZFP2	4	HLA.DPB1	1	ADCYAP1	1	PALB2
3	NPBWR1	3	FLNB	1	BUB1	1	IL5RA	2	ZNF643	4	SLC8A1	1	COLEC12	1	ATP6V1A
3	KCND1	3	WAS	1	PSMD2	1	DEFB132	2	EZR	4	HLA.DPA1	1	EMR1	1	LSM5
3	GRID1	3	POLR3H	1	POLE	1	KCNN3	2	PLCG1	4	HLA.DRA	1	SLC34A2	1	IFNAR2
3	HSD17B6	3	PSMB4	1	RAP1A	1	IL17A	2	ZNF343	4	KCNK6	1	SLC12A6	1	HNRNPD
3	GALR1	3	RASGRP3	1	ZNF250	1	MYLK2	2	DYRK1A	4	PRCP	1	CLCA1	1	NUP155
3	CYP7B1	3	KIF23	1	FUS	1	F10	2	ZNF17	4	LY86	1	CYP4A22	1	NBN
3	FZD2	3	PPP2R3B	1	ALG1	1	FGF22	2	CCT7	4	MGST2	1	SHC3	1	EIF4G1
3	RDH5	3	TGFBR2	1	TIRAP	1	CNTN2	2	POLR2C	4	KCNQ1	1	ADRA1D	1	HLA.DRB1
3	SLC32A1	3	SLC16A3	1	CCNB1	1	ACE	2	CDC5L	4	CFD	1	OR2B6	1	IFITM3
3	DEFB125	3	MAP3K12	1	ZNF354B	1	NGF	2	MVD	4	CYBRD1	1	DSG2	1	KIF18A
3	KCNH8	3	IRS1	1	HIST1H3F	1	CYP21A2	2	BTRC	4	TLR5	1	CHRNA3	1	PARN
3	ITGB1	3	ANAPC11	1	HNRNPM	1	SST	2	STIM1	4	PECAM1	1	WASL	1	EIF2B3
3	ABCC11	3	CCL23	1	PSMD11	1	MUC5B	2	ACACA	4	CD36	1	SLC17A5	1	OAS1
3	AQP1	3	UBA1	1	CDC25A	1	SLC6A11	2	ZNF300	4	CSF3R	1	CALCB	1	ACTR3
3	GRIN2A	3	CCT7	1	ZNF45	1	COL8A2	3	FUS	4	TXNIP	1	IFNA21	1	SLBP
3	RASGRF2	3	PPM1A	1	SNRPD1	1	CTSG	3	KPNB1	4	WNT10B	1	CHAD	1	IL12RB2
3	CHRNA4	3	IL2RG	1	ZNF354A	1	RDH12	3	MEN1	4	GLB1	1	HSD11B2	1	EIF3B
3	ABCA2	3	PSMC5	1	POLR3H	1	APOC2	3	CEL	4	PLA2G15	1	MASP2	1	LARS
3	CACNA2D3	3	SRSF7	1	POLR3G	1	SLC8A2	3	ZNF43	4	ACAA2	1	PRDM9	1	ALG6
3	PIK3R3	3	PPP3R2	1	SLC41A1	1	ESR1	3	ZNF625	4	MANBA	1	GPX3	1	NDUFA6

Appendix

3	ABCG1	3	RPL10	1	FEN1	1	NFKBIA	3	IL27RA	4	SERPINF2	1	PGR	1	CNOT2
3	GLRB	3	RHOA	1	DOLPP1	1	ALOX15B	3	C8G	4	TAB1	1	CYP3A43	1	YWHAH
3	PENK	3	POLR2E	1	PRF1	1	CD14	3	GRK4	4	CDH3	1	PTN	1	SUPT16H
3	TRPV1	3	MAP3K11	1	ACSL4	1	HCN4	3	GTF3C2	4	KCNJ1	1	CPA3	1	PDK1
3	CBS	3	FYB	1	HNRNPU	1	FLNC	3	PRKX	4	ARHGAP26	1	FGF20	1	POLE3
3	PRKCB	3	HPSE	1	GEMIN6	1	GABBR2	3	MAML1	4	CACNA2D3	1	TAS2R3	1	BIRC5
3	CHRNA1	3	USP9X	1	CCNB2	1	CYP11B2	3	MKNK2	4	ALDH2	1	CHRNA10	1	ORC2
3	CYP2D6	3	SRSF5	1	HIST1H2A B	1	ATP1A3	3	POLD2	4	LRP1	1	NPY5R	1	MAGED1
3	PTGER1	3	PSMC4	1	FANCB	1	MMP25	3	ZNF263	4	PLBD1	1	AKT3	1	IL21R
3	MET	3	PSMA4	1	RRN3	1	PRG2	3	IDI1	4	RNF135	1	CHP2	1	POU2F1
3	GRIK5	3	POLR2J	1	STAG2	1	AMN	3	NOP56	4	CD33	1	GNAS	1	CLNS1A
3	SLC9A9	3	LCK	1	ZFP28	1	DHH	3	KIRREL	4	PYCARD	1	MMP19	1	CNTRL
3	KCNK2	3	KIFAP3	1	DDX23	1	KLB	3	SLC25A20	4	ITGAM	1	NID1	1	TGFBR1
3	DMGDH	3	AZIN1	1	UBE2Q2	1	CCL21	3	PRDM9	4	CD4	1	TREM2	1	PANX1
3	SHMT2	3	EIF1AX	1	HIST1H3G	1	TPM1	3	RAP1B	4	C5	1	SLC28A3	1	ABL1
3	GJB6	3	XRCC5	1	KPNA5	1	HCK	3	UBTF	4	CD1D	1	GH1	1	LARS2
3	ATP7B	3	WBP11	1	RIPK1	1	HTR6	3	KIF3C	4	NAGA	1	IL19	1	THOC2
3	SLC44A1	3	SOCS4	1	HSD17B4	2	LHCGR	3	SREBF2	4	PLA2G2A	1	CNGB1	1	CDC23
3	CHAT	3	RPS15	1	ITGAV	2	MOS	3	NCK2	4	GAA	1	SLC6A7	1	IMPDH2
3	KCNJ11	3	NDUFB6	1	SLC15A2	2	CRH	3	STK24	4	CHRM1	1	GUCA1A	1	NFKB1
3	PLCB4	3	ICAM2	1	ZNF643	2	WNT8B	3	ZNF154	4	FGD4	1	MFGES	1	APC
3	SLC9A5	3	CCT3	1	KLF8	2	P2RY1	3	ZNF274	4	MEFV	1	TNXB	1	ITGA4
3	MYLK	3	BUB3	1	TUSC3	2	GP5	3	DAXX	4	WNT9A	1	INHBB	1	MSH6
3	NUP43	3	CCT2	1	MSH6	2	LTA	3	PPP1R13B	4	HK3	1	STAR	1	MCCC2
3	KLK6	3	ORC2	1	OR56B4	2	ITGA7	3	ZAP70	4	OR6N1	1	WNT2	1	SMC3
3	SLC37A4	3	CCNH	1	NRG1	2	NR1D1	3	PRPF8	4	OR6K2	1	EPB41	1	STAT3
3	CHRNA6	3	SKP1	1	LARS2	2	PPBP	3	ZNF225	4	HTR1E	1	ACACA	1	TUBB
3	TAS1R1	3	EXOSC8	1	SORT1	2	KCNA3	3	PARD6A	4	MAN2B1	1	ADRA1A	1	NUP93

Appendix

3	SLC15A4	3	PPIL1	1	CCT5	2	MUCL1	3	NUP214	4	OR8H1	1	MYL10	1	CASP3
3	SLC7A11	3	NUP93	1	POLA2	2	FGF10	3	KIFAP3	4	CD1A	1	CDH9	1	MAP2K7
3	AQP6	3	PDP1	1	PSMC2	2	GABRA1	3	STK36	4	TAAR8	1	WNT8B	1	METTL3
3	ADCY9	3	RAP1A	1	ZNF226	2	IL22RA2	3	ENDOD1	4	OR1S2	1	FGF16	1	MYD88
3	GNAL	3	GADD45A	2	ABCB4	2	ACTC1	3	TH1L	4	GBA3	1	VIPR2	1	ITPA
3	DEFB1	3	NFKB1	2	UGT2B17	2	SAG	3	CDK7	4	MYH10	1	F13A1	1	CD74
3	KCNG4	3	AP2M1	2	ZNF585A	2	GGT6	3	CTDP1	4	RAB3A	1	ADRB3	1	IRF2
3	CYP21A2	3	PSMD4	2	CENPN	2	KCNQ1	3	RPS6KA1	4	OR10J3	1	RAMP3	1	GANAB
3	ICAM4	3	TH1L	2	ZNF304	2	ADCY4	3	POLE2	4	MAP2K6	1	TGFB1	1	CD7
3	HTR4	3	CUL1	2	NBN	2	GNRHR	3	APITD1	4	ADRB2	1	ADAMTS12	1	MTA1
3	GLDC	3	RAN	2	BRCA1	2	OPN3	3	CENPO	4	ALDH3A2	1	AKAP5	1	ATP2A3
3	SLC15A3	3	POLR1C	2	MAT2A	2	EDA	3	MAGOH	4	GALK1	1	HIST1H2B O	1	CALR
3	SLC28A1	3	H2AFV	2	AP3M2	2	LPAR4	3	HIST1H4H	4	CREB3L4	1	SV2B	1	PARVB
3	THBS3	3	HNRNPK	2	RAD51	2	PMM1	3	GZMA	4	OR4D2	1	TGFA	1	TLN1
3	ABCA3	3	NDUFA6	2	ZNF564	2	C3AR1	3	PTPN1	4	FZD1	1	CDH18	1	UCP2
3	AOX1	3	PPP2R2A	2	PAFAH1B 1	2	COX4I2	3	POLR2J	4	TLR4	1	PRL	1	TGS1
3	CYP4A22	3	MAP3K1	2	SGOL1	2	FLT3LG	3	TERT	4	ATP6V0A1	1	COLEC10	1	NME4
3	PTCH2	3	ARF4	2	TFDP2	2	GRK1	3	CYP1A2	4	HYAL1	1	RDH16	1	CD55
3	OR2B3	3	RPS6KB1	2	ACSL5	2	IL6R	3	MAP3K10	4	DEFA5	1	GSTT2	1	DARS
3	ATP2B3	3	CDC26	2	POT1	2	FGF11	3	MVK	4	HLA.DQA1	1	SLC44A4	1	HLA.DQA1
3	HCN3	3	PSMB1	2	PIK3CB	2	CYP3A7	3	SF3B3	4	HLA.DQA2	1	OXT	1	HLA.DRB5
3	GABRG1	3	TUBB3	2	ZNF480	2	LPAR3	3	ORAI1	4	HLA.DMA	1	ARHGEF4	1	CTSO
3	PROS1	3	NDUFB9	2	ZNF222	2	DEFB127	3	PHKA2	4	TRIO	1	IFNA16	1	KPNB1
3	SLC44A3	3	PPP2CA	2	PIK3R3	2	MMP8	3	MAPK1	5	DAGLA	1	IFNA17	1	GALNT12
3	FZD10	3	NCOR2	2	GTF2B	2	SLC18A1	3	MAPK3	5	PIK3R6	1	PLA2G1B	1	MCM4
3	SLC13A1	3	ORAI1	2	ZNF398	2	IL3RA	3	TNFRSF1B	5	TLR7	1	MUC3A	1	TPX2
3	HSD3B1	3	NDUFAB1	2	SMC4	2	RDH8	3	HSD17B7	5	EEA1	1	GABRB1	1	MAN1A1
3	GALR3	3	ZMAT2	2	MAD2L1	2	HRH1	3	MTOR	5	CTSD	1	OR1D2	1	PSME3

Appendix

3	SNAP25	3	CWC15	2	SHMT1	2	IFNA1	3	MED1	5	ACP5	1	NAT2	1	KIT
3	STAT5B	3	RAB11A	2	CTSS	2	P2RY12	3	KPNA3	5	GNPDA1	1	ROCK2	1	KLRC3
3	GP6	3	SMNDC1	2	POLD3	2	RXRG	3	PIK3R5	5	NANP	1	OR1E2	1	CCR7
3	SLC34A3	3	RAD23B	2	ZNF92	2	ABCA1	3	HIST2H2A C	5	SLC18A3	1	CA8	1	DSE
3	SLC39A6	3	PARP1	2	EIF5	2	DEFB121	3	ZNF169	5	SLC16A1	1	GRP	1	RNF138
3	SLC16A8	3	PSMA6	2	RAD50	2	CACNG3	3	PPP2R5D	5	CAMK1	1	SLC5A1	1	SEH1L
3	SLC25A10	3	EIF2S1	2	OR2B6	2	SLC44A5	3	CARD11	5	CD276	1	TCF7L2	1	IDH3A
4	PLA2G6	3	CD48	2	UBE2T	2	PCOLCE2	3	CD247	5	ARHGAP31	1	CACNB1	1	STAG1
4	CGA	3	SNRPD3	2	CCNE1	2	PTHLH	3	ZNF3	5	RILP	1	GRIN1	1	MICB
4	MGST3	3	ROCK1	2	CLSPN	2	CCL27	3	DNM2	5	ATP6V1F	1	P2RY2	1	PAK1
4	SLC35A3	3	UQCRRF5	2	ZNF311	2	ABAT	3	PRPF4	5	CTS2	1	LAMA1	1	PDHX
4	PRKACB	3	POLR1B	2	CENPQ	2	DEFB118	3	TERF2IP	5	COX5A	1	NPY1R	1	USP1
4	TPM3	3	EIF4G1	2	DBF4	2	UGT2B15	3	CCL24	5	YWHAH	1	SLC2A10	1	DDB1
4	ADH5	3	TAOK1	2	GTF2H3	2	GLS2	3	EGLN3	5	NNT	1	SLC39A2	1	CCR5
4	CTNND1	3	VBP1	2	MOCOS	2	RASGRP4	3	RARG	5	GNPTAB	1	HSP90AA1	1	CTSC
4	AQP4	3	SRSF9	2	GTF2E1	2	GAD1	3	PXN	5	GM2A	1	CHRNA5	1	GRK5
4	IBSP	3	PDHA1	2	ZNF607	2	NR4A2	3	PPP2R5C	5	G6PD	1	KAL1	1	IL27RA
4	NUP133	3	RAC2	2	TAS2R10	2	GRIN2B	3	YWHAQ	5	CYC1	1	GIP	1	NHP2L1
4	KCNH7	3	SDHD	2	DZIP3	2	CYP11A1	3	SH2B1	5	ATP2A2	1	INS.IGF2	1	CD40LG
4	ATP6V1D	3	CLP1	2	CSTF3	2	RGS9BP	3	SF3B2	5	BCAT2	1	L1CAM	1	NOP56
4	SLC38A1	3	E2F3	2	KIF15	2	CXCL3	3	PIK3CD	5	ATP6V1B2	1	OR7A5	1	ITPR3
4	MYL6	3	SRSF8	2	HIST1H3E	2	KCNMB1	3	FASLG	5	GPX4	1	CTRC	1	MDC1
4	FTCD	3	PPP2R1A	2	PRKCQ	2	NTSR2	3	GNAS	5	PLA2G4C	1	GHRHR	1	EGR1
4	OR1Q1	3	PSMD5	2	CDC23	2	DRD1	3	ZNF587	5	ARHGAP10	1	GLI2	1	SHC1
4	GSTK1	3	SRP68	2	RPS6KA3	2	FGF9	3	NR2C2	5	IDH3G	1	PRELP	1	MGAT4B
4	PRKAR1A	3	MYD88	2	PIGW	2	KCNJ6	3	CPT1C	5	AP2S1	1	COL11A1	1	MOGS
4	COL18A1	3	ARAF	2	AGL	2	PGR	3	CFLAR	5	CD74	1	GUCY1B3	1	PPP1R13B
4	ITPR2	3	EIF2B1	2	CROT	2	EMR1	3	CDKN2D	5	ND3	1	IL36RN	1	PLEC

Appendix

4	CP	4	ATP5G1	2	ALG11	2	IL27	3	AKT2	5	CANX	1	PRKAB2	1	MYH9
4	FBXW11	4	NDUFB3	2	ANAPC1	2	GFRA1	3	BGN	5	SUMF2	1	SLC5A5	1	SOS1
4	GNAQ	4	PSMD13	2	LCMT2	2	IL3	3	XIAP	6	RENB	1	THBS1	1	ARPC1A
4	RPA1	4	PSMA8	2	ZNF175	2	ELMO1	3	TBXAS1	6	HLA.DQB1	1	WISP3	1	PPP2R5D
4	DDC	4	SEC61G	2	FANCM	2	GRIA3	3	SUN2	6	HLA.DRB3	1	ABL2	1	STAT5A
4	WNT10A	4	NDUFA12	2	RPS6KA6	2	HLA.DPA1	3	SOCS4	6	GNB4	1	ITIH2	1	WDR61
4	FANCI	4	POLA1	2	CDC7	2	CCL11	3	SLC2A1	6	ATP1B1	1	ASIP	1	B4GALT5
4	KCNH2	4	XCL1	2	OR2D2	2	CXCL10	3	SLC1A7	6	OPLAH	1	ANXA10	1	SRRM1
4	FKBP1A	4	CSF2	2	SKP2	2	UQCRB	3	SF3A2	6	KCNMB1	1	TACR3	1	E2F1
4	SLC26A1	4	FCER2	2	ZNF600	2	DNTT	3	RORA	6	CD1E	1	GUCY1A3	1	SIPA1
4	PARK2	4	COX6B1	2	ZNF493	2	MC2R	3	PRKCQ	6	CIITA	1	KCNMB1	1	TACC1
4	GPR17	4	E2F1	2	MRE11A	2	MAS1	3	PRF1	6	TAAR2	1	OR2C1	1	ENO1
4	GRM2	4	IL12RB2	2	ZNF248	2	GNG3	3	PREX1	6	C3	1	IFNA5	1	CFL1
4	ATP1B3	4	TNFRSF9	2	MND1	2	IL2	3	PDGFRB	6	PTPRF	1	MAT1A	1	SF3B4
4	FANCB	4	SNRPD2	2	ABCD3	2	ICAM2	3	PAPOLA	6	TNFSF13B	1	COL4A2	1	PIGT
4	ADCY1	4	TIMP1	2	BLM	2	RXFP2	3	NCOR1	6	AMDHD2	1	IL24	1	PTBP1
4	ATP4B	4	ERCC5	2	RFWD2	2	MUC17	3	MED16	6	PDCD1LG2	1	TBXAS1	1	NUP210
4	OR1J2	4	SUV420H1	2	BRIP1	2	SLC6A5	3	KIF3B	6	SERPING1	1	UNC13B	1	OGDH
4	DNM2	4	RPL13A	2	CENPI	2	DEFB110	3	KIF26A	6	IL2	1	GPR50	1	HDAC3
4	LCMT2	4	SUPT4H1	2	PEX3	2	KCNH6	3	HIST1H3I	6	GNG4	1	CST8	1	PPP3CB
4	KCNH6	4	TNFRSF4	2	TAS2R9	2	MYL2	3	FOXO3	6	CTSC	1	GLYAT	1	VPS37B
4	ATP11A	4	DUSP3	2	MED23	2	EDAR	3	EIF4G3	6	C2	1	FLG	1	ANAPC2
4	ABCA4	4	ITGA4	2	TAS2R14	2	IL12B	3	DAPK3	6	C1QC	1	SLC11A1	1	UBA1
4	GRIK4	4	AP1B1	2	TAS2R31	2	TGFB3	3	DAPK2	6	C1QA	1	FGF14	1	PRPF8
4	ROCK2	4	SPCS1	2	ZNF595	2	NPSR1	3	CYP7B1	6	C1QB	1	IL1B	1	PTPN6
4	SLC30A7	4	PSMD11	2	TAS2R3	2	COX8C	3	CYP4F2	6	NAGK	1	LAMC1	1	TGFB1
4	TFRC	4	SOCS2	2	ZNF528	2	AVPR2	3	CHST12	6	PDGFB	1	FCGR1A	1	AKT1
4	ATP11C	4	RXRA	2	PCF11	2	IFNE	3	CDKN1B	7	AP1M2	1	SLC4A8	1	CISH
4	PIK3CB	4	SNRPB2	2	ZNF223	2	DEFB119	3	CD8A	7	DRD3	1	S100A5	1	SNRPA

Appendix

4	CDH1	4	VASP	2	ITPR2	2	EDN1	3	CCR4	7	OR5L2	1	TAS2R9	1	VASP
4	WDR48	4	ACVR1B	2	TTK	2	CXCR4	3	CCL5	7	CREB3L1	1	ERBB3	1	ARPC2
4	ART1	4	PFDN5	2	ZNF655	2	TSHB	3	CCL25	7	GPX3	1	F9	1	ITGB7
4	ANXA1	4	PRKAR2A	2	ZNF320	2	FBP2	3	B3GAT2	7	CD86	1	SLC7A8	2	STAT4
4	GJB3	4	POLR2I	2	ZNF658	2	HLA.DQA1	3	BCR	7	PLA2G1B	1	CACNB4	2	CEP76
4	MCHR1	4	L2HGDH	2	SPC25	2	RGS18	3	EDA	7	ARHGAP44	1	FPR3	2	CASP1
4	FN1	4	RNF144B	2	TAS2R20	2	CD1B	3	LAMC2	7	ICA1	1	CYP11A1	2	CD19
4	RFC3	4	ATP5I	2	OR6A2	2	FCGR3A	3	ACIN1	7	HGSNAT	1	NPY2R	2	HNRNPA1
4	INSR	4	POLR1D	2	TAS2R50	2	FGF14	3	PRKAB1	7	OR1N2	1	COL11A2	2	XAF1
4	PTGER3	4	SH3BP2	2	CDC6	2	GADD45B	3	EXOC7	7	GNAQ	1	BDKRB1	2	HSPA9
4	CHRNA2	4	RPL26	2	FANCL	2	CD3D	3	CPT1B	7	OR10G8	1	USH2A	2	IDH3G
4	NCALD	4	RPL36A	2	PIGN	2	COX7B2	3	SHC1	7	PCK1	1	HIST1H2BH	2	IL15
5	NBN	4	RPSA	2	GPAM	2	CCR1	3	MAP2K1	7	DGKD	1	SLC7A9	2	NME1.NME2
5	MC1R	4	CCNT1	2	ZNF347	2	NR4A3	3	TINF2	7	PRKCB	1	TIAM2	2	MTMR4
5	CD36	4	ICA1	2	TAS2R13	2	CALCB	3	CHPF2	7	BACE2	1	ACTN3	2	POLR3C
5	RRAS	4	TAF13	2	ZNF234	2	COL8A1	3	DHCR24	7	CD180	1	KCNQ3	2	COX7B
5	CAMK2D	4	SEC61B	2	BRCA2	2	CYP8B1	3	NRG2	7	KCNH8	1	CYP11B1	2	DHX9
5	SLC6A11	4	HLA.DRB1	2	ZNF141	2	ITGA10	3	CKAP5	7	ITGA4	1	OR1G1	2	EIF3I
5	PPP1R12B	4	IFNA8	2	MBD4	2	PRL	3	RXRB	7	ARHGEF4	1	PENK	2	MOCS2
5	COL4A6	4	TUBB6	2	SLCO1B3	2	HLA.DMA	3	FBXW4	7	CTSG	1	LGI1	2	HNRNPM
5	UGT2A3	4	MAF	2	ZNF484	2	IFNA5	3	ZNF434	7	DEFA4	1	PAMR1	2	MAPKAP1
5	CACNB1	4	PRKCE	2	CENPE	2	CACNG1	3	ZNF45	7	MEF2C	1	TRPV1	2	PPP1CC
5	SORBS1	4	POLR2K	2	POLE2	2	TDGF1	3	RAF1	8	CTSK	1	IFNA14	2	RBMX
5	H2AFX	4	HLA.E	2	ZNF225	2	GYPA	3	DDX23	8	STARD13	1	EMCN	2	ALG8
5	KCNMB4	4	TGFBR1	2	RFC4	2	PLA2G2A	3	MCM3	8	GP9	1	GABRA4	2	NUP107
5	DEFB127	4	SMAD7	2	XRCC2	2	TNFSF11	4	SLC6A6	8	ASAH1	1	HPSE2	2	IRAK4
5	ATP2A1	4	HCST	2	SLC26A2	2	CACNB4	4	TRIP10	8	PPT1	1	NR5A2	2	DZIP3
5	CDH11	4	LDHA	2	DCLRE1C	2	CYP2A13	4	TRAF2	8	NPL	1	RYS3	2	USP6NL

Appendix

5	SLC24A1	4	DUSP4	2	PTTG2	2	DEFB113	4	PDK1	8	RHOB	1	TCF7L1	2	APH1A
5	PLA2G12A	4	APOBEC3G	2	CCNE2	2	KCNK10	4	TNFRSF4	8	EPB41L1	1	SCN4A	2	NACA
5	PRKG1	4	CDK6	2	CTSK	2	ABCG1	4	HIST1H2BJ	8	CTSL2	1	CD36	2	HRAS
5	GABRA1	4	TXNL4A	2	CEP290	2	PLA2G1B	4	LAMC1	8	ITSN1	1	PRSS1	2	POLR2J
5	SLC40A1	4	SHC3	2	EIF2AK2	2	CYP17A1	4	EHD2	8	HMOX1	1	CGA	2	PAPD7
5	ADAM28	4	SF3B14	2	TAS2R19	2	CHP2	4	H1F0	8	OSBPL1A	1	GABBR1	2	EDEM1
5	KCNS2	4	CXCR4	2	APC	2	SCARB1	4	HIST1H2A C	8	PLTP	1	FZD3	2	RBCK1
5	STEAP3	4	POLR2L	2	TAS2R4	2	GABRQ	4	MSMO1	8	OPHN1	1	FBN2	2	SNRPB2
5	PIK3R2	4	BCL2L14	2	ACACA	2	GPR68	4	CXCR6	8	SLC46A1	1	SLC22A11	2	AP1B1
5	CAMKK1	4	ATP5J2	2	HIST1H3A	2	GRK7	4	LAMA5	8	PDK4	1	GABRB2	2	SRP54
5	CLTB	4	RAC1	2	KIF18A	2	ITGAX	4	NDRG1	8	F13A1	1	GSTA1	2	ATP6V1H
5	CYP17A1	4	IKZF1	2	TAS2R46	2	CACNA1I	4	NQO1	8	KCNMB4	1	IL1RN	2	PSMB6
5	PLCB1	4	CD79B	2	MSH2	2	CNR2	4	CCL17	8	VIPR1	1	SLCO1C1	2	HNRNPR
5	SLC4A7	4	CDKN1A	2	ZNF215	2	GRM5	4	CDH1	8	TPM2	1	S100A7	2	RAD51
5	ABCB6	4	EHD1	2	PGM3	2	THRA	4	INHBA	8	NRCAM	1	TAS2R1	2	NDUFV2
5	LMOD1	4	DHX8	2	ZNF28	2	GP1BA	4	TGFA	8	ARHGAP22	1	NDNF	2	CDK7
5	SLC39A5	4	EIF3G	2	SEPSECS	2	CCL26	4	PDGFA	8	FUCA1	1	COL2A1	2	GTF3C3
5	GPR156	4	WWTR1	2	GCLM	2	ITPR1	4	ENG	8	FAM13A	1	CTRB2	2	NDUFA8
5	FANCA	4	SSR4	2	MCM8	2	MAF	4	PPBP	8	HEXA	1	EPHX2	2	UBE2E1
5	ATP6V0D2	4	XAB2	2	CDC26	2	CYP4B1	4	VDR	8	GPX1	1	MMP12	2	PSMD12
5	CDH15	4	HLA.DRA	2	CCDC99	2	LCP2	4	RELB	8	GCLC	1	EPHX1	2	MT2A
5	COL7A1	4	TGFBR3	2	FANCD2	2	AGT	4	PFKFB3	8	GP1BB	1	FRZB	2	PSMA7
5	DNM1	4	PSMB8	2	FBXO5	2	GLT25D2	4	MAP3K4	8	HGF	1	HPX	2	ATP5G3
5	DEFB126	4	PLK1	2	MCM10	2	GABRD	4	IL3RA	8	SLC40A1	1	NOV	2	NCL
5	SLC6A20	4	NDUFA7	2	KIF11	2	PLCG1	4	LAMB3	8	OCRL	1	SLC18A1	2	SLC25A5
5	RAMP3	4	NAGLU	2	PCNA	2	SHC2	4	HSPA2	8	GSTT2	1	CACNG2	2	SLC2A3
5	KCNJ16	4	MNAT1	2	PDE3B	2	ATP2B3	4	RAB27A	8	DEFA3	1	SERPINA4	2	PPP3CC
5	CYP3A5	4	CIITA	2	SLC30A5	2	CD8B	4	TNIK	8	DEPDC7	1	SLC1A1	2	SNRNP27

Appendix

5	MRVI1	4	GZMA	2	NEIL3	2	CYR61	4	SREBF1	9	SIGLEC1	1	THY1	2	APRT
5	GLYCTK	4	HIST3H2A	2	RBBP8	2	KCNG3	4	SPHK1	9	KCNJ9	1	FGF23	2	CDC16
5	CADM1	4	LIG4	2	NCAPG	2	FGF1	4	SLC7A11	9	LGMN	1	NTSR2	2	SF3A3
5	EIF2AK2	4	UQCRQ	2	CENPK	2	HLA.DQB1	4	SLC16A10	9	UCHL1	1	HIST1H4F	2	FARSA
5	ABCD2	4	DUSP7	2	DDX58	2	HTR1A	4	SDC2	9	CXCL12	1	IGFBP3	2	PSMD2
5	PLA2G12B	4	ETS2	2	TLR4	2	PDGFRB	4	S100A12	9	GRIK3	1	CUBN	2	COX7A2
6	CSH1	4	CCDC12	2	SKA2	2	NOS1	4	NKX3.1	9	UNC13B	1	HABP2	2	UBE2M
6	OXTR	4	HLA.B	2	GINS1	2	ST3GAL1	4	MT2A	9	DNASE2	1	PDE11A	2	DYNC1I2
6	STX6	4	NDUFS6	2	MBOAT1	2	CYP2A6	4	KIFC3	9	IDH1	1	OR10H3	2	MAN2A1
6	HCRT2	4	HGF	2	MAPK14	2	COX5B	4	IL26			1	ALDH1A3	2	NFX1
6	POM121	4	AP2S1	2	PARP1	2	FMO2	4	IL1B			1	GAD2	2	RPL6
6	CYP7A1	4	TADA2B	2	ENTPD5	2	DUSP16	4	HIST3H2A			1	LEPR	2	NAMPT
6	KCNV2	4	DAPK2	2	NUDT12	2	NFKB1	4	HIST2H2AA4			1	HIST1H2B M	2	RUVBL2
6	OR8D1	4	TAP1	2	OR2M5	2	PRKX	4	HIST1H2A E			1	ARTN	2	PRPF40A
6	UGT1A6	4	D2HGDH	2	KIF20A	2	FGF3	4	HIST1H2AD			1	ATP4B	2	DPM1
6	ITGA10	4	PSME1	2	ALG5	2	PPYR1	4	HIF1A			1	AVPR1A	2	GRPEL1
6	NMUR2	4	DDB2	2	RARS	2	NPBWR2	4	FIGF			1	BCL2	2	SDHC
6	CYP2A7	4	MCM4	2	TGS1	2	IFNA13	4	EREG			1	CLDN10	2	MLH1
6	FZD6	4	SH2D1B	2	ZNF224	2	GHRL	4	CYP1B1			1	NTNG1	2	GNB1
6	SLC9A3	4	STAT4	2	TPMT	2	HSPA6	4	CXCL1			1	SLC38A1	2	NDUFB4
6	OPN4	4	LOC646626	2	GMNN	2	IL2RA	4	CHST15			1	OR1F1	2	SRP19
6	GPR35	4	CD2	2	HIST1H2B J	2	NR3C2	4	CCL7			1	ARSB	2	ATP5L
6	MYLK3	4	PPP3CC	2	IDH1	2	NPBWR1	4	CCL23			1	IFNA13	2	CCT7
6	CHRNA2	4	HIST2H2A A3	2	ZDHHC21	2	SLC44A4	4	CCL18			1	C8B	2	CSNK2A1
6	ABCC5	4	GRAP2	2	KNTC1	2	LRAT	4	CACNA1G			1	TAAR2	2	UQCRRS1
6	CADM2	4	CD3G	2	SLC16A7	2	HCRT1	4	BIRC3			1	ITGA6	2	ASB1
6	FGG	4	IRF8	2	HERC4	2	ATF3	4	BHLHE41			1	NRCAM	2	MAPK1

Appendix

6	AGXT2	4	PSME2	2	RNGTT	2	DUSP1	4	BCL6	1	SLC6A12	2	MUT
6	CYP19A1	4	GBP2	2	PIGM	2	DUSP10	4	ACSL1	1	SEMA7A	2	FYN
6	OR7A10	4	KIR2DL3	2	TOPBP1	2	F2	4	ADM	1	PRKAR2B	2	TPR
6	GNB4	4	NCOA1	2	ITGB3BP	2	NR1I3	4	WNT5B	1	ADAMTS2	2	ATP6V1C1
6	SLC6A1	4	KIR2DL2	2	ZNF273	2	CD40LG	4	VCAN	1	C6	2	MANEA
6	CYP11B1	4	SRPR	2	ROCK2	2	DUSP9	4	THBS1	1	NEU3	2	ACTR2
6	BHMT	4	GZMB	2	ZWILCH	2	FAS	4	SLC32A1	1	IFNB1	2	DCK
6	KCNH4	4	IFNA7	2	GK	2	GRIK2	4	NCS1	1	HTR1B	2	SPCS3
6	CDH13	4	IL2RB	2	ABCA12	2	P2RX7	4	MAMLD1	1	AKR1C2	2	ALG13
6	NEFL	4	TYROBP	2	CEP70	2	CAMK2D	4	IL8	1	GNA14	2	PMAIP1
6	GSTA1	4	PRMT1	2	KIF4A	2	NR1D2	4	IL1A	1	SLPI	2	MAP3K5
6	ADH1A	4	PSIP1	2	TRMT11	2	CACNG2	4	IL17F	1	LNPEP	2	UBE2G1
6	PVR	4	BCL10	2	ZNF10	2	HTR7	4	HIST1H3D	1	CYP3A5	2	USP39
6	ATP9B	4	SKP2	2	SLC12A2	2	MAP3K4	4	HIST1H2BN	1	COL5A3	2	ATM
6	RFC5	4	NDUFA9	2	ZNF519	2	PLA2G5	4	HIST1H2BH	1	AMBN	2	TGIF2
6	SLC12A4	4	CABIN1	2	TFAM	2	STX8	4	HIST1H2BG	1	AQP7	2	PSMC6
6	HSD17B2	4	EXOSC7	2	CEP76	2	EDN2	4	HIST1H2BF	1	ACTN2	2	CTSH
6	THRB	4	FBXO5	2	MFSD8	2	PLCE1	4	HIST1H2BD	1	FZD4	2	GALNT7
6	ADH1B	4	PSMB10	2	DNA2	2	MYC	4	HIST1H2B C	1	GCG	2	CEP135
6	WNT3	4	FAM120B	2	CUL3	2	PHKG1	4	HIST1H1C	1	KCNJ14	2	APPL1
6	PLCB3	4	PDIA3	2	RPE	2	PHPT1	4	GADD45G	1	HTR2C	2	PPP1R12A
6	AVPR2	4	PTPN11	2	ZNF138	2	IL20RA	4	FLT1	1	TYRP1	2	HDAC1
6	GAL	4	PPIH	2	NUF2	2	TNNT1	4	EBI3	1	SRGAP3	2	NPEPPS
6	PDGFB	4	PIK3R5	2	SESN1	2	GCGR	4	CXCL5	1	APOC3	2	IFNGR2
6	SLC27A4	4	RPS24	2	SLC33A1	2	CYP4F8	4	CXCL3	1	ADAMTSL 3	2	TACC3
6	FBN1	4	KIR2DS5	2	MLH3	2	GHRHR	4	AGFG1	1	FGFR2	2	SUCLA2
6	OR2C1	4	GTF2H4	2	UGGT2	2	CRKL	4	CCL20	1	PPY	2	SUPT3H
6	KCNC3	4	RAC3	2	ZNF587	2	TLL2	4	HIST2H2B E	1	GNA13	2	PSMD4

Appendix

6	SLC46A1	4	ZNF385A	2	GNPNAT1	2	RLN3	4	CDK5RAP2	1	HIST1H3J	2	MPG
6	HCN4	4	TNFSF13	2	NEDD4	2	CLCF1	4	JAK3	1	TLL2	2	VPS4A
6	CACNA1B	4	CDK9	2	ZNF468	2	ADM	4	HIST3H2B B	1	HTR1F	2	POLR2B
6	COL4A2	4	RPLP2	2	HIST2H4A	2	COL23A1	4	HIST1H2BM	1	SLC5A7	2	SNRPD2
6	OR6B1	4	ADAR	2	ATR	3	GUCA1B	4	HIST1H2B L	1	CD59	2	ATP5J
6	FANCE	4	RIPK2	2	EDEM3	3	PIK3R5	4	HIST1H2BI	1	CAMK4	2	TOMM40
6	THBS4	4	PARD3	2	ZNF492	3	HK2	4	HIST1H2BK	1	VIP	2	TBCB
6	CDH8	4	YWHAH	2	EXO1	3	NGFR	4	SQLE	1	VWA5A	2	EIF4E2
6	KCNF1	4	GTF2A1	2	GTF2H2	3	TNF	4	CYP27A1	1	SLC17A1	2	NFKBIE
6	LHB	4	HLA.F	2	ZNF605	3	IL28RA	4	CASP7	1	MYC	2	OAZ1
6	CCR2	4	ASAH1	2	SGOL2	3	ITGA11	4	TCF7L2	1	CNGA3	2	GTF2H3
6	HTR1E	4	CD247	2	SOS2	3	MLN	4	AICDA	1	MEP1A	2	RASGRP1
6	CLTCL1	4	NUP37	2	CHUK	3	KCNK9	4	SDC1	1	TECTA	2	SLC25A6
6	FZD8	4	RPS6KA4	2	MOCS2	3	WNT5A	4	VAV1	1	GNAO1	2	NMI
6	SLC17A7	4	EEF2K	2	PIGF	3	RGS6	4	SRC	1	NMBR	2	POLR2F
6	BRCA2	4	SEC11C	2	NME7	3	GRM2	4	CEBPB	1	FSHB	2	RELB
6	CDH6	4	NDUFS3	2	TPTE2	3	EDA2R	4	CD44	1	LRP2	2	HLA.DRB4
6	ADRA1B	4	HLA.G	2	DGKH	3	KCNJ12	4	CXCL16	1	WNT6	2	PPP1CB
6	RIMS1	4	STK39	2	POLR2B	3	HTR2B	4	IL12A	1	SLC6A9	2	NDUFB6
6	GRIK2	4	IDH3G	2	CASP1	3	OSM	4	PPP1R3B	1	CHRM2	2	ATP5J2
6	SLC2A8	4	ZNF473	2	CASP8	3	AVPR1B	4	CYP27B1	1	CACNA1H	2	HPRT1
6	OR10C1	4	POLR2H	2	ANAPC4	3	HTR2C	4	SPRED2	1	CCR3	2	ATP5F1
6	ACHE	4	RCHY1	2	ATM	3	DARC	5	TBK1	1	PLA2G2D	2	NDUFC1
6	KCNB2	4	CD2BP2	2	MTERF	3	PFKFB1	5	PIM1	1	SLCO1A2	2	PSEN2
6	SLC2A13	4	GAB2	2	SUCLA2	3	DEFB130	5	TANK	1	COL4A4	2	CHMP4A
6	ALDH1A1	4	UQCRC1	2	ERCC4	3	GLP1R	5	NFKB1	1	DSP	2	NDUFA1
6	CHRM1	4	BANF1	2	POLH	3	MTMR7	5	SDC4	1	HSD17B2	2	PSMA6
6	GAD1	4	TSG101	2	ADHFE1	3	MTNR1B	5	NFKBIA	1	IHH	2	PTMA

Appendix

6	DAO	4	SMAD1	2	ZNF416	3	PRKCB	5	PSMA6	1	KCNA5	2	BCAS2
6	CDH7	4	PAPOLA	2	CTPS	3	S1PR5	5	MAP2K3	1	SLC18A3	2	SUGP1
6	LEP	4	PSMD2	2	RRM1	3	KCNK4	5	STAT5B	1	TNNI3	2	RPL17
6	OPRM1	4	RNPS1	2	KLHL13	3	BMP2	5	SLC1A3	1	CYP19A1	2	SMAD7
6	P2RY6	4	RPL26L1	2	OR2AG2	3	CYP2C9	5	BAK1	1	TAS2R4	2	COX5A
6	KCNH1	5	HOXA7	2	RDX	3	DEFB104A	5	DOCK4	1	PDGFC	2	NDUFC2
6	ADCY2	5	TNFRSF10C	2	ZW10	3	RGS1	5	NOLC1	1	MMP8	2	GSK3B
6	CALD1	5	EIF6	2	NSL1	3	CACNA2D2	5	PSMA7	1	FCGR2A	2	TRIM24
6	XRCC3	5	TICAM1	2	NUPL2	3	GNG7	5	TNFAIP3	1	MASP1	2	PSMF1
6	SLC2A10	5	COX8A	3	NUP210	3	RIMS1	5	CSF2RB	1	RGS11	2	AP2S1
6	SLC22A3	5	IFITM3	3	OR5D16	3	GADD45G	5	REC8	1	CHRNA3	2	CD2BP2
6	OR10J1	5	IL8	3	OR10H4	3	PROKR1	5	NFKB1B	1	PTGDS	2	SNRPD3
6	NPY1R	5	MAP3K8	3	PPP2R5A	3	ICAM1	5	PSME1	1	CDH10	2	ARHGEF6
6	KLKB1	5	TNFSF14	3	GTF2A1	3	PLA2G4E	5	CTLA4	1	KCNJ15	2	CTDP1
6	KCNJ10	5	IFITM2	3	UPRT	3	PCOLCE	5	TAP2	1	RDH8	2	DYNC2H1
6	HSD17B3	5	SDC4	3	PPT1	3	PFKFB3	5	CD40	1	SCARB1	2	YARS2
6	GP9	5	XPA	3	FBXO4	3	CTF1	5	NLRC5	1	CD14	2	NDUFAB1
6	GABRP	5	FOSL1	3	FUT8	3	CACNB2	5	UBE2L6	1	CTNNA3	2	TUBA3D
6	AVP	5	CCND2	3	ABCG5	3	COL13A1	5	TXN	1	CXCL12	2	LSM3
6	GABRA6	5	NDRG1	3	PIK3R1	3	FGF18	5	TAP1	1	HIST1H2BN	2	ANKRA2
6	SLC24A4	5	USH1C	3	CEP63	3	ITGB2	5	SOCS1	1	DKK1	2	PNP
6	ATP2A3	5	TIAM1	3	DPAGT1	3	BDKRB2	5	RIPK2	1	SERPINB2	2	GPX1
6	TGM2	5	PDK1	3	MGAT4A	3	PTGFR	5	PSMB9	1	TRIO	2	UBE2L6
6	CXCR1	5	JUND	3	SYNJ1	3	DRD2	5	PSMA3	1	NMU	2	CD96
6	F2RL3	5	HIC1	3	HNRNP1	3	ACHE	5	PML	1	IAPP	2	MSH3
6	VCAM1	5	FYN	3	TAF2	3	GNGT2	5	MCL1	1	ADRA2A	2	CLASP1
6	PLCB2	5	CR2	3	FANCC	3	CXCL2	5	LTA	1	CLEC1A	2	HERC1
6	SLC7A8	5	ERCC2	3	ERCC6	3	PENK	5	IRF1	1	PFKFB1	2	MAGOH

Appendix

6	CHEK1	5	TUBA3C	3	EDEM2	3	QRFP	5	GBP4	1	GABRA1	2	IFI6
6	OR2J3	5	PPP3CA	3	SLC16A1	3	PFKL	5	FLT3LG	1	CYBRD1	2	ATP5I
6	OR2W3	5	SHFM1	3	MTHFD2L	3	CCL3	5	GBP2	1	CCL27	2	RNF7
6	CCL2	5	NDUFA11	3	XRCC4	3	IL24	5	LMNB1	1	COLQ	2	ATP5D
6	OR2S2	5	PAK4	3	NUP93	3	CXCL9	5	ACSL5	1	OR1A2	2	POLD4
6	OR5K1	5	LIF	3	TPP2	3	PC	5	CDKN1A	1	ABCC3	2	PPIE
6	HRH2	5	CDC45	3	SLC22A3	3	FGG	5	AKT1S1	1	PTGFR	2	EIF3K
6	OR2C3	5	NFKBIE	3	FBXO11	3	RARA	5	CYP3A7	1	MMP2	2	MGAT2
6	DEFB118	5	REL	3	GTF3C3	3	CALML5	5	CD80	1	FGG	2	SART1
6	METTL2B	5	TNFRSF1B	3	ZNF33B	3	COL9A2	5	PMAIP1	1	AKR1C1	2	NUP43
6	GLRA2	5	SERPINE1	3	FAR1	3	CYP26C1	5	IFITM3	1	CCL25	2	NDUFB2
6	SSTR2	5	PPP2R5C	3	RFC5	3	NFATC4	5	IRF7	1	PDE2A	2	TBCA
6	NOS2	5	PLK3	3	KRAS	3	CYP2A7	5	TNIP2	1	FZD9	2	LSM7
6	ALB	5	PAG1	3	ZNF611	3	HRH2	5	TRAF4	1	GLUL	2	ADCY3
6	KCNAB3	5	LTA	3	ERCC5	3	DES	5	ETS1	1	WISP1	2	DAPP1
6	ATP9A	5	IRF4	3	POLR1C	3	COL5A1	5	FAS	1	GPX2	2	SRPRB
6	MTNR1A	5	IL2	3	ZNF256	3	CYP26A1	5	IL15RA	1	ITGA8	2	PSMB3
6	ATP1A4	5	HSP90B1	3	NARS2	3	AVP	5	KIF18A	1	PZP	2	SEC11A
6	GLS	5	GADD45B	3	MAN2A1	3	TNFRSF25	5	KPNA5	1	SRD5A1	2	B4GALT2
6	UGT2B15	5	CREM	3	SOS1	3	CHRNA7	5	PSMD11	1	FGF5	2	CR1
6	CALCRL	5	BCL2A1	3	OR9A2	3	GABRA6	5	RPS6KA5	1	CA1	2	RAP1A
6	OR51M1	5	CCL20	3	RFC3	3	PDE1B	6	MED12	1	RGS7	2	SNRPG
6	CSF1R	5	USP6NL	3	RAD54B	3	RGS7	6	YWHAZ	1	AQP9	2	DCTN3
6	ATP8B1	5	TRAF1	3	OCLN	3	RPS6KA2	6	PPP2R1A	1	CCL11	2	PPP3CA
6	CACNG8	5	SPRY1	3	PEX12	3	FOSL1	6	DHX15	1	ITGB6	2	TNFRSF4
7	ABCC8	5	SH2D2A	3	PIGX	3	JUNB	6	NUP153	1	PLAUR	2	PDHB
7	GNG12	5	PTPN6	3	BUB1B	3	GPR83	6	FZD6	1	HSD17B3	2	C19orf40
7	PXN	5	POMC	3	PRIM2	3	KCNJ3	6	HMGA1	1	CLDN15	2	SH2D2A
7	KCNG3	5	JUNB	3	NUP133	3	MAP4K1	6	CHPF	1	SLC30A5	2	PPCS

Appendix

7	SLC2A11	5	ICOS	3	PSMD5	3	IFNA4	6	NR1H2	1	SLC4A1	2	RPS26
7	CD4	5	FOSB	3	AMY2B	3	TAS1R2	6	SRSF4	1	GDF9	2	NAE1
7	CHRM2	5	DUSP5	3	LOC643997	3	VIP	6	PSMD5	1	SFRP4	2	OFD1
7	SLC2A5	5	DDIT4	3	CASC5	3	IL18RAP	6	ANAPC7	1	KCNB1	2	ATR
		5	BTLA	3	USP8	3	MUC15	6	CDK2	1	ITSN1	2	NUP133
		5	CD40LG	3	ERCC8	3	GJD2	6	MAP3K8	1	FCN2	2	VPS4B
		5	TNF	3	POLR3B	3	NFKB1B	6	TNFRSF18	1	HTR1E	2	ADSL
		5	IFITM1	3	PPP3CB	3	SGCG	6	SOS1	1	SLC2A9	2	PSMC4
		5	IL3	3	ZNF17	3	THY1	6	HSP90AA1	1	SLC17A7	2	CKAP5
		5	NFKBIA	3	UBE4A	3	CCL2	6	MAFF	1	ADAM20	2	U2SURP
		5	NR4A1	3	USP9X	3	COL2A1	6	XAB2	1	SCN2A	2	PIK3CA
		5	FOSL2	3	IARS	3	GDF5	6	ADRBK1	1	HIST1H3H	2	UQCRC1
		5	FOS	3	TAF9B	3	ATP1A4	6	SRSF2	1	GC	2	NDUFA10
		5	EGR3	3	PGAP1	3	CCL16	6	HNRNP	1	EDN1	2	ACAT2
		5	EGR2	3	SLC7A11	3	CD19	6	BCL2L1	1	ADCY1	2	DLD
		5	EGR1	3	MED31	3	KCNS1	6	ZNF267	1	CASR	2	PPAT
		5	BTG2	3	SNRPE	3	ATF4	6	TNFSF14	1	CYP26A1	2	MIS18BP1
		5	DOCK2	3	SLC25A32	3	SRF	6	PSMC4	1	FSHR	2	RAN
		5	ZFP36	3	TAF1B	3	CACNA1G	6	IL2RB	1	PCOLCE2	2	HSD17B10
		5	C4BPA	3	POLA1	3	ADRA2B	6	PIK3R2	1	A2M	2	CENPF
		5	MYC	3	PNPT1	3	IL6	6	SYNE2	1	KALRN	2	FAS
		5	ARHGEF18	3	CEP135	3	COL4A4	6	TAF4B	1	ADAMTS9	2	EIF2B5
		5	TUBA1C	3	RASA1	3	LTC4S	6	TNFRSF12A	1	AKT2	2	CD79A
		5	UCP3	3	RARS2	3	PMM2	6	IL23R	1	CA14	2	PFDN2
		5	YWHAG	3	ZNF567	3	CALML6	6	NFKB2	1	KCNC3	2	PSMB4
		5	ETFB	3	JAK2	3	XCR1	6	FASN	1	MXRA5	2	RPS6KB2
		5	TUBA4A	3	RASA2	3	RGR	6	ZFP37	1	MIP	2	XYLT2
		5	NDUFA1	3	CTH	3	CACNG5	6	ZFP28	1	MMP16	2	NUP62

Appendix

5	IL21R	3	PLK4	3	CSH1	6	YWHAЕ	1	GRIA4	2	UQCRQ
5	EIF4A1	3	OFD1	3	PDE6G	6	XRN2	1	IGF1R	2	IP6K2
5	TUBB2A	3	AGK	3	MUC21	6	TP53	1	KCNJ1	2	TIMM17A
5	CSK	3	ZNF624	3	GP9	6	TERF2	1	ABCC2	2	EIF2B2
5	TUBA1A	3	IDH3A	3	PLAUR	6	STAT3	1	HAL	2	ETFDH
5	HNRNPA0	3	UBR5	3	FUK	6	SNRPB	1	KCNAB1	2	GSTO1
5	BCAS2	3	ACOX1	3	MPI	6	SLC3A2	1	G6PC2	2	POLR3K
5	EGR4	3	ZNF169	3	CD37	6	SF3B4	1	AVPR1B	2	IFI30
5	ANP32A	3	PIKFYVE	3	PROKR2	6	SERPINE1	1	BMP2	2	SAMM50
5	CCR4	3	CAPN2	3	EPX	6	RDX	1	JUP	2	ARPC3
5	LIMK1	3	EPRS	3	LHB	6	RBM22	1	PLXNB2	2	LCP2
5	PTTG1	3	ALG8	3	FLT4	6	RASSF5	1	CACNA1F	2	INPPL1
5	SNRPF	3	DET1	3	RGS20	6	RASGRP1	1	PKLR	2	PSMD8
5	NDUFA13	3	LARS	3	HTR5A	6	PTK2	1	CLEC7A	2	SNRPF
5	SMAD2	3	PDHX	3	ADORA2A	6	PTBP1	1	MYL2	2	SAE1
5	HSPG2	3	REV3L	3	CXCL1	6	PSMD13	1	MYO10	2	NDUFS3
5	GTF2H2	3	RNMT	3	MMP17	6	PKM2	1	NR1D1	2	TAF11
5	PSMA7	3	CDC27	3	APOA1	6	KDM1A	1	RPS6KA6	2	HNRNPF
5	HIST1H3B	3	PHKB	3	MAP2K3	6	JAK1	1	SLC12A5	2	SF3B5
5	MAP2K3	3	ALMS1	3	COL6A1	6	ITGA5	1	OR2S2	2	DYNC1H1
5	TROVE2	3	FCGR1A	3	PYY	6	IRF9	1	IMPG1	2	ISG15
5	SH2B3	3	ATIC	3	GRID1	6	IRAK2	1	BMP15	2	SEC61B
5	BIRC2	3	CCAR1	3	PRKCG	6	IL2RG	1	CCL7	2	FBXW4
5	LIN7A	3	RGNEF	3	PLOD1	6	IL18R1	1	SEMA6D	2	NADSYN1
5	POMP	3	CDK8	3	ELK1	6	IL10	1	MEP1B	2	ADAM17
5	NDUFV3	3	SAR1B	3	KCNC2	6	IARS	1	SERPINB5	2	GMPR2
5	PSMD8	3	ZNF613	3	IL27RA	6	HSP90AB1	1	ADAMTS6	2	PPP1CA
5	MAP4K2	3	MAP3K7	3	APOC3	6	FURIN	1	GRIK3	2	GRAP2
5	BIRC3	3	MLH1	3	IL22RA1	6	EP300	1	SEMA3G	2	MIS12

Appendix

5	MAPK6	3	E2F6	3	TBX21	6	CSK	1	PRKG2	2	ACAA1
5	CD7	3	NEDD1	3	CYP4A22	6	CSF1	1	GALR2	2	TUBB2A
5	PDPK1	3	PHKA1	3	WNT16	6	CPSF7	1	AQP4	2	NDUFA7
5	RASGRP2	3	ZNF383	3	CCL22	6	CD7	1	CYP2A7	2	PSMC3
5	EIF3C	3	NUP43	3	FGF23	6	CCND3	1	CA6	2	CCNH
5	HSPA2	3	AKR1B1	3	BMP1	6	CCL1	1	KCNS3	2	TUBA4A
5	PRPF4	3	KLC1	3	CAMK2B	6	CASP2	1	NR2E1	2	NCF4
5	TAF9	3	TAF11	3	CYP11B1	6	ACTN4	1	HIST1H4D	2	NUDT21
		3	MDM4	3	BCAR1	6	ADAR	1	MAOB	2	RNASEH2A
		3	MYD88	3	KCNJ11	7	DNM1L	1	MAS1	2	UCKL1
		3	SLC38A1	3	TNNT2	7	B2M	1	SLC28A1	2	SYCP2
		3	ALG13	3	CACNA1F	7	GCK	1	GPLD1	2	COX17
		3	NMT2	3	FSHB	7	NDEL1	1	CACNG3	2	DPM3
		3	CSTF2	3	RDH16	7	SLU7	1	IFNA4	2	COX6B1
		3	PITPNB	3	GABRG3	7	H2AFX	1	PLA2G3	2	COX8A
		3	ANAPC7	3	KCNB2	7	HIST2H2AB	1	ABCC6	2	DCLRE1C
		3	TRRAP	3	COL3A1	7	HIST2H4B	1	RHO	2	AUH
		3	YARS	3	CTLA4	7	POLR3G	1	LRP6	2	NDUFA13
		3	POLR1B	3	MYLK3	7	HNRNPH2	1	BAIAP2	2	CASP7
		3	FBXO3	3	VAMP2	7	B9D2	1	LAMB3	2	UQCR11
		3	UBE2J1	3	CACNA1S	7	ADAM17	1	EPB41L3	2	CYBA
		3	H3F3B	3	CCL3L1	7	SHC2	1	MBOAT2	2	NDUFA4
		3	PIK3C2A	3	FOXP3	7	WHSC2	1	CLDN7	2	POLH
		3	SMARCA5	3	KCNQ5	7	IKBKB	1	KCNN2	2	RBX1
		3	GTF2I	3	CCL17	7	ZNF583	1	CPN2	2	NDUFB8
		3	RRM2B	3	GPR17	7	BCAR1	1	SULF1	2	SMN2
		3	MAP2K4	3	IFNA8	7	MAVS	1	TLE1	2	SHFM1
		3	SMUG1	3	ADAMTS2	7	PRKDC	1	ASH1L	2	LSM2
		3	DUT	3	KCNQ2	7	ZNF441	1	LGALS13	2	PIN1

Appendix

3	UGP2	4	ITGB7	7	NUPL2	1	SGCG	2	CHUK
3	TAF12	4	CXCR5	7	TRA2A	1	SERPINB4	2	DROSHA
3	UBR1	4	NOS3	7	SMAD4	1	TRH	2	CD72
3	WDR77	4	POMC	7	TGIF1	1	AGTR2	2	NDUFS8
3	CCT2	4	VAV1	7	PLD2	1	TGFB2	2	DUSP4
3	SLC22A15	4	CHRM1	7	PSMD1	1	MYH7B	2	EXOSC8
3	NR5A2	4	FGF4	7	CENPI	1	AGT	2	HLA.DQA2
3	ZNF184	4	GRK4	7	PPIH	1	S100G	2	BCL2A1
3	AHCTF1	4	FPR3	7	BIRC2	1	ANGPTL2	2	CTSS
3	AP4B1	4	CRHR2	7	U2AF1	1	THRB	2	NDUFS1
3	HSD17B12	4	MS4A2	7	CREBBP	1	HIST1H2A G	2	TNRC6B
3	SEC61A2	4	VAV3	7	PTK2B	1	PTPRF	2	UBE2E3
3	SEC23A	4	COL1A2	7	TCERG1	1	SST	2	BATF3
3	PARP2	4	DEFB135	7	SNAPC3	1	MFAP5	2	TNFSF10
3	XIAP	4	PFKFB4	7	DHX9	1	LEPREL1	2	RASA1
3	NCOA3	4	KCNJ10	7	DDX3Y	1	HTR3B	2	RASGRP3
3	CDK6	4	FMO3	7	TNFRSF25	1	KERA	2	CSGALNACT2
3	SLC30A7	4	DMP1	7	DNA2	1	JAM2	2	PSMD13
3	ZNF230	4	MYBPC1	7	FADD	1	GLRA1	2	OSM
3	RNF34	4	PTGER3	7	PSMA1	1	OR1A1	2	CDKN1A
3	TEAD1	4	KLKB1	7	UGCG	1	SCN1B	2	EIF1AX
3	NFX1	4	COX7A1	7	NEDD1	1	TNFSF18	2	NUPL1
3	GNPTAB	4	CYP3A5	7	ZNF569	1	CYP2E1	2	SKP2
3	PREB	4	CXCR7	7	CAB39	1	RARA	2	FANCG
3	PDP2	4	GPX5	7	SLC38A2	1	BRS3	2	IFI35
3	PSMB2	4	FZD7	7	RIPK1	1	GNRHR	2	TNFRSF14
3	SEC31A	4	GRM1	7	U2AF2	1	CHRM5	2	HLA.DPB1
3	ANAPC5	4	HLA.DPB1	7	PYGM	1	CYP26B1	2	PSMD3

Appendix

			3	SLC9A9	4	FGF16	7	ZNF20			1	NTF3	2	PRKCZ
			3	PIAS2	4	IFNB1	7	EXOG			1	CAP2	2	TRAF3
			3	ZNF544	4	PCK1	7	CREB1			1	FMO4	2	MAD1L1
			3	CUL1	4	CCBP2	7	CCNE1			1	ATP10B	2	SRPR
			3	ZNF426	4	EDNRB	7	TAF4			1	CLDN17	2	PSMB10
			3	DLD	4	TNFSF18	7	FOXO1			1	F2RL3	2	USP18
			3	NPC1	4	CCL23	7	HNRNPM			1	GABRE	2	ETFB
			3	SLC9A6	4	P2RY8	7	ZNF347			1	VIPR1	2	ITPR1
			3	RHOA	4	STX1B	7	SELP			1	GUCY2F	2	PTPRK
			3	ZNF431	4	CHRNA3	7	CD28			1	DOCK1	2	RPA3
			3	ZNF37A	4	CXCL5	7	RPS6KB1			1	LGALS1	2	FXN
			3	TRIM32	4	DEFB136	7	HIST1H4F			1	LEFTY1	2	MBD4
			3	CUL2	4	CD1A	7	HIST1H4I			1	MYL1	2	POMP
			3	SLC44A1	4	KL	7	ZNF442			1	GNAT3	2	BIRC3
			3	TEX15	4	CD38	7	SLC43A2			1	CBLN1	2	CD79B
			3	OR1J2	4	GPR65	7	SDCCAG8			1	COL10A1	2	UBE2S
			3	PIK3C2G	4	FGB	7	TAF1A			1	CNTN1	2	DFFB
			3	PAPSS2	4	IFNA21	7	RB1			1	KCNJ2	2	JUN
			3	SLC5A7	4	TLR1	7	ADA			1	F2	2	HIST1H4C
			3	OR9K2	4	TLR2	7	ZNF582			1	KCNMB3	2	CREM
			3	ZNF615	4	CYP2C19	7	ZNF101			1	FGA	2	UBE2V2
			3	SLC25A4	4	RHCG	7	TCP1			1	GJB3	2	CSGALNACT1
			3	UPF3B	4	CYP24A1	7	NUP205			1	PROL1	2	TJP2
			3	OR52E8	4	MUC19	7	RFFL			1	CDH7	2	CHMP2A
			3	PIK3CA	4	PDE6A	7	ATF2			1	ABCC9	2	ICOS
			3	PDHB	4	AWAT2	7	ZNF567			1	TIMP2	2	KIF2C
			3	MED17	4	FGA	7	PSMC6			1	ESRRG	2	KPNA2
			3	OR4M2	4	KCNA2	7	SNAPC1			1	HIST2H2BE	2	DBF4
			3	CDC16	4	BST1	7	ZNF224			1	PRKG1	2	MAVS

Appendix

3	RAD17	4	IL5	7	KAT2A	1	C5	2	POLR2K
3	WWP1	4	FSHR	7	NUP107	1	VWF	2	CDK8
3	MED7	4	DUSP8	7	PCF11	1	CYP2C18	2	RFXANK
3	ZNF382	4	MCHR2	7	ZNF445	1	SERPIND1	2	XPA
3	UGT2B4	4	CXCL14	7	ZNF552	1	ADAM29	2	BLNK
3	ZNF415	4	KCNA7	7	GSK3A	1	ADCY6	2	CASP4
3	FIG4	4	PF4	7	USP39	1	CNTN2	2	HERC2
3	PDC	4	CYP4A11	7	ZNF708	1	PLCB1	2	IRF8
3	LIPF	4	ICAM4	7	ZNF430	1	PMCH	2	CDC20
3	ZFP112	4	NCAM1	7	ZNF417	1	S100B	2	TIMM13
3	PRKDC	4	MMP10	7	ZNF354A	1	UGT2B28	2	CKS1B
3	ATP6V0A1	4	ADH1B	7	TDP2	1	FMO1	2	SYNE2
3	BLMH	4	GPX2	7	RUNX1	1	TAAR5	2	DHFR
3	L2HGDH	4	NTF4	7	RAD21	1	PARD6A	2	GBP2
3	PPP2R5C	4	CPE	7	PRKAG2	1	PLA2G2E	2	REV1
3	SLC7A6	4	PTGIS	7	MED26	1	UNC5B	2	FUS
3	AP3M1	4	SLC1A2	7	MAPK8	1	EDNRA	2	MAP4K3
3	OR2A4	4	IL2RG	7	IL4	1	GRIP2	2	POLR2I
3	CPSF3	4	NTS	7	FBXO5	1	WNT2B	2	BAG4
3	BTRC	4	CYP2E1	7	EXOC6	1	CYP2B6	2	TCEB1
3	RAD1	4	CGA	7	CHERP	1	WASF3	2	CHMP3
3	ZNF619	4	HTR2A	7	CEP135	1	NR4A3	2	CCL17
3	LY96	4	GLRA3	7	CENPT	1	F7	2	HDAC4
3	ZNF510	4	MMP7	7	CCT4	1	GJD2	2	CCL22
3	MTM1	4	GRAP2	7	CDK8	1	OPRK1	2	LONP2
3	PIK3R4	4	MUC12	7	CLIP1	1	IL37	2	GTF2A2
3	SEC24D	4	AGTR2	7	EIF4E3	1	CYP17A1	2	PCCB
3	UBE4B	4	INHBC	7	CXCL2	1	HTRA1	2	REV3L

Appendix

			3	RCHY1	4	CR1	7	NUP43		1	TFPI	2	HNRNPH2
			3	UBE2A	4	CXCL13	7	EGLN1		1	SLC6A15	2	PTPN2
			3	DOCK2	4	UGT2B28	7	PRKACB		1	CYP7B1	2	VBP1
			3	PDPR	4	ADH4	7	DOCK5		1	NTS	2	BANF1
			3	EEF1E1	4	NPY5R	7	PLD1		1	LTF	2	IL7
			3	ALG14	4	CYP2B6	7	HIST1H2AJ		1	PPBP	2	SNAPC5
			3	ZNF540	4	KDR	7	STX1A		1	CCKBR	2	DCTN2
			3	OR4K15	4	FBP1	7	HK2		1	LPAR1	2	EBI3
			3	PRKACB	4	NPFFR2	7	PDE1B		1	RORB	2	C1D
			3	ASB17	4	RPE65	8	CYCS		1	CYP2J2	2	LY96
			3	PDK3	4	CYP3A4	8	HIST1H3F		1	CSH2	2	ISG20
			3	UBA5	4	IL21	8	CAMKK1		1	SLC1A7	2	TBCE
			3	EIF4E3	4	TRPC6	8	CREB5		2	VEGFC	2	NT5C2
			3	RNF7	4	CACNA1E	8	CDKN2B		2	GRIA2	2	UGCG
			3	B4GALT6	4	DEFB128	8	HCK		2	HIST1H2AI	2	PAPOLA
			3	RBL2	4	MMP13	8	GTF2E2		2	FPR1	2	TOMM22
			3	CENPJ	4	ADORA3	8	KIF5B		2	SLC16A8	2	CYFIP2
			3	IARS2	4	GIP	8	PRPF18		2	PLA2G2F	2	UQCRC2
			3	CENPC1	4	KCNA4	8	SC5DL		2	SLC24A2	2	NDUFB1
			3	PPM1D	4	DEFB115	8	TRAF3		2	DGKI	2	ACP5
			3	RBMX	4	SLC6A14	8	RAPGEF1		2	NRP2	2	L2HGDH
			3	YES1	4	MYH8	8	SMPD2		2	SERPINA1	2	TIMM8B
			3	ZNF649	4	ALDOB	8	CSNK1D		2	ALDH1A1	2	SNRPE
			3	UBR2	4	DEFB114	8	MCM6		2	GLS2	2	ACADM
			3	ALG2			8	SORBS1		2	ATP1A2	2	SSB
			3	APAF1			8	CHST11		2	PARM1	2	NCBP1
			3	NUP205			8	HLX		2	IL3	2	CTNNBL1
			3	GBE1			8	CAMKK2		2	RHOB	2	GALNT1
			3	TRAF6			8	CDK1		2	CTNND1	2	RPL26L1

Appendix

3	HECTD3	8	SGOL1	2	MUC16	2	PLCB3
3	ENTPD4	8	CD2BP2	2	MUC5B	2	PSMA5
3	SLC9A2	8	PIK3AP1	2	CREBBP	2	TERF1
3	ZNF26	8	PPP2CA	2	KLK2	2	TLR1
3	EVI5	8	AP3M1	2	PTPN1	2	NARS2
3	SLC35D1	8	PRKAR1A	2	C9	2	AHCTF1
3	MED8	8	NUP188	2	SOX17	2	NDUFS4
3	ZNF254	8	PGK1	2	S100A12	3	ACAA2
3	TBP	8	CASP8	2	CDH3	3	CSTF2
3	MDM2	8	JUNB	2	TNR	3	QARS
3	NFIB	8	PSME2	2	TNC	3	ERCC2
3	PI3	8	MDM2	2	OPRL1	3	MLST8
3	TARSL2	8	SNRNP40	2	PLG	3	PEBP1
3	AP1G1	8	CCNA2	2	PSPN	3	IL13
3	HACL1	8	RPS6KB2	2	TLN2	3	RAC3
3	ESR2	8	NCOA4	2	ADRBK1	3	BCR
3	ZNF300	8	CAB39L	2	ATP2B3	3	SRP72
3	NUP155	8	CSF2RA	2	AQP6	3	NFATC3
3	CTSO	9	MAP3K5	2	CHRD	3	CREB3L2
3	DHX9	9	SLC38A5	2	GAS1	3	RPS6KA5
3	LNPEP	9	PI3	2	MAPK10	3	RHOC
3	SACM1L	9	LDHA	2	CYP24A1	3	IRAK1
3	ZNF350	9	SLC7A5	2	GLP2R	3	RPN1
3	CDC25C	9	GZMB	2	STARD5	3	DUSP6
3	HNRNPR	9	TIAM2	2	TMOD1	3	TIMM10
3	SLC22A5	9	TRAF1	2	PRKAG2	3	CUL4A
3	TFDP1	9	ITGA1	2	SSTR4	3	HTRA2
3	ATP6V0A2	9	PTPN2	2	FZD2	3	C17orf70

Appendix

3	CHEK1	9	FLOT1	2	SLC1A2	3	DDX42
3	EIF2S1	9	SOCS3	2	VDR	3	CD3G
3	NF1	9	IFITM1	2	C7	3	SEC61G
3	BIRC6	9	ISG20	2	SLC29A3	3	HNRNPUL1
3	CALM2	9	IL18RAP	2	GRM7	3	SSR2
3	PIP4K2A	9	IL12RB2	2	ABCG5	3	XRCC1
3	OR2T6	9	SLC1A1	2	SLIT1	3	FBXW7
3	SETD7	9	ZNF254	2	CACNA1S	3	PDIA3
3	CTNNB1	9	XAF1	2	ITGA10	4	MTX1
3	OR2A7	9	TNFRSF6B	2	KCNH6	4	NUP85
3	OR5AS1	9	PTGS2	2	HIST1H2BE	4	CSF2RB
3	ALDH9A1	9	PLAT	2	HGFAC	4	MAPKAPK3
3	MCM6	9	MAPK11	2	HIST1H2B G	4	BMPRI1A
3	RANBP2	9	IL6	2	GJB5	4	DDOST
3	ACSL1	9	IL2RA	2	EDN2	4	HLA.E
3	TAF13	9	IL24	2	FTH1	4	GMPPB
3	NUP107	9	IL19	2	PLCD1	4	MAP3K8
3	GARS	9	IFNG	2	COL7A1	4	ITGA1
3	SMC3	9	DCN	2	CACNB2	4	NOD2
3	FANCI	9	CCL19	2	GRM2	4	PPP1R14B
3	CPSF2	9	CRLF2	2	THBD	4	TYMP
3	OR9I1	9	CFB	2	NR1I3	4	CLIP1
3	SLC1A1	9	ITGA3	2	WNT11	4	DCTN6
3	GTF2E2			2	EHMT2	4	PTPN7
3	PSMA4			2	MMP27	4	SH2B3
3	GTF2H1			2	ABCA7	4	ERCC3
3	LCLAT1			2	ATP2A1	4	PSME2
3	FARSB			2	SLC3A2	4	PPA2

Appendix

3	GSK3B	2	CELA3B	4	SSBP1
3	MED4	2	HIST2H2A A4	4	PSMB9
3	RNF4	2	FGF7	4	UNG
3	ERAP1	2	CRK	4	DTYMK
3	PIK3C3	2	ACE2	4	EXOSC10
3	POLR2D	2	GDNF	4	NDUFS6
3	CEP192	2	SLC38A3	4	ALG5
3	XYLB	2	SLC13A2	4	GRK6
3	SLC35A3	2	SERPINB13	4	PMM2
3	ZNF432	2	P2RX1	4	BARD1
3	SDCCAG8	2	BTRC	4	PPCDC
3	CDCA8	2	HSD17B1	4	EGR2
3	KLHL9	2	LGALS4	4	RGS1
3	OCRL	2	SLC27A6	4	CCL3L3
3	PPP1CC	2	LAMB4	4	IL2RA
3	ZNF200	2	CALM1	4	TNF
3	AGPAT5	2	SPARC	4	STAM
3	RACGAP1	2	F3	4	CHST11
3	MAN1A1	2	GABRB3	4	DDB2
3	CCT8	2	GSTM5	4	MX1
3	DARS2	2	PDE8B	4	GALNT6
3	SF3A3	2	SSTR1	4	GZMA
3	MAPK8	2	HNF4G	4	SF3B3
3	NUDT21	2	RGS6	4	SRSF2
3	TARS	2	CLDN8	4	IARS2
3	CENPL	2	CDH12	4	UBE3C
3	CP	2	COL6A1	4	BLM
3	SLC35A1	2	COL5A1	4	MIS18A

Appendix

3	HNRNPA2B1	2	FPR2	4	GMNN
3	SLC35B2	2	PPP2R1B	4	POLA2
3	NCBP1	2	THBS2	4	CDC6
3	RBBP7	2	ANXA5	4	IFNG
3	WARS	2	EMR2	4	EIF4G3
3	ZNF140	2	CHRNA4	4	CDK2
3	NNT	2	FGFR3	4	FOXM1
3	PSMC6	2	PTCH1	4	AK4
3	PPIL2	2	HTR2A	4	DSN1
3	RUVBL1	2	MYBPC2	4	PSMD1
3	ZNF266	2	ST8SIA2	4	ANAPC1
3	CCNH	2	COL14A1	4	HLA.DMA
3	PARP4	2	FGFBP1	4	SLC25A16
3	FAR2	2	GFRA4	4	WBP11
3	NT5E	2	TGFB3	4	JUNB
3	PTS	2	TGM3	4	ZBTB32
3	UBE3B	2	IL11	4	TUBB3
3	ALG10	2	OR2J2	4	CDC37
3	SEH1L	2	FGF17	4	CHMP6
3	PEX1	2	GRIK5	4	SMC1A
3	PIGU	2	ATP10A	4	MYBL2
3	KPNA2	2	CDK5R1	4	FUT8
3	XPO1	2	EGFR	4	CDT1
3	NCK2	2	KCNA10	4	PFAS
3	SF3B3	2	ADRA1B	4	CDC45
3	PTK2	2	MMP15	4	TK1
3	SNRNP40	2	GNAZ	4	MCM6
3	PSMD1	2	BDKRB2	4	POLR2E
3	CDC40	2	GBA3	4	VDAC1

Appendix

3	EIF3A	2	GHRH	4	CBX5
3	RB1	2	NRTN	4	MEF2A
3	ZFP90	2	SPON2	4	RFC5
3	HERC5	2	SYT5	4	RPA1
3	ZNF33A	2	GSK3A	4	XRCC4
3	MARS2	2	ENAH	4	SEC63
3	HIST1H2BK	2	NRG1	4	CCNE1
3	KPNB1	2	PCSK1	4	TOPBP1
3	TAF1A	2	CHRNA9	4	NUP37
3	PSMA2	2	ATP9A	4	EP300
3	TUBA4A	2	GRM3	4	FBXO5
3	ATG7	2	HGF	4	GTF2H5
3	TUBGCP3	2	F13B	4	IL1R2
3	CCNT1	2	ITGA9	4	RRM1
3	FBXW11	2	MDK	4	ANAPC10
3	SMURF2	2	PLOD3	4	MSH2
3	CEP57	2	HIST1H3E	4	TPK1
3	SLC38A2	2	COL4A3	4	HJURP
3	PIGS	2	GDF10	4	RSF1
3	PSMD14	2	CXCL1	4	SMAD2
3	E2F3	2	BMP6	4	SNRPB
3	NUP88	2	GRIN2D	4	SLC19A2
3	PLK1	2	NR2F1	4	TAP1
3	PPP1R13B	2	WNT10B	4	NFKBIA
3	ERCC3	2	DMD	4	LIG1
3	SDHA	2	INSL6	4	MCM5
3	ZNF621	2	MRAS	4	KIF3B
3	GRB2	2	CLTC	4	RNASEH1
3	TUBA1B	2	PRKACA	4	PSMB7

Appendix

3	CDKN1B	2	PLA2G4A	4	TNFRSF1B
3	UBE3C	2	CDH4	4	PRKX
3	HIST1H4H	2	COL8A2	4	CDC7
3	MAPK1	2	GJA4	4	COQ2
3	SLC36A1	2	PTGER1	4	PIGB
3	MIS12	2	ACAN	4	MGMT
3	PSME4	2	HSD3B2	4	TFDP1
3	RAN	2	PLXND1	4	TUBGCP5
3	EIF4G3	2	TNNC1	4	UPF3A
3	POLR2G	2	CDX2	4	EGR3
3	RAE1	2	GRIA3	4	TFRC
3	AIMP1	2	INHBA	4	IFT57
3	EIF1AX	2	KCNH2	4	KIF23
3	SDHD	2	GRIN2A	4	SPC25
3	MAX	2	GPR39	4	SMC2
3	TAF5	2	CHRND	4	TUBGCP3
3	PIGQ	2	PTPRB	4	GBP1
3	VCL	2	TSHB	4	SRSF10
3	OR5A2	2	THSD4	4	ECT2
3	UBE2G1	2	ARHGEF17	4	NCAPG
3	NCBP2	2	STS	4	DLGAP5
3	RBL1	2	BMP8A	4	KIF15
3	RBPJ	2	SDC3	4	HLA.DRA
3	PIGT	2	SLC6A2	4	NDC80
3	TCP1	2	KCNA1	4	SKA1
3	TUBA1A	2	FSTL1	4	PRPS2
3	TUBGCP5	2	HNF4A	4	LTA
3	QARS	2	CDH15	4	GINS2
3	LMAN1	2	MFAP3	4	CXCR3

Appendix

3	ZNF154	2	TNFAIP6	4	KIF11
3	METTL6	2	RHAG	4	FANCI
3	PSMA3	2	PCOLCE	4	BRIP1
3	MANEA	2	SEMA6A	4	FEN1
3	STAT1	2	CHRNA6	4	ITGB3BP
3	FH	2	ECM1	4	TTK
3	SLC4A7	2	SERPINC1	4	POLE2
3	NUP37	2	CAV3	4	UCK2
3	METTL3	2	CCL16	4	BUB1B
3	SMC2	2	GJC1	4	PLK4
3	DHFR	2	AREG	4	IL18RAP
3	ZNF551	2	C8A	4	MCM7
3	CSNK1G3	2	CD46	4	ZWINT
3	MBD2	2	GP1BA	4	TIPIN
3	TRIM37	2	PRSS2	4	IL18R1
3	ABCC4	2	GAST	4	CDK1
3	SUPT16H	2	PRKCG	4	RRM2
3	RFC1	2	AGPAT1	4	CCNE2
3	SKA1	2	CRLF1	4	CHEK1
3	NDC80	2	COL17A1	4	CHSY1
3	TAS2R5	2	ELK1	4	FANCL
3	RAP1B	2	FGF9	4	RFC2
3	TAF9	2	ARSJ	4	MAD2L1
3	ITCH	2	FGF18	4	SMURF2
3	RAD51C	2	CHRNA6	4	TCERG1
3	POLR3F	2	TNFSF8	4	KNTC1
3	CCT6A	2	CYP4B1	4	RBL1
3	LAMP3	2	C4B	4	GINS1
3	ELOVL6	2	ADH6	4	GZMB

Appendix

				3	PPP2R1B					2	TINAGL1	4	XCL2	
				3	CUL5					2	SLC5A2	4	CENPN	
				3	SLC25A16					2	HTR1D	4	BCOR	
				3	PAPOLA					2	LEFTY2	4	PRC1	
				3	SLC35B4					2	DPT	4	AURKB	
				3	CCNC					2	FSTL3	4	BRCA1	
				3	NUP50					2	LPCAT4	4	MCM2	
				3	MTFMT					2	PROS1	4	APITD1	
				3	PALB2					2	APOA1	4	ADK	
				3	ZNF192					2	ARHGAP35	4	ZWILCH	
				3	AHCYL1					2	BMP4	4	CCNA2	
				3	BARD1					2	OR7C2	4	OAS3	
				3	AKAP9					2	SLC6A6	4	ORC1	
				3	AP4E1					2	TAC3	4	DUT	
				4	ALG6					2	GABRA5	4	HLA.DMB	
				4	ZNF189					2	SLC28A2	4	MTHFD1	
				4	SYCP2					2	ATP2A2	4	POLA1	
				4	ZNF214					2	LAMB1	4	GPHN	
				4	OR13C4					2	KCNAB3	4	WARS2	
				4	OR5T1					2	P2RX6	4	CRTAM	
				4	SLC1A6					2	SRC	4	CXCL9	
				4	OR6C3					2	CTNNA1	4	RFC3	
				4	OR8H3					2	GNAI1	4	KLRC2	
				4	OR5I1					2	GABRG2	4	MCM10	
				4	OR2A12					2	MMP17	4	ETF1	
				4	PI4K2B					2	IFNA6	4	SNRPD1	
				4	PNLIPRP2					2	PRLR	4	ORC6	
				4	MOGS					2	PXDN	4	ARPC4	
				4	CD28					2	GFRA2	4	FLT3LG	

Appendix

4	LBP	2	SCN8A	4	UMPS
4	OR4K2	2	GJA9	4	POLR3E
4	OR6C4	2	MYLK3	4	RARS
4	OR5M3	2	COL4A1	4	EIF4H
4	OR4C3	2	TACR1	4	IFITM2
4	CD36	2	FMOD	4	PPIH
4	SLCO1B1	2	KCNG1	4	APEX1
4	OR6C76	2	LAMB2	4	CDC25A
4	SLCO1C1	2	NR2E3	4	TRRAP
4	OR5J2	2	CA5A	4	SLC25A12
4	ABCC9	2	CYP2U1	4	LMNB2
4	PEX11B	2	CNGA1	4	ARPC1B
4	OR52A5	2	EFEMP2	4	UCHL5
4	OR14J1	2	SULT4A1	4	CUL5
4	OR1A2	2	STX1A	4	GALNT2
4	OR4C13	2	UGT1A9	4	ESPL1
4	OR4C16	2	ACSM1	4	EXOSC9
4	SLC38A4	2	ADH1B	4	AK2
4	ZNF552	2	GPR17	4	MTX2
4	ABCD2	2	GCLC	4	WARS
4	ZNF599	2	PODNL1	4	ACAT1
4	AMY2A	2	PRKACB	4	RAP1GAP2
4	UIMC1	2	FZD1	4	FOXO1
4	LIPC	2	MAOA	4	MS4A1
4	OR5R1	2	SMO	4	PSMA3
4	ZNF708	2	SPHK1	4	BID
4	ZNF566	2	CASC5	4	CEP290
4	ZNF14	2	TNFSF15	4	TLE4
4	ZNF506	2	FGD2	4	RMI1

Appendix

4	OR2T29	2	PDE6A	4	RUVBL1
4	TAS2R8	2	TAS2R10	4	NNT
4	OR2T5	2	TYR	4	PSMA4
4	OR4K5	2	SCNN1G	4	DOCK2
4	OR2F2	2	COLEC11	4	STAT1
4	OR9G1	2	NYX	4	SNRPA1
4	AMY1A	2	PRLH	4	STK39
4	ASB15	2	CACNA1B	4	CENPQ
4	OR1D4	2	DAG1	4	RCOR1
4	OR51G1	2	RAE1	4	CUL2
4	OR8K1	2	CTSB	4	DNA2
4	SLC7A9	2	PDE4C	4	NME1
4	OR2J3	2	CRH	4	DLG1
4	OR8I2	2	CXCL2	4	ETFA
4	UGT2A1	2	SDC1	4	CDKN2C
4	OR2L8	2	SULT1A4	4	GTF2F2
4	ROCK1	2	ACTA2	4	MAPK9
4	OR4F17	2	NUP88	4	NHP2
4	UGT2B11	2	C3	4	EIF5B
4	OR4M1	2	CYR61	4	POLD2
4	OR5B2	2	CCL2	4	PSMB8
4	OR7A10	2	LRP1	4	TUBG1
4	OR8B3	2	COL1A2	4	GNG10
4	ACP5	2	NUP188	4	PRIM1
4	CEBPA	2	RXRG	4	ZEB1
4	ZNF20	2	MMP10	4	ERCC6L
4	ZNF490	2	KCND1	4	CCR6
4	OR1C1	2	ZP2	4	PGAP1
4	OR5T2	2	PTPRJ	4	TFAM

Appendix

	4	SLCO1A2						2	PARVA	4	ACTL6A
	4	ZNF563						2	PDE1A	4	CXCL10
	4	ELL						2	CHRNE	4	UBE2C
	4	OR13C3						2	ARSE	4	DPYD
	4	OR5W2						2	MUC5AC	4	CCL20
	5	ERCC6L						2	AP2B1	4	CCNB1
	5	RNF144B						2	GRID2	4	PSMA2
	5	ZNF479						2	ADCY9	4	CCNB2
	5	LYPLA1						2	FRAT1	4	KIF20A
	5	PMAIP1						2	P2RX2	4	NEK2
								2	ADCYAP1R1	4	RACGAP1
								2	NR1D2	4	SGMS1
								2	CTGF	4	JAK2
								2	ERBB4	4	OIP5
								2	EMILIN1	4	XYLT1
								2	TPMT	4	SMS
								2	SEMA3F	4	ENTPD1
								2	CYP2F1	4	LYN
								2	DNM1	4	RBBP8
								2	PF4	4	IL26
								2	CELA3A	4	SLC25A32
								2	GLRA3	5	PIGH
								2	CLTA	5	TNFSF13
								2	GJA1	5	ITGB2
								2	SCTR	5	IL5
								2	PYGO1	5	IL4
								2	H2AFX	5	TNFRSF11A
								2	SPOCK3	5	GAB2

[illegible]

[illegible]

	2	CACNG1	
	2	OR2F2	
	2	CDH6	
	2	OPLAH	
	2	CYP3A4	
	2	GREM1	
	2	DES	
	2	PAPPA	
	2	GRIK2	
	2	NR1I2	
	2	CACNA2D3	
	2	COL9A1	
	2	MMP9	
	2	CYP21A2	
	2	FCN1	
	2	KNG1	
	2	MTNR1A	
	2	HTR5A	
	2	KCNF1	
	2	AGR1	
	2	RRAS	
	2	ALDH3B2	
	2	AR	
	2	OR10H2	
	2	HIST2H4B	
	2	PLCB2	
	2	PNLIP	
	2	PDE1B	
	2	PFN2	

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Appendix

												2	IGFBP7		
												2	APLNR		
												2	CYSLTR1		
												2	CXCL11		
												2	KCNH4		
												2	NCAN		
												2	CLDN11		
												2	DRD3		
												2	SEMA3C		
												2	DAGLA		
												2	CAMK2B		
												2	RGR		
												2	DGKG		
												2	AQP5		
												2	HRH2		
												2	LAMA5		
												2	ITGA5		
												2	VEGFA		
												2	IMPG2		
												2	MCHR1		
												2	OR7E24		
												2	SCNN1B		
												2	CRHBP		
												2	KCNN1		

Appendix

Table A.4. Functional Enrichment of PPA networks

EXPERIMENT: E-GEOD-4209, Downregulated Canonical Pathways						
Maximum stable clusters : 5						
Functional Enrichment p -value <10⁻⁵						
cluster	KEGG pathway	p-value	# of proteins	Reactome pathway	p-value	# of proteins
1	Spliceosome	1.759e ⁻²³	31	Spliceosomal B complex	1.410e ⁻¹⁷	24
1	T cell receptor signaling pathway	1.759e ⁻²³	29	Formation of the Spliceosomal B complex	1.410e ⁻¹⁷	24
1	Metabolic pathways	1.209e ⁻¹⁶	70	mRNA splicing- major pathway	3.230e ⁻¹⁷	25
1	<i>Neurotrophin signaling pathway</i>	9.749e ⁻¹⁵	23	Spliceosomal active C complex with lariat containing 5'-end	1.190e ⁻¹⁶	24
1	<i>Natural killer cell mediated cytotoxicity</i>	2.729e ⁻¹³	22	cleaved pre-mRNP: CBC complex		
1	RNA transport	3.880e ⁻¹³	23	Lariat formation and 5'- splice site cleavage	1.190e ⁻¹⁶	24
1	<i>Chronic myeloid leukemia</i>	4.599e ⁻¹³	17	Formation of an intermediate spliceosomal C complex	1.239e ⁻¹⁶	24
1	<i>Fc epsilon RI signaling pathway</i>	1.800e ⁻¹²	17	Exon junction complex	1.239e ⁻¹⁶	24
1	<i>Measles</i>	2.279e ⁻¹²	21	Cleavage at the 3'-splice site and exon ligation	1.239e ⁻¹⁶	24
1	Cell cycle	4.360e ⁻¹²	20	Spliceosomal intermediate C complex	7.859e ⁻¹⁶	23
1	B cell receptor signaling pathway	1.009e ⁻¹¹	16	Spliceosomal active C complex	7.859e ⁻¹⁶	23
1	Pyrimidine metabolism	1.049e ⁻¹¹	18	mRNA splicing-minor pathway	4.580e ⁻¹²	14
1	Huntington's disease	1.730e ⁻¹¹	23	ATAC B complex	4.580e ⁻¹²	14
1	<i>Pancreatic cancer</i>	1.799e ⁻¹¹	15	ATAC C complex with lariat containing 5'-end cleaved mRNA	4.580e ⁻¹²	14
1	Pathways in cancer	1.799e ⁻¹¹	30	<i>ATAC C complex</i>	4.580e ⁻¹²	14
1	Oocyte meiosis	3.780e ⁻¹¹	18	<i>Formation AT-AC C complex</i>	4.580e ⁻¹²	14

Appendix

1	<i>ErbB signaling pathway</i>	6.739e ⁻¹¹	16	ATAC spliceosome mediated 3' splice site cleavage, exon	4.580e ⁻¹²	14
1	<i>Pathogenic Escherichia coli infection</i>	1.949e ⁻¹⁰	13	ligation		
1	<i>Fc gamma R-mediated phagocytosis</i>	2.210e ⁻¹⁰	16	Spliceosomal A complex	4.809e ⁻¹²	18
1	Focal adhesion	4.780e ⁻¹⁰	22	<i>Transport of mature Mrna derived from an Intron-containing</i>	4.809e ⁻¹²	15
1	Purine metabolism	4.980e ⁻¹⁰	20	<i>Docking of the TAP:EJC complex with the NPC</i>	7.590e ⁻¹¹	14
1	Alzheimer's disease	5.980e ⁻¹⁰	20	RNA polymerase 2 pre-transcription events	1.080e ⁻⁸	13
1	Chemokine signaling pathway	6.870e ⁻¹⁰	21	Degradation of multiubiquitinated Securin	5.860e ⁻⁸	13
1	MAPK signaling pathway	7.090e ⁻¹⁰	25	Formation of the Spliceosomal E complex	5.860e ⁻⁸	13
1	<i>Influenza A</i>	8.890e ⁻¹⁰	20	Spliceosomal E complex	5.860e ⁻⁸	13
1	<i>Regulation of actin cytoskeleton</i>	8.890e ⁻¹⁰	22	Degradation of multiubiquitinated cell cycle proteins	1.290e ⁻⁷	13
1	Parkinson's disease	3.359e ⁻⁹	17	<i>APC/C:Cdh1 mediated degradation of Cdc20 and other</i>	1.509e ⁻⁷	13
1	Nucleotide excision repair	6.139e ⁻⁹	11	APC/C:Cdh1 targeted proteins in late mitosis/early G1		
1	<i>Glioma</i>	1.480e ⁻⁸	12	snRNP assembly	1.790e ⁻⁷	11
1	<i>Non-small cell lung cancer</i>	3.019e ⁻⁸	11	<i>Release from the NPC and disassembly of the mRNP</i>	2.519e ⁻⁷	10
1	<i>Hepatitis C</i>	3.369e ⁻⁸	16	Degradation of multiubiquitinated Cdh1	2.820e ⁻⁷	12
1	<i>Progesteron-mediated oocyte maturation</i>	4.540e ⁻⁸	13	Autodegradation of Cdh1 by Cdh1:APC/C	2.820e ⁻⁷	12
1	<i>Renal cell carcinoma</i>	4.799e ⁻⁸	12	<i>RNA pol 3 simple start sequence initiation at type 3</i>	3.060e ⁻⁷	9
1	<i>Acute myeloid leukemia</i>	5.009e ⁻⁸	11	<i>promoters</i>		
1	<i>Jak-STAT signaling pathway</i>	5.359e ⁻⁸	17	<i>RNA pol 3 promoter opening at type 3 promoters</i>	3.060e ⁻⁷	9
1	<i>Leukocyte transendothelial migration</i>	2.179e ⁻⁷	14	<i>RNA pol 3 transcription initiation from type 3 promoter</i>	3.060e ⁻⁷	9
1	<i>Toxoplasmosis</i>	2.320e ⁻⁷	15	<i>RNA pol 3: TF3B:SNAPc:type 3 open promoter complex</i>	3.060e ⁻⁷	9

Appendix

1	<i>GnRH signaling pathway</i>	2.990e ⁻⁷	13	<i>RNA pol 3 :TF3B:SNAPc:type 3 promoter complex</i>	3.060e ⁻⁷	9
1	Ubiquitin mediated proteolysis	2.990e ⁻⁷	15	Intron-containing complex	3.590e ⁻⁷	10
1	<i>Toll-like receptor signaling pathway</i>	3.730e ⁻⁷	13	Formation of RNA pol 2 elongation complex	4.529e ⁻⁷	10
1	Osteoclast differentiation	5.770e ⁻⁷	14	Formation of HIV-1 elongation complex in the absence of	4.529e ⁻⁷	10
1	<i>Colorectal cancer</i>	7.190e ⁻⁷	10	HIV-1 Tat		
1	<i>Shigellosis</i>	7.190e ⁻⁷	10	<i>APC/C:Cdh1-mediated degradation of Skp2</i>	4.899e ⁻⁷	12
1	Oxidative phosphorylation	1.089e ⁻⁶	14	<i>RNA pol 3 abortive and retractive initiation</i>	5.689e ⁻⁷	10
1	<i>Epithelial cell signaling in Helicobacter pylori infection</i>	2.430e ⁻⁶	10	<i>Transport of the SLBP dependent mature mRNA</i>	7.579e ⁻⁷	9
1	<i>Viral myocarditis</i>	3.660e ⁻⁶	10	Cdc20:phospho-APC/C mediated degradation of Cyclin A	7.859e ⁻⁷	12
1	<i>Gap junction</i>	3.750e ⁻⁶	11	<i>Degradation multiubiquitinated Cyclin A</i>	7.859e ⁻⁷	12
1				<i>CD28 dependent Vav1 pathway</i>	8.129e ⁻⁷	6
1				<i>3'-polyadenylated, capped mRNA complex</i>	9.159e ⁻⁷	8
1				ATAC A complex	9.350e ⁻⁷	9
1				<i>TAP:3'-polyadenylated, capped mRNA complex</i>	1.289e ⁻⁶	8
1				<i>Transport of the export-competent complex through the NPC</i>	2.129e ⁻⁶	9
1				<i>Signaling by SCF-KIT</i>	2.129e ⁻⁶	9
1				mRNA capping	2.600e ⁻⁶	8
1				<i>RNA pol 2 transcription elongation</i>	3.289e ⁻⁶	9
1				Recruitment of elongation factors to form HIV-1 elongation	3.289e ⁻⁶	9
1				complex		
1				Addition of nucleotides leads to transcript elongation	3.289e ⁻⁶	9
1				Elongation complex	3.289e ⁻⁶	9

Appendix

1				HIV-1 elongation complex	3.289e ⁻⁶	9
2	Metabolic pathways	4.34E-30	83	Respiratory electron transport	2.290e ⁻²⁹	30
2	Alzheimer's disease	7.53E-40	44	Exon Junction Complex	7.959e ⁻²⁰	26
2	Huntington's disease	1.19E-37	44	Cleavage at the 3'-Splice site and Exon Ligation	7.959e ⁻²⁰	26
2	Oxidative phosphorylation	9.59E-35	37	mRNA Splicing- Major Pathway	1.039e ⁻¹⁹	26
2	Parkinson's disease	3.06E-32	35	Spliceosomal Intermediate C Complex	3.549e ⁻¹⁹	25
2	Spliceosome	5.33E-14	20	Spliceosomal Active C Complex	3.549e ⁻¹⁹	25
2	Proteasome	1.53E-22	19	Spliceosomal Active C Complex with lariat containing, 5'-end	3.549e ⁻¹⁹	25
2	Pathways in cancer	4.11E-06	19	Cleaved pre-Mrnp: CBC complex		
2	Purine metabolism	3.49E-09	17	Lariat Formation and 5'-Splice Site Cleavage	3.549e ⁻¹⁹	25
2	B cell receptor signaling pathway	9.96E-14	16	Spliceosomal A Complex	3.730e ⁻¹⁹	23
2	RNA transport	6.46E-09	16	Formation of an Intermediate Spliceosomal C Complex	3.730e ⁻¹⁹	25
2	T cell receptor signaling pathway	4.47E-09	14	Degradation of ubiquitinated p27/p21 by the 26s proteasome	2.370e ⁻¹⁸	19
2	Osteoclast differentiation	2.86E-08	14	SCF(Skp2)-mediated degradation of p27/p21	2.370e ⁻¹⁸	19
2	Ubiquitin mediated proteolysis	9.56E-08	14	APC/C:Cdh1 mediated degradation of Cdc20 and other	3.240e ⁻¹⁸	21
2	Cell cycle	1.96E-07	13	APC/C:Cdh1 targeted proteins in late mitosis/early G1		
2	Pyrimidine metabolism	1.50E-07	12	Spliceosomal B Complex	3.829e ⁻¹⁸	23
2	Focal adhesion	1.89E-04	12	Formation of the Spliceosomal B Complex	3.829e ⁻¹⁸	23
2	Long-term potentiation	2.24E-08	11	Complex 1 - NADH: Ubiquinone oxidoreductase	5.799e ⁻¹⁸	18
2	Apoptosis	2.80E-07	11	NADH enters the respiratory chain at Complex 1	5.799e ⁻¹⁸	18
2	Oocyte meiosis	2.89E-06	11	APC/C: Cdh 1 - mediated degradation of Skp2	8.690e ⁻¹⁸	20
2	Cardiac muscle contraction	6.50E-07	10	Degradation of multiubiquitinated Securin	8.690e ⁻¹⁸	20

Appendix

2	<i>Valine, leucine and isoleucine degradation</i>	8.51E-07	8	<i>APC/C:Cdc20 mediated degradation of Securin</i>	8.690e ⁻¹⁸	20
2	<i>Protein export</i>	6.58E-08	7	Formation of the Spliceosomal E complex	8.690e ⁻¹⁸	20
2	<i>RNA polymerase</i>	5.41E-07	7	Spliceosomal E Complex	8.690e ⁻¹⁸	20
2				<i>26s proteasome degrades ODC holoenzyme complex</i>	1.139e ⁻¹⁷	18
2				<i>26s proteasome</i>	1.290e ⁻¹⁷	17
2				<i>Cdc20: Phospho-APC/C mediated degradation of Cyclin A</i>	2.159e ⁻¹⁷	20
2				<i>Degradation of multiubiquitinated Cyclin A</i>	2.159e ⁻¹⁷	20
2				<i>Degradation of multiubiquitinated cell cycle proteins</i>	2.930e ⁻¹⁷	20
2				<i>Autodegradation of the E3 ubiquitin ligase COP1</i>	3.519e ⁻¹⁷	18
2				<i>Ubiquitin mediated degradation of phosphorylated Cdc25A</i>	3.519e ⁻¹⁷	18
2				Degradation of multiubiquitinated Cdh1	3.980e ⁻¹⁷	19
2				Autodegradation of Cdh1 by Cdh1:APC/C	3.980e ⁻¹⁷	19
2				<i>Destruction of AUF1 and Mrna</i>	4.340e ⁻¹⁷	17
2				<i>Proteasomal cleavage of exogenous antigen</i>	4.340e ⁻¹⁷	17
2				<i>Antigen processing : Ubiquitination and Proteasome degradation</i>	6.460e ⁻¹⁷	18
2				<i>SCF-mediated degradation of Emi 1</i>	6.460e ⁻¹⁷	18
2				<i>SCF-beta-</i>	6.460e ⁻¹⁷	18
2				<i>Cross-presentation of soluble exogenous antigens(endosomes)</i>	1.600e ⁻¹⁶	17
2				<i>Activation of NF-kappaB in B cells</i>	2.029e ⁻¹⁶	18
2				<i>Proteasomal cleavage of substrate</i>	2.320e ⁻¹⁶	17
2				<i>Regulation of activated PAK-2p34 by proteasome mediated</i>	2.989e ⁻¹⁶	17
2				<i>Proteasome mediated degradation of COP1</i>	2.989e ⁻¹⁶	17
2				<i>Proteolytic degradation of ubiquitinated-Cdc25A</i>	2.989e ⁻¹⁶	17
2				<i>Ubiquitinated Cdc6 is degraded by the proteasome</i>	2.989e ⁻¹⁶	17

Appendix

2				<i>Ubiquitinated Orc1 is degraded by the proteasome</i>	2.989e ⁻¹⁶	17
2				<i>Proteasome mediated degradation of PAK-2p34</i>	2.989e ⁻¹⁶	17
2				<i>Proteasome mediated degradation of Cyclin D1</i>	2.989e ⁻¹⁶	17
2				<i>Ubiquitinated geminin is degraded by the proteasome</i>	2.989e ⁻¹⁶	17
2				<i>Ubiquitin - dependent degradation of Cyclin D1</i>	4.380e ⁻¹⁶	17
2				<i>CDK-mediated phosphorylation and removal of Cdc6</i>	4.380e ⁻¹⁶	17
2				<i>Destabilization of mRNA by AUF1(hnRNP DO)</i>	9.659e ⁻¹⁶	17
2				<i>Degradation of beta-catenin by the destruction complex</i>	2.009e ⁻¹⁵	17
2				<i>Degradation of ubiquitinated-beta catenin by the proteasome</i>	2.009e ⁻¹⁵	17
2				<i>CDT1 association with the CDC6: ORC: origin complex</i>	5.850e ⁻¹⁵	17
2				Orc1 removal from chromatin	6.089e ⁻¹⁵	18
2				<i>ER- phagosome pathway</i>	8.099e ⁻¹⁵	18
2				<i>HP subcomplex</i>	3.139e ⁻¹⁴	14
2				Intron-containing complex	1.320e ⁻¹⁰	12
2				ATAC A Complex	3.759e ⁻¹⁰	11
2				mRNA splicing - Minor pathway	5.960e ⁻⁹	11
2				ATAC B Complex	5.960e ⁻⁹	11
2				ATAC C Complex with lariat containing 5'-end cleaved mRNA	5.960e ⁻⁹	11
2				<i>ATAC C Complex</i>	5.960e ⁻⁹	11
2				<i>Formation of AT-AC C Complex</i>	5.960e ⁻⁹	11
2				ATAC spliceosome mediated 3' splice site cleavage, exon ligation	5.960e ⁻⁹	11
2				<i>Capped, methylated pre-mRNP: CBC complex</i>	1.460e ⁻⁸	10
2				RNA polymerase 2 pre-transcription events	1.619e ⁻⁸	12
2				<i>Hypophosphorylation of RNA pol 2 CTD by FCP1P protein</i>	1.790e ⁻⁸	9
2				snRNP assembly	2.279e ⁻⁸	11

Appendix

2				<i>U2 snRNP</i>	3.669e ⁻⁸	8
2				<i>snRNP nuclear import and release</i>	6.490e ⁻⁸	10
2				<i>Formation of the Early Elongation Complex</i>	8.239e ⁻⁸	9
2				<i>Formation of the HIV-1 Early Elongation Complex</i>	8.239e ⁻⁸	9
2				Formation of RNA pol 2 elongation complex	8.239e ⁻⁸	10
2				Formation of HIV-1 elongation complex in the absence of HIV-1 Tat	8.239e ⁻⁸	10
2				Microtubule-bound kinetochore	6.330e ⁻⁷	12
2				Kinetochore	6.330e ⁻⁷	12
2				mRNA capping	6.959e ⁻⁷	8
2				<i>RNA polymerase 2 transcription elongation</i>	6.959e ⁻⁷	9
2				<i>Recruitment of elongation factors to form HIV-1 elongation complex</i>	6.959e ⁻⁷	9
2				<i>Addition of nucleotides 5 through 9 on the growing transcript</i>	6.959e ⁻⁷	9
2				Addition of nucleotides leads to transcript elongation	6.959e ⁻⁷	9
2				Elongation complex	6.959e ⁻⁷	9
2				<i>HIV-1 transcription complex containing 4 nucleotide long transcript</i>	6.959e ⁻⁷	9
2				<i>Addition of nucleotides 5 through 9 on the growing HIV-1 transcript</i>	6.959e ⁻⁷	9
2				<i>pol 2 transcription complex containing 4 nucleotide long transcript</i>	6.959e ⁻⁷	9
2				HIV-1 elongation complex	6.959e ⁻⁷	9
2				<i>snRNP Sm core complex</i>	6.959e ⁻⁷	5
2				<i>U11 snRNP</i>	6.959e ⁻⁷	5
2				<i>U4 ATAC snRNP</i>	6.959e ⁻⁷	5
2				<i>U4 ATAC snRNP: U6 ATAC snRNP</i>	6.959e ⁻⁷	5
2				<i>Spliceosomal m3G capped snRNA loaded with the SM complex</i>	6.959e ⁻⁷	5
2				<i>U6 ATAC snRNP</i>	6.959e ⁻⁷	5
2				<i>RNA polymerase 2 transcription initiation</i>	6.959e ⁻⁷	9

Appendix

2				<i>RNA polymerase 2 transcription pre-initiation and promoter opening</i>	6.959e ⁻⁷	9
2				<i>RNA polymerase 2 promoter opening: first transition</i>	6.959e ⁻⁷	9
2				<i>newly formed phosphodiester bond stabilized and Ppi released</i>	6.959e ⁻⁷	9
2				<i>RNA polymerase 2 promoter escape</i>	6.959e ⁻⁷	9
2				<i>pol 2 promoter escape complex</i>	6.959e ⁻⁷	9
2				<i>HIV-1 transcription initiation</i>	6.959e ⁻⁷	9
2				<i>Transcription of the HIV genome</i>	6.959e ⁻⁷	9
2				<i>Addition of the fourth nucleotide on the nascent transcript: second transition</i>	6.959e ⁻⁷	9
2				<i>HIV-1 transcription complex containing 3 nucleotide long transcript</i>	6.959e ⁻⁷	9
2				HIV-1 initiation complex with phosphodiester-Ppi intermediate	6.959e ⁻⁷	9
2				HIV-1 open pre-initiation complex	6.959e ⁻⁷	9
2				HIV-1 promoter escape complex	6.959e ⁻⁷	9
2				RNA polymerase 2 HIV-1 promoter escape	6.959e ⁻⁷	9
2				NTP binds active site of RNA polymerase 2	6.959e ⁻⁷	9
2				Nucleophilic attack by 3'- hydroxyl Oxygen of nascent transcript on the alpha phosphate of NTP	6.959e ⁻⁷	9
2				Addition of the third nucleotide on the nascent transcript	6.959e ⁻⁷	9
2				HIV-1 transcription complex	6.959e ⁻⁷	9
2				pol 2 transcription complex containing 3 nucleotide long transcript	6.959e ⁻⁷	9
2				pol 2 initiation complex with phosphodiester-Ppi intermediate	6.959e ⁻⁷	9
2				pol 2 colsed pre-initiation complex	6.959e ⁻⁷	9
2				pol 2 initiation complex	6.959e ⁻⁷	9
2				pol 2 open pre-initiation complex	6.959e ⁻⁷	9
2				HIV-1 promoter opening: first transition	6.959e ⁻⁷	9
2				Addition of the third nucleotide on the nascent HIV-1 transcript	6.959e ⁻⁷	9

Appendix

2				Addition of the fourth nucleotide on the nascent HIV-1 transcript: second transition	6.959e ⁻⁷	9
2				<i>Fall back to closed pre-initiation complex</i>	6.959e ⁻⁷	9
2				<i>Nucleophilic attack by 3'-hydroxyl Oxygen of nascent HIV-1 transcript on the Alpha phosphate of NTP</i>	6.959e ⁻⁷	9
2				<i>newly formed phosphodiester bond stabilized and Ppi released pre-initiation complex</i>	6.959e ⁻⁷	9
2				<i>NTP binds active site of RNA polymerase 2 in HIV-1 open</i>	6.959e ⁻⁷	9
2				<i>HIV-1 closed pre-initiation complex</i>	6.959e ⁻⁷	9
2				<i>pol 2 transcription complex</i>	6.959e ⁻⁷	9
2				<i>HIV-1 initiation complex</i>	6.959e ⁻⁷	9
2				<i>U12 snRNP</i>	1.419e ⁻⁶	6
2				<i>Addition of nucleotides between position +11 and +30</i>	2.349e ⁻⁶	7
2				<i>pol 2 transcription complex with (ser5) phosphorylated CTD containing extruded transcript to +30</i>	2.349e ⁻⁶	7
2				<i>pol 2 transcription complex containing extruded transcript to +30</i>	2.349e ⁻⁶	7
2				<i>pol 2 transcription complex containing 11 nucleotide long transcript</i>	2.349e ⁻⁶	7
2				<i>pol 2 transcription complex containing 9 nucleotide long transcript</i>	2.349e ⁻⁶	7
2				<i>Phosphorylation (ser5) of RNA pol 2 CTD</i>	2.349e ⁻⁶	7
2				<i>Addition of nucleotides 10 and 11 on the growing transcript: third transition</i>	2.349e ⁻⁶	7
2				<i>Addition of nucleotides between position +11 and +30 on HIV-1 transcript</i>	2.349e ⁻⁶	7
2				<i>RNA polymerase 2: NTP: TF2F complex</i>	2.349e ⁻⁶	7
2				<i>pol 2 transcription complex containing transcript to +30</i>	2.349e ⁻⁶	7
2				<i>HIV-1 transcription complex containing 9 nucleotide long transcript</i>	2.349e ⁻⁶	7
2				<i>HIV-1 transcription complex containing 11 nucleotide long transcript</i>	2.349e ⁻⁶	7
2				<i>Addition of nucleotides 10 and 11 on the growing HIV-1 transcript: third transition</i>	2.349e ⁻⁶	7
2				<i>HIV-1 transcription complex containing extruded transcript to +30</i>	2.349e ⁻⁶	7

Appendix

2				<i>HIV-1 transcription complex containing transcript to +30</i>	2.349e ⁻⁶	7
2				<i>HIV-1 transcription complex with (ser5) phosphorylated CTD containing extruded transcript to +30</i>	2.349e ⁻⁶	7
2				<i>RNA polymerase 2 CTD(phosphorylated) binds to CE</i>	3.279e ⁻⁶	7
2				<i>RNA pol 2 with phosphorylated CTD: CE complex with activated GT</i>	3.279e ⁻⁶	7
2				<i>RNA pol 2 with phosphorylated CTD: CE complex</i>	3.279e ⁻⁷	7
2				<i>cNAP-1 depleted centrosome</i>	3.539e ⁻⁶	10
2				<i>Nlp-depleted centrosome</i>	3.539e ⁻⁶	10
2				<i>Plk-1-mediated phosphorylation of Nlp</i>	3.850e ⁻⁶	10
2				<i>Loss of C-Nap-1 from centrosome</i>	3.850e ⁻⁶	10
2				<i>Centrosome containing recruited CDK11p58</i>	3.850e ⁻⁶	10
2				<i>Centrosome</i>	3.850e ⁻⁶	10
2				<i>Centrosome associated Plk1</i>	3.850e ⁻⁶	10
2				<i>Loss of Nlp from mitotic centrosomes</i>	3.850e ⁻⁶	10
2				<i>Centrosome containing phosphorylated Nlp</i>	3.850e ⁻⁶	10
2				<i>U1 snRNP</i>	3.850e ⁻⁶	5
2				<i>u6 snRNP</i>	3.850e ⁻⁶	5
2				<i>U4 snRNP:U6 snRNP complex</i>	3.850e ⁻⁶	5
2				<i>Transfer of GMP from the capping enzyme GT site to 5'-end of mRNA</i>	3.850e ⁻⁶	7
2				<i>Methylation of GMP-cap by RNA methyltransferase</i>	3.850e ⁻⁶	7
2				<i>SPT5 subunit of pol 2 binds the RNA triphosphatase (RTP)</i>	3.850e ⁻⁶	7
2				<i>Capping complex (initial)</i>	3.850e ⁻⁶	7
2				<i>Hydrolysis of the 5'- end of the nascent transcript by the capping enzyme</i>	3.850e ⁻⁶	7
2				<i>RNA pol 2 CTD phosphorylation and interaction with CE</i>	3.850e ⁻⁶	7
2				<i>Formation of the CE:GMP intermediated complex</i>	3.850e ⁻⁶	7
2				<i>pol 2 transcription complex containing 4-9 nucleotide long transcript</i>	3.850e ⁻⁶	7

Appendix

2				<i>CE:Pol 2 CTD: Spt5 complex</i>	3.850e ⁻⁶	7
2				<i>HIV-1 transcription complex containing 4-9 nucleotide long transcript</i>	3.850e ⁻⁶	7
2				<i>Capping complex(intermediate)</i>	3.850e ⁻⁶	7
2				<i>Capping complex(hydrolyzed)</i>	3.850e ⁻⁶	7
2				<i>Covalent CE: GMP intermediate complex</i>	3.850e ⁻⁶	7
2				<i>Capping complex (with freed 5'-GMP)</i>	3.850e ⁻⁶	7
2				Capping complex(GpppN..) protein	3.850e ⁻⁶	7
3	No match			No match		
4	Cell cycle	1.369e ⁻²⁴	27	<i>Activation of the pre-replicative complex</i>	5.280e ⁻²⁰	16
4	DNA replication	1.139e ⁻²³	18	<i>Activation of ATR in response to replication stress</i>	8.130e ⁻¹⁹	16
4	Pyrimidine metabolism	8.950e ⁻¹⁵	18	<i>DNA polymerase alpha: primase binds at the origin</i>	8.629e ⁻¹⁹	15
4	Purine metabolism	4.200e ⁻⁹	16	Microtubule-bound kinetochore	2.170e ⁻¹⁴	17
4	Metabolic pathways	4.200e ⁻⁹	41	Kinetochore	2.170e ⁻¹⁴	17
4	Nucleotide excision repair	4.200e ⁻⁹	10	<i>DNA polymerase epsilon binds at the origin</i>	2.170e ⁻¹⁴	12
4	Mismatch repair	5.149e ⁻⁹	8	<i>RPA: Cdc45:CDK:DDK:Mcm10:pre-replicative complex</i>	4.289e ⁻¹³	11
4	Chemokine signaling pathway	1.559e ⁻⁷	15	<i>Cdc45:CDK:DDK:Mcm10:pre-replicative complex</i>	2.099e ⁻¹²	10
4	Cytokine-cytokine receptor interaction	4.309e ⁻⁷	17	<i>Activation of claspin</i>	3.189e ⁻¹²	10
4	Oocyte meiosis	1.480e ⁻⁶	11	<i>Cdc45:CDK:DDK:Mcm10:claspin:pre-replicative complex</i>	3.189e ⁻¹²	10
4	p53 signaling pathway	1.839e ⁻⁶	9	<i>Cdc45:CDK:DDK:Mcm10:activated claspin:pre-replicative complex</i>	3.189e ⁻¹²	10
4				<i>Activation of E2F target genes at G1/S</i>	1.310e⁻¹¹	9
4				Orc1 removal from chromatin	1.790e ⁻¹¹	14

Appendix

4				<i>Removal of licensing factors from origins</i>	2.569e ⁻¹¹	9
4				<i>CDK:DDK:Mcm10:pre-replicative complex</i>	5.069e ⁻¹¹	9
4				<i>Mcm10:pre-replicative complex</i>	9.900e ⁻¹⁰	8
4				<i>Cdt1 displaced from the pre-replicative complex</i>	9.900e ⁻¹⁰	8
4				<i>Unwinding of DNA</i>	1.499e ⁻⁹	7
4				<i>Unwinding complex at replication fork</i>	1.499e ⁻⁹	7
4				Cdc20:phospho-APC/C mediated degradation of Cyclin A	5.470e ⁻⁹	12
4				<i>Degradation multiubiquitinated Cyclin A</i>	5.470e ⁻⁹	12
4				<i>Orc1 is phosphorylated by Cyclin A/CDK2</i>	8.830e ⁻⁹	8
4				<i>Pre-replicative complex(Orc1-minus)</i>	1.750e ⁻⁸	7
4				<i>RSF Complex binds the centromere</i>	2.070e ⁻⁸	8
4				Deposition of new CENPA-containing nucleosomes at the centromere	2.070e ⁻⁸	8
4				<i>Mcm10:active pre-replicative complex</i>	2.619e ⁻⁸	7
4				<i>Polymerase switching</i>	2.619e ⁻⁸	7
4				<i>Removal of the Flap Intermediate</i>	2.619e ⁻⁸	7
4				<i>Polymerase switching on the C-strand of the telomere</i>	2.619e ⁻⁸	7
4				<i>Pre-replicative complex</i>	2.619e ⁻⁸	7
4				<i>RFC Heteropentamer:RNA primer-DNA primer:origin duplex</i>	5.589e ⁻⁸	6
4				<i>RFC dissociates after sliding clamp formation</i>	1.270e ⁻⁷	6
4				<i>Loading of PCNA-sliding clamp formation</i>	1.270e ⁻⁷	6
4				<i>RFC Heteropentamer : RNA primer-DNA primer: origin duple:PCNA homotrimer</i>	1.270e ⁻⁷	6
4				<i>CDT1 association with the CDC6:ORC:origin complex</i>	2.010e ⁻⁷	10
4				Deposition of New CENPA-containing nucleosomes at the centromere	2.949e ⁻⁷	7
4				<i>Centromeric chromatin: New CENPA nucleosome:Mis18: HJURP complex</i>	2.949e ⁻⁷	7
4				<i>Centromeric chromatin : CENPH-1:Mis18:HJURP:CENPA complex</i>	2.949e ⁻⁷	7

Appendix

4				<i>Regulation of APC/C activators between G1/S and early anaphase</i>	4.379e ⁻⁷	7
4				<i>G0 and early G1</i>	6.360e ⁻⁷	7
4				<i>Removal of RNA primer and dissociation of RPA and Dna2 active pre-replicative complex</i>	8.099e ⁻⁷	6
4				<i>Processive complex: Okazaki fragment:Flap:RPA heterotrimer:dna2</i>	8.099e ⁻⁷	6
4				Degradation of multiubiquitinated cell cycle proteins	1.036e ⁻⁶	10
4				<i>APC/C:Cdh1 mediated degradation of Cdc20 and other</i>	1.190e ⁻⁶	10
4				APC/C:Cdh1 targeted proteins in late mitosis/early G1	2.210e ⁻⁶	7
4				<i>Mcm2-7 is phosphorylated by DDK</i>	2.210e ⁻⁶	5
5	Cytokine-cytokine receptor interaction	8.689e ⁻¹⁰	11	No match		
EXPERIMENT: E-GEOD-38836, Upregulated Canonical Pathways						
Maximum stable clusters : 7						
Functional enrichment p -value <10⁻⁵						
cluster	KEGG pathway	p-value	# of proteins	Reactome pathway	p-value	# of proteins
1	Neuroactive ligand-receptor interaction	8.47E-10	10	No Match		
1	Calcium signaling pathway	3.86E-06	6			
2	Neuroactive ligand-receptor interaction	4.88E-22	26	G alpha (i) signalling events	2.23E-15	17
2	Focal adhesion	6.66E-14	17	The Ligand:GPCR:Gi complex dissociates	2.67E-15	16
2	ECM-receptor interaction	5.52E-10	10	<i>The Ligand:GPCR:Gs complex dissociates</i>	1.67E-08	8
2	Retinol metabolism	4.15E-06	6	<i>G alpha (s) signalling events</i>	1.95E-08	9

Appendix

2				The Ligand:GPCR:Gq complex dissociates	2.72E-06	8
2				G alpha (q) signalling events	6.16E-06	8
3	Neuroactive ligand-receptor interaction	2.57E-26	30	No Match		
3	Glutamatergic synapse	2.39E-10	12			
3	<i>Glycine, serine and threonine metabolism</i>	4.00E-09	7			
3	Cholinergic synapse	1.45E-08	10			
3	Calcium signaling pathway	9.66E-08	11			
3	<i>Pathways in cancer</i>	1.20E-06	13			
3	Focal adhesion	2.94E-06	10			
3	ECM-receptor interaction	4.17E-06	7			
3	Pancreatic secretion	8.74E-06	7			
4	Glutamatergic synapse	1.89E-09	8			
4	Vascular smooth muscle contraction	2.44E-08	7			
4	Gastric acid secretion	4.23E-08	6			
4	<i>GnRH signaling pathway</i>	2.76E-07	6			
4	Cholinergic synapse	6.14E-07	6			
4	<i>Salivary secretion</i>	2.75E-06	5			
4	Pancreatic secretion	6.42E-06	5			

Appendix

5	Vascular smooth muscle contraction	1.31E-10	8	No match		
6	Neuroactive ligand-receptor interaction	6.22E-22	24	G alpha (q) signalling events	3.77E-10	11
6	Calcium signaling pathway	4.95E-09	11	<i>Activation of voltage gated Potassium channels</i>	4.14E-09	7
6	<i>Steroid hormone biosynthesis</i>	2.56E-08	7	<i>Octamer of Voltage gated K⁺ channels</i>	4.14E-09	7
6	Gastric acid secretion	1.77E-07	7	The Ligand:GPCR:Gq complex dissociates	3.52E-08	9
6	Retinol metabolism	1.15E-06	6	G alpha (i) signalling events	3.89E-07	9
6	<i>Drug metabolism - cytochrome P450</i>	3.19E-06	6	The Ligand:GPCR:Gi complex dissociates	1.19E-06	8
6	Cholinergic synapse	3.74E-06	7	<i>OR - G Protein Trimer Complex</i>	<i>2.82E-05</i>	<i>10</i>
6	Vascular smooth muscle contraction	3.97E-06	7	<i>Olfactory Receptor - G Protein olfactory trimer complex formation</i>	<i>2.82E-05</i>	<i>10</i>
6	Glutamatergic synapse	8.22E-06	7			
6	<i>Olfactory transduction</i>	<i>1.14E-05</i>	<i>11</i>			
7	No match			No match		
EXPERIMENT: E-GEOD-38836, Downregulated Canonical Pathways						
Maximum stable clusters : 5						
Functional enrichment p-value <10⁻⁵						
cluster	KEGG pathway	p-value	# of proteins	Reactome pathway	p-value	# of proteins
1	Jak-STAT signaling pathway	1.75E-11	9	None		
1	Cytokine-cytokine receptor interaction	7.91E-11	10			

Appendix

1	Natural killer cell mediated cytotoxicity	2.03E-07	6			
1	Systemic lupus erythematosus	1.42E-06	5			
1	Measles	6.02E-06	5			
2	Ribosome	2.72E-19	17	GTP hydrolysis and joining of the 60S ribosomal subunit	3.02E-32	26
2	Spliceosome	9.90E-18	18	L13a-mediated translational silencing of Ceruloplasmin expression	1.11E-30	25
2	T cell receptor signaling pathway	7.32E-15	15	Formation of a pool of free 40S subunits	1.50E-28	23
2	Natural killer cell mediated cytotoxicity	2.57E-12	14	SRP-dependent cotranslational protein targeting to membrane	2.96E-27	23
2	RNA transport	2.17E-10	13	Translated mRNA Complex with Premature Termination Codon Preceding Exon Junction	1.61E-26	22
2	Measles	1.20E-06	9	80S:Met-tRNAi:mRNA:eIF5B:GTP	1.95E-26	21
2	Renal cell carcinoma	1.33E-06	7	eIF5B:GTP is hydrolyzed and released	1.95E-26	21
2	Cytokine-cytokine receptor interaction	1.55E-06	12	The 60S subunit joins the translation initiation complex	1.95E-26	21
2	B cell receptor signaling pathway	1.97E-06	7	Signal peptide cleavage from ribosome-associated nascent protein	2.09E-26	22
2	Chemokine signaling pathway	2.51E-06	10	eRF3-GDP:eRF1:80S Ribosome:mRNA:peptidyl-tRNA Complex	2.63E-26	21
2	Neurotrophin signaling pathway	7.13E-06	8	eRF3-GTP:eRF1:80S Ribosome:mRNA:peptidyl-tRNA Complex	2.63E-26	21
2	Regulation of actin cytoskeleton	7.23E-06	10	GTP Hydrolysis by eRF3 bound to the eRF1:mRNA:polypeptide:80S Ribosome complex	2.63E-26	21
2				Translocation of ribosome by 3 bases in the 3' direction	2.63E-26	21
2				eRF3-GDP:eRF1:80S Ribosome:mRNA:tRNA Complex	2.63E-26	21
2				Peptide chain elongation	2.63E-26	21
2				Eukaryotic Translation Termination	2.63E-26	21

Appendix

2				Polypeptide release from the eRF3-GDP:eRF1:mRNA:80S Ribosome complex	2.63E-26	21
2				Translocation of signal-containing nascent peptide to Endoplasmic Reticulum	3.50E-26	22
2				SMG1:Phosphorylated UPF1:EJC:Translated mRNP	4.52E-26	22
2				SMG1 Phosphorylates UPF1 (Enhanced by Exon Junction Complex)	4.52E-26	22
2				SMG1:UPF1:EJC:Translated mRNP	4.52E-26	22
2				SRP:polypeptide+signal:ribosome	6.26E-26	21
2				Translated mRNA Complex with Premature Termination Codon Not Preceding Exon Junction	8.28E-26	21
2				Phosphorylated UPF1:SMG5:SMG7:SMG6:PP2A:Translated mRNP	9.48E-26	22
2				SMG6 Cleaves mRNA with Premature Termination Codon	9.48E-26	22
2				UPF1:eRF3 Complex on Translated mRNA	1.09E-25	21
2				SRP receptor:SRP:ribosome:polypeptide+signal	1.09E-25	21
2				Nonsense Mediated Decay Enhanced by the Exon Junction Complex	1.94E-25	22
2				Phosphorylated UPF1 Recruits SMG5, SMG7, SMG6, and PP2A	1.94E-25	22
2				cleaved polypeptide:Translocon	2.45E-25	21
2				polypeptide+signal:Translocon	2.45E-25	21
2				80S Ribosome:mRNA:peptidyl-tRNA with elongating peptide	7.15E-25	20
2				Peptide transfer from P-site tRNA to the A-site tRNA	7.15E-25	20
2				Elongation complex with growing peptide chain	7.15E-25	20
2				80S ribosome	7.15E-25	20
2				membrane-bound ribosome:mRNA:cleaved polypeptide	7.15E-25	20
2				80S:Met-tRNAi:mRNA:aminoacyl-tRNA	7.15E-25	20
2				ribosome:mRNA:polypeptide+signal	7.15E-25	20
2				80S:Met-tRNAi:mRNA	7.15E-25	20
2				membrane-bound ribosome:mRNA:polypeptide+signal	7.15E-25	20

Appendix

2				Hydrolysis of eEF1A:GTP	9.51E-25	20
2				80S:aminoacyl tRNA:mRNA:eEF1A:GTP	9.51E-25	20
2				<i>60s ribosomal complex lacking L13a subunit</i>	1.49E-20	15
2				<i>60S ribosomal complex</i>	2.14E-20	15
2				mRNA Splicing - Major Pathway	8.63E-18	17
2				Spliceosomal B Complex	1.61E-17	16
2				Formation of the Spliceosomal B Complex	1.61E-17	16
2				Formation of an intermediate Spliceosomal C complex	1.38E-16	16
2				Exon Junction Complex	1.63E-16	16
2				Cleavage at the 3'-Splice Site and Exon Ligation	1.63E-16	16
2				Spliceosomal Intermediate C Complex	2.52E-15	15
2				Spliceosomal Active C Complex	2.52E-15	15
2				Spliceosomal A Complex	2.60E-15	14
2				Spliceosomal active C complex with lariat containing, 5'-end cleaved pre-mRNP:CBC complex	2.95E-15	15
2				Lariat Formation and 5'-Splice Site Cleavage	2.95E-15	15
2				<i>48S complex</i>	7.59E-12	10
2				<i>43S:mRNA:eIF4F:eIF4B:eIF4H</i>	7.59E-12	10
2				<i>Ribosomal scanning</i>	7.59E-12	10
2				<i>Ribosomal scanning and start codon recognition</i>	9.38E-12	10
2				<i>eIF2:GTP is hydrolyzed, eIFs are released</i>	9.38E-12	10
2				<i>Formation of translation initiation complexes yielding circularized Ceruloplasmin mRNA in a 'closed-loop' conformation</i>	9.38E-12	10
2				<i>43S: Ceruloplasmin mRNA:eIF4F:eIF4B:eIF4H:PABP</i>	9.38E-12	10
2				<i>phospho-L13a associated with the 3' UTR GAIT element of ceruloplasmin mRNA within the translation initiation complex</i>	1.15E-11	10
2				Formation of the Spliceosomal E complex	7.31E-11	10

Appendix

2				Spliceosomal E Complex	7.31E-11	10
2				mRNA Splicing - Minor Pathway	8.08E-10	8
2				ATAC B Complex	8.08E-10	8
2				ATAC C Complex with lariat containing 5'-end cleaved mRNA	8.08E-10	8
2				ATAC C Complex	8.08E-10	8
2				Formation of AT-AC C complex	8.08E-10	8
2				ATAC spliceosome mediated 3' splice site cleavage, exon ligation	8.08E-10	8
2				<i>40S:eIF3:eIF1A</i>	1.24E-09	8
2				<i>Formation of the ternary complex, and subsequently, the 43S complex</i>	2.25E-09	8
2				<i>Formation of the 43S pre-initiation complex</i>	2.25E-09	8
2				<i>43S complex</i>	2.25E-09	8
2				intron-containing complex	1.75E-08	7
2				ATAC A Complex	1.65E-07	6
2				<i>U2 snRNP</i>	2.64E-07	5
2				RNA Polymerase II Pre-transcription Events	2.67E-07	7
2				Docking of the TAP:EJC Complex with the NPC	2.16E-06	6
2				HIV-1 promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF complex*	2.34E-06	5
2				pol II promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF complex	2.34E-06	5
2				Transport of Mature mRNA derived from an Intron-Containing Transcript	2.45E-06	6
2				pol II promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF:TFIIE complex	3.40E-06	5
2				<i>40S ribosomal complex</i>	4.05E-06	5
2				<i>40S:Met-tRNAi:mRNA</i>	4.05E-06	5
2				<i>capped, methylated pre-mRNP:CBC complex</i>	5.65E-06	5
3	Spliceosome	8.96E-18	18	Cdc20:Phospho-APC/C mediated degradation of Cyclin A	3.32E-18	15

Appendix

3	MAPK signaling pathway	4.93E-13	19	Degradation multiubiquitinated Cyclin A	3.32E-18	15
3	Cell cycle	9.70E-13	14	Degradation of beta-catenin by the destruction complex	4.13E-18	14
3	Proteasome	1.41E-12	10	Degradation of ubiquitinated -beta catenin by the proteasome	4.13E-18	14
3	Huntington's disease	2.44E-10	14	Degradation of multiubiquitinated Cdh1	2.27E-17	14
3	<i>Hepatitis C</i>	6.19E-10	12	Autodegradation of Cdh1 by Cdh1:APC/C	2.27E-17	14
3	<i>TGF-beta signaling pathway</i>	1.20E-09	10	APC/C:Cdh1-mediated degradation of Skp2	6.33E-17	14
3	Natural killer cell mediated cytotoxicity	9.23E-08	10	Degradation of multiubiquitinated Securin	6.33E-17	14
3	<i>RNA polymerase</i>	1.04E-07	6	APC/C:Cdc20 mediated degradation of Securin	6.33E-17	14
3	Measles	1.07E-07	10	Degradation of ubiquitinated p27/p21 by the 26S proteasome	8.27E-17	13
3	<i>Ubiquitin mediated proteolysis</i>	1.33E-07	10	SCF(Skp2)-mediated degradation of p27/p21	8.27E-17	13
3	<i>Oocyte meiosis</i>	2.19E-07	9	SCF-mediated degradation of Emi1	1.48E-16	13
3	Cytosolic DNA-sensing pathway	4.80E-07	7	SCF-beta-TrCP mediated degradation of Emi1	1.48E-16	13
3	Parkinson's disease	8.81E-07	9	Degradation of multiubiquitinated cell cycle proteins	1.64E-16	14
3	Chagas disease (American trypanosomiasis)	1.54E-06	8	APC/C:Cdh1 mediated degradation of Cdc20 and other APC/C:Cdh1 targeted proteins in late mitosis/early G1	2.06E-16	14
3	Wnt signaling pathway	3.18E-06	9	Regulation of activated PAK-2p34 by proteasome mediated degradation	2.20E-15	12
3	<i>mRNA surveillance pathway</i>	4.21E-06	7	Proteasome mediated degradation of COP1	2.20E-15	12
3	Osteoclast differentiation	6.44E-06	8	Proteolytic degradation of ubiquitinated-Cdc25A	2.20E-15	12
3	Alzheimer's disease	7.05E-06	9	Proteasome mediated degradation of PAK-2p34	2.20E-15	12
3	Oxidative phosphorylation	9.76E-06	8	Autodegradation of the E3 ubiquitin ligase COP1	3.82E-15	12
3				Ubiquitin Mediated Degradation of Phosphorylated Cdc25A	3.82E-15	12
3				Destabilization of mRNA by AUF1 (hnRNP D0)	4.99E-15	12

Appendix

3				Antigen processing: Ubiquitination & Proteasome degradation	6.49E-15	12
3				26S proteasome	9.74E-15	11
3				CDT1 association with the CDC6:ORC:origin complex	1.74E-14	12
3				Destruction of AUF1 and mRNA	2.44E-14	11
3				Proteasomal cleavage of exogenous antigen	2.44E-14	11
3				Cross-presentation of soluble exogenous antigens (endosomes)	5.66E-14	11
3				Proteasomal cleavage of substrate	7.38E-14	11
3				26S proteasome degrades ODC holoenzyme complex	7.38E-14	11
3				Ubiquitinated Cdc6 is degraded by the proteasome	9.57E-14	11
3				Ubiquitinated Orc1 is degraded by the proteasome	9.57E-14	11
3				Proteasome mediated degradation of Cyclin D1	9.57E-14	11
3				Ubiquitinated geminin is degraded by the proteasome	9.57E-14	11
3				Ubiquitin-dependent degradation of Cyclin D1	1.23E-13	11
3				CDK-mediated phosphorylation and removal of Cdc6	1.23E-13	11
3				Orc1 removal from chromatin	1.49E-13	12
3				Activation of NF-kappaB in B Cells	5.01E-13	11
3				Spliceosomal Intermediate C Complex	1.20E-12	13
3				Spliceosomal Active C Complex	1.20E-12	13
3				Spliceosomal active C complex with lariat containing, 5'-end cleaved pre-mRNP:CBC complex	1.37E-12	13
3				Lariat Formation and 5'-Splice Site Cleavage	1.37E-12	13
3				Formation of an intermediate Spliceosomal C complex	1.57E-12	13
3				Spliceosomal A Complex	1.72E-12	12
3				Exon Junction Complex	1.79E-12	13
3				Cleavage at the 3'-Splice Site and Exon Ligation	1.79E-12	13
3				mRNA Splicing - Major Pathway	2.31E-12	13

Appendix

3				ER-Phagosome pathway	5.17E-12	11
3				Spliceosomal B Complex	6.48E-12	12
3				Formation of the Spliceosomal B Complex	6.48E-12	12
3				Formation of the Spliceosomal E complex	6.92E-11	10
3				Spliceosomal E Complex	6.92E-11	10
3				TRAF6 mediated NF-kB activation	7.29E-10	6
3				Interleukin-1 signaling	1.11E-08	7
3				ATAC A Complex	1.60E-07	6
3				<i>mRNA 3'-end processing</i>	2.86E-07	6
3				<i>Cleavage of mRNA at the 3'-end</i>	2.86E-07	6
3				<i>Ligated exon containing complex</i>	2.86E-07	6
3				mRNA Splicing - Minor Pathway	6.70E-07	6
3				ATAC B Complex	6.70E-07	6
3				ATAC C Complex with lariat containing 5'-end cleaved mRNA	6.70E-07	6
3				ATAC C Complex	6.70E-07	6
3				Formation of AT-AC C complex	6.70E-07	6
3				ATAC spliceosome mediated 3' splice site cleavage, exon ligation	6.70E-07	6
3				<i>3'-polyadenylated, capped mRNA complex</i>	7.68E-07	5
3				<i>TAP:3'-polyadenylated, capped mRNA complex</i>	9.75E-07	5
3				Docking of the TAP:EJC Complex with the NPC	9.75E-07	5
3				<i>Prefoldin mediated transfer of substrate to CCT/TriC</i>	1.22E-06	5
3				<i>Actin/tubulin:prefoldin complex associates with CCT/TriC</i>	1.22E-06	5
3				Phosphorylated UPF1:SMG5:SMG7:SMG6:PP2A:Translated mRNP	1.43E-06	8
3				SMG6 Cleaves mRNA with Premature Termination Codon	1.43E-06	8
3				Nonsense Mediated Decay Enhanced by the Exon Junction Complex	1.79E-06	8

Appendix

3				Phosphorylated UPF1 Recruits SMG5, SMG7, SMG6, and PP2A	1.79E-06	8
3				<i>mRNA polyadenylation</i>	2.28E-06	5
3				<i>3' end cleaved, ligated exon containing complex</i>	2.28E-06	5
3				Transport of Mature mRNA derived from an Intron-Containing Transcript	2.37E-06	6
3				Respiratory electron transport	2.50E-06	7
3				Formation of the Early Elongation Complex	2.76E-06	5
3				Formation of the HIV-1 Early Elongation Complex	2.76E-06	5
3				Cyclin D associated events in G1	2.76E-06	5
3				RNA Polymerase II Pre-transcription Events	4.80E-06	6
3				<i>HP subcomplex</i>	7.50E-06	5
4	Natural killer cell mediated cytotoxicity	1.67E-15	16	SRP-dependent cotranslational protein targeting to membrane	4.81E-13	13
4	Huntington's disease	5.38E-12	15	Signal peptide cleavage from ribosome-associated nascent protein	4.28E-12	12
4	Oxidative phosphorylation	1.95E-10	12	ER-Phagosome pathway	5.32E-11	10
4	Proteasome	2.68E-08	7	Translocation of signal-containing nascent peptide to Endoplasmic Reticulum	1.14E-10	11
4	Parkinson's disease	3.33E-08	10	RNA Polymerase II Pre-transcription Events	2.07E-10	9
4	Alzheimer's disease	3.42E-08	11	cleaved polypeptide:Translocon	9.91E-10	10
4	Protein export	4.87E-07	5	polypeptide+signal:Translocon	9.91E-10	10
4	Cytokine-cytokine receptor interaction	5.49E-07	12	Orc1 removal from chromatin	1.16E-09	9
4	Chronic myeloid leukemia	7.65E-07	7	APC/C:Cdh1 mediated degradation of Cdc20 and other APC/C:Cdh1 targeted proteins in late mitosis/early G1	1.53E-09	9
4	Jak-STAT signaling pathway	2.08E-06	9	Degradation of ubiquitinated p27/p21 by the 26S proteasome	2.56E-09	8
4	Spliceosome	3.92E-06	8	SCF(Skp2)-mediated degradation of p27/p21	2.56E-09	8

Appendix

4	Cytosolic DNA-sensing pathway	4.78E-06	6	Spliceosomal Intermediate C Complex	2.89E-09	10
4	Measles	5.57E-06	8	Spliceosomal Active C Complex	2.89E-09	10
4				Spliceosomal active C complex with lariat containing, 5'-end cleaved pre-mRNP:CBC complex	3.19E-09	10
4				Lariat Formation and 5'-Splice Site Cleavage	3.19E-09	10
4				Formation of an intermediate Spliceosomal C complex	3.52E-09	10
4				Exon Junction Complex	3.89E-09	10
4				Cleavage at the 3'-Splice Site and Exon Ligation	3.89E-09	10
4				Respiratory electron transport	4.27E-09	9
4				mRNA Splicing - Major Pathway	4.71E-09	10
4				Activation of NF-kappaB in B Cells	5.73E-09	8
4				<i>Early elongation complex with hyperphosphorylated Pol II CTD</i>	7.25E-09	6
4				<i>HIV-1 early elongation complex with hyperphosphorylated Pol II CTD</i>	7.25E-09	6
4				<i>RNA Polymerase II Transcription Elongation</i>	7.26E-09	7
4				<i>Recruitment of elongation factors to form HIV-1 elongation complex</i>	7.26E-09	7
4				<i>Addition of nucleotides leads to transcript elongation</i>	7.26E-09	7
4				<i>Elongation complex</i>	7.26E-09	7
4				<i>HIV-1 elongation complex</i>	7.26E-09	7
4				<i>Formation of RNA Pol II elongation complex</i>	1.08E-08	7
4				<i>Formation of HIV-1 elongation complex in the absence of HIV-1 Tat</i>	1.08E-08	7
4				<i>Hyperphosphorylation (Ser2) of RNA Pol II CTD by P-TEFb complex</i>	1.33E-08	6
4				APC/C:Cdh1-mediated degradation of Skp2	1.77E-08	8
4				Spliceosomal B Complex	1.79E-08	9
4				Formation of the Spliceosomal B Complex	1.79E-08	9
4				Destruction of AUF1 and mRNA	2.26E-08	7
4				Proteasomal cleavage of exogenous antigen	2.26E-08	7

Appendix

4				26S proteasome	2.26E-08	7
4				Degradation of multiubiquitinated cell cycle proteins	2.94E-08	8
4				<i>Formation of transcription-coupled NER (TC-NER) repair complex</i>	2.98E-08	6
4				<i>Dual incision reaction in TC-NER</i>	2.98E-08	6
4				<i>Assembly of repair proteins at the site of Pol II blockage</i>	2.98E-08	6
4				<i>Displacement of stalled Pol II from the lesion site</i>	2.98E-08	6
4				Cross-presentation of soluble exogenous antigens (endosomes)	3.73E-08	7
4				Proteasomal cleavage of substrate	4.37E-08	7
4				26S proteasome degrades ODC holoenzyme complex	4.37E-08	7
4				<i>HIV-1 arrested processive elongation complex</i>	4.82E-08	6
4				<i>HIV-1 paused processive elongation complex</i>	4.82E-08	6
4				<i>2-4 nt.backtracking of Pol II complex on the template leading to elongation pausing</i>	4.82E-08	6
4				<i>Resumption of elongation after recovery from pausing</i>	4.82E-08	6
4				<i>Abortive termination of HIV-1 elongation after arrest</i>	4.82E-08	6
4				<i>Elongation arrest and recovery</i>	4.82E-08	6
4				<i>HIV-1 elongation arrest and recovery</i>	4.82E-08	6
4				<i>Pausing and recovery of HIV-1 elongation</i>	4.82E-08	6
4				<i>Processive elongation complex</i>	4.82E-08	6
4				<i>Arrested processive elongation complex</i>	4.82E-08	6
4				<i>Paused processive elongation complex</i>	4.82E-08	6
4				<i>HIV-1 processive elongation complex</i>	4.82E-08	6
4				<i>Resumption of elongation of HIV-1 transcript after recovery from pausing</i>	4.82E-08	6
4				<i>HIV-1 aborted elongation complex after arrest</i>	4.82E-08	6
4				<i>Abortive termination of elongation after arrest</i>	4.82E-08	6
4				<i>Elongating transcript encounters a lesion in the template</i>	4.82E-08	6

Appendix

4				<i>Aborted elongation complex after arrest</i>	4.82E-08	6
4				<i>Elongation complex prior to separation</i>	4.82E-08	6
4				<i>Elongation complex with separated and uncleaved transcript</i>	4.82E-08	6
4				<i>2-4 nt.backtracking of Pol II complex on the HIV-1 template leading to elongation pausing</i>	4.82E-08	6
4				Proteasome mediated degradation of Cyclin D1	5.11E-08	7
4				Regulation of activated PAK-2p34 by proteasome mediated degradation	5.11E-08	7
4				Proteasome mediated degradation of COP1	5.11E-08	7
4				Proteolytic degradation of ubiquitinated-Cdc25A	5.11E-08	7
4				Ubiquitinated geminin is degraded by the proteasome	5.11E-08	7
4				Ubiquitinated Cdc6 is degraded by the proteasome	5.11E-08	7
4				Ubiquitinated Orc1 is degraded by the proteasome	5.11E-08	7
4				Proteasome mediated degradation of PAK-2p34	5.11E-08	7
4				Ubiquitin-dependent degradation of Cyclin D1	5.94E-08	7
4				CDK-mediated phosphorylation and removal of Cdc6	5.94E-08	7
4				Autodegradation of the E3 ubiquitin ligase COP1	6.89E-08	7
4				Ubiquitin Mediated Degradation of Phosphorylated Cdc25A	6.89E-08	7
4				Destabilization of mRNA by AUF1 (hnRNP D0)	7.96E-08	7
4				Antigen processing: Ubiquitination & Proteasome degradation	9.17E-08	7
4				SCF-mediated degradation of Emi1	9.17E-08	7
4				SCF-beta-TrCP mediated degradation of Emi1	9.17E-08	7
4				Degradation of beta-catenin by the destruction complex	1.05E-07	7
4				Degradation of ubiquitinated -beta catenin by the proteasome	1.05E-07	7
4				Spliceosomal A Complex	1.23E-07	8
4				CDT1 association with the CDC6:ORC:origin complex	1.57E-07	7
4				SRP receptor:SRP:ribosome:polypeptide+signal	2.19E-07	8

Appendix

4				<i>Addition of Nucleotides 5 through 9 on the growing Transcript</i>	2.37E-07	6
4				<i>HIV-1 transcription complex containing 4 nucleotide long transcript</i>	2.37E-07	6
4				<i>Addition of nucleotides 5 through 9 on the growing HIV-1 transcript</i>	2.37E-07	6
4				<i>pol II transcription complex containing 4 nucleotide long transcript</i>	2.37E-07	6
4				<i>RNA Polymerase II Transcription Initiation</i>	2.82E-07	6
4				<i>RNA Polymerase II Transcription Pre-Initiation And Promoter Opening</i>	2.82E-07	6
4				<i>RNA Polymerase II Promoter Opening: First Transition</i>	2.82E-07	6
4				Newly Formed Phosphodiester Bond Stabilized and PPi Released	2.82E-07	6
4				<i>RNA Polymerase II Promoter Escape</i>	2.82E-07	6
4				<i>Pol II Promoter Escape Complex</i>	2.82E-07	6
4				<i>HIV-1 Transcription Initiation</i>	2.82E-07	6
4				<i>Transcription of the HIV genome</i>	2.82E-07	6
4				<i>Addition of the third nucleotide on the nascent HIV-1 transcript</i>	2.82E-07	6
4				<i>Addition of the fourth nucleotide on the nascent HIV-1 transcript: Second Transition</i>	2.82E-07	6
4				<i>Fall Back to Closed Pre-initiation Complex</i>	2.82E-07	6
4				<i>Nucleophilic attack by 3'-hydroxyl oxygen of nascent HIV-1 transcript on the Alpha phosphate of NTP</i>	2.82E-07	6
4				Newly formed phosphodiester bond stabilized and PPi released	2.82E-07	6
4				<i>NTP binds active site of RNA Polymerase II in HIV-1 open pre-initiation complex</i>	2.82E-07	6
4				<i>Addition of the fourth nucleotide on the Nascent Transcript: Second Transition</i>	2.82E-07	6
4				<i>HIV-1 transcription complex containing 3 nucleotide long transcript</i>	2.82E-07	6
4				<i>HIV-1 initiation complex with phosphodiester-PPi intermediate</i>	2.82E-07	6
4				<i>HIV-1 open pre-initiation complex</i>	2.82E-07	6
4				<i>HIV-1 Promoter Escape Complex</i>	2.82E-07	6
4				<i>RNA Polymerase II HIV-1 Promoter Escape</i>	2.82E-07	6

Appendix

4				<i>HIV-1 closed pre-initiation complex</i>	2.82E-07	6
4				<i>pol II transcription complex</i>	2.82E-07	6
4				<i>NTP Binds Active Site of RNA Polymerase II</i>	2.82E-07	6
4				<i>Nucleophilic Attack by 3'-hydroxyl Oxygen of nascent transcript on the Alpha Phosphate of NTP</i>	2.82E-07	6
4				<i>Addition of the third nucleotide on the nascent transcript</i>	2.82E-07	6
4				<i>HIV-1 initiation complex</i>	2.82E-07	6
4				<i>HIV-1 Promoter Opening: First Transition</i>	2.82E-07	6
4				<i>HIV-1 transcription complex</i>	2.82E-07	6
4				<i>pol II transcription complex containing 3 Nucleotide long transcript</i>	2.82E-07	6
4				<i>Pol II Initiation complex with phosphodiester-PPi intermediate</i>	2.82E-07	6
4				<i>pol II closed pre-initiation complex</i>	2.82E-07	6
4				<i>Pol II initiation complex</i>	2.82E-07	6
4				<i>pol II open pre-initiation complex</i>	2.82E-07	6
4				Formation of a pool of free 40S subunits	3.13E-07	8
4				Degradation of multiubiquitinated Securin	3.64E-07	7
4				APC/C:Cdc20 mediated degradation of Securin	3.64E-07	7
4				mRNA Splicing - Minor Pathway	3.91E-07	6
4				ATAC B Complex	3.91E-07	6
4				ATAC C Complex with lariat containing 5'-end cleaved mRNA	3.91E-07	6
4				ATAC C Complex	3.91E-07	6
4				Formation of AT-AC C complex	3.91E-07	6
4				ATAC spliceosome mediated 3' splice site cleavage, exon ligation	3.91E-07	6
4				Translated mRNA Complex with Premature Termination Codon Preceding Exon Junction	4.05E-07	8
4				Cdc20:Phospho-APC/C mediated degradation of Cyclin A	5.06E-07	7
4				Degradation multiubiquitinated Cyclin A	5.06E-07	7

Appendix

4				SMG1:Phosphorylated UPF1:EJC:Translated mRNA	5.62E-07	8
4				SMG1 Phosphorylates UPF1 (Enhanced by Exon Junction Complex)	5.62E-07	8
4				SMG1:UPF1:EJC:Translated mRNA	5.62E-07	8
4				L13a-mediated translational silencing of Ceruloplasmin expression	7.12E-07	8
4				Phosphorylated UPF1:SMG5:SMG7:SMG6:PP2A:Translated mRNA	7.12E-07	8
4				SMG6 Cleaves mRNA with Premature Termination Codon	7.12E-07	8
4				GTP hydrolysis and joining of the 60S ribosomal subunit	7.69E-07	8
4				<i>Complex I - NADH:Ubiquinone oxidoreductase</i>	8.20E-07	6
4				<i>NADH enters the respiratory chain at Complex I</i>	8.20E-07	6
4				Nonsense Mediated Decay Enhanced by the Exon Junction Complex	8.96E-07	8
4				Phosphorylated UPF1 Recruits SMG5, SMG7, SMG6, and PP2A	8.96E-07	8
4				HIV-1 promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF complex*	1.45E-06	5
4				pol II promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF complex	1.45E-06	5
4				Formation of the Early Elongation Complex	1.75E-06	5
4				Formation of the HIV-1 Early Elongation Complex	1.75E-06	5
4				Cyclin D associated events in G1	1.75E-06	5
4				80S Ribosome:mRNA:peptidyl-tRNA with elongating peptide	1.77E-06	7
4				Peptide transfer from P-site tRNA to the A-site tRNA	1.77E-06	7
4				Elongation complex with growing peptide chain	1.77E-06	7
4				80S ribosome	1.77E-06	7
4				membrane-bound ribosome:mRNA:cleaved polypeptide	1.77E-06	7
4				80S:Met-tRNAi:mRNA:aminoacyl-tRNA	1.77E-06	7
4				ribosome:mRNA:polypeptide+signal	1.77E-06	7
4				80S:Met-tRNAi:mRNA	1.77E-06	7
4				membrane-bound ribosome:mRNA:polypeptide+signal	1.77E-06	7

Appendix

4				Hydrolysis of eEF1A:GTP	1.93E-06	7
4				80S:Met-tRNAi:mRNA:eIF5B:GTP	1.93E-06	7
4				80S:aminoacyl tRNA:mRNA:eEF1A:GTP	1.93E-06	7
4				eIF5B:GTP is hydrolyzed and released	1.93E-06	7
4				The 60S subunit joins the translation initiation complex	1.93E-06	7
4				eRF3-GDP:eRF1:80S Ribosome:mRNA:peptidyl-tRNA Complex	2.10E-06	7
4				eRF3-GTP:eRF1:80S Ribosome:mRNA:peptidyl-tRNA Complex	2.10E-06	7
4				GTP Hydrolysis by eRF3 bound to the eRF1:mRNA:polypeptide:80S Ribosome complex	2.10E-06	7
4				Translocation of ribosome by 3 bases in the 3' direction	2.10E-06	7
4				eRF3-GDP:eRF1:80S Ribosome:mRNA:tRNA Complex	2.10E-06	7
4				Peptide chain elongation	2.10E-06	7
4				Eukaryotic Translation Termination	2.10E-06	7
4				Polypeptide release from the eRF3-GDP:eRF1:mRNA:80S Ribosome complex	2.10E-06	7
4				pol II promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF:TFIIE complex	2.11E-06	5
4				SRP:polypeptide+signal:ribosome	2.70E-06	7
4				Translated mRNA Complex with Premature Termination Codon Not Preceding Exon Junction	2.93E-06	7
4				ATAC A Complex	2.98E-06	5
4				UPF1:eRF3 Complex on Translated mRNA	3.17E-06	7
4				Degradation of multiubiquitinated Cdh1	4.32E-06	6
4				Autodegradation of Cdh1 by Cdh1:APC/C	4.32E-06	6
4				Formation of the Spliceosomal E complex	6.40E-06	6
4				Spliceosomal E Complex	6.40E-06	6
4				Interferon alpha/beta signaling	7.03E-06	6
4				intron-containing complex	7.37E-06	5

Appendix

5	T cell receptor signaling pathway	2.82E-13	12	Interferon alpha/beta signaling	9.22E-07	6
5	MAPK signaling pathway	1.29E-08	12	Respiratory electron transport	2.60E-06	6
5	Osteoclast differentiation	1.40E-08	9	<i>Expression of IFN-induced genes</i>	3.59E-06	5
5	B cell receptor signaling pathway	9.49E-08	7			
5	Cytokine-cytokine receptor interaction	1.16E-07	11			
5	<i>Pathogenic Escherichia coli infection</i>	2.89E-07	6			
5	Chagas disease (American trypanosomiasis)	9.15E-07	7			
5	Jak-STAT signaling pathway	1.38E-06	8			
5	Cell cycle	3.12E-06	7			
5	Chemokine signaling pathway	5.02E-06	8			
5	<i>Toxoplasmosis</i>	5.31E-06	7			
5	<i>Apoptosis</i>	5.41E-06	6			
EXPERIMENT: E-GEOD-43582, Upregulated Canonical Pathways						
Maximum stable clusters : 5						
Functional enrichment p-value <10⁻⁵						
	KEGG pathway	p-value	# of proteins	Reactome pathway	p-value	# of proteins
1	Olfactory transduction	7.26E-19	37	Olfactory Receptor - G Protein olfactory trimer complex formation	2.18E-13	29
1	Cell cycle	3.33E-16	21	OR - G Protein Trimer Complex	2.18E-13	29
1	<i>DNA replication</i>	1.99E-13	12	DNA polymerase alpha:primase binds at the origin	2.44E-13	11

Appendix

1	Oocyte meiosis	1.45E-11	16	Activation of the pre-replicative complex	3.88E-13	11
1	Progesterone-mediated oocyte maturation	4.38E-11	14	Spliceosomal Intermediate C Complex	3.61E-12	16
1	Metabolic pathways	1.77E-07	42	Spliceosomal Active C Complex	3.61E-12	16
1	Lysosome	5.04E-07	12	Spliceosomal active C complex with lariat containing, 5'-end cleaved pre-mRNP:CBC complex	4.24E-12	16
1	N-Glycan biosynthesis	7.05E-07	8	Lariat Formation and 5'-Splice Site Cleavage	4.24E-12	16
1	Ubiquitin mediated proteolysis	1.53E-06	12	Formation of an intermediate Spliceosomal C complex	4.97E-12	16
1	Aminoacyl-tRNA biosynthesis	3.48E-06	7	Exon Junction Complex	5.83E-12	16
1	Selenocompound metabolism	5.29E-06	5	Cleavage at the 3'-Splice Site and Exon Ligation	5.83E-12	16
1	Nucleotide excision repair	6.64E-06	7	mRNA Splicing - Major Pathway	7.95E-12	16
1				Spliceosomal B Complex	9.01E-12	15
1				Formation of the Spliceosomal B Complex	9.01E-12	15
1				Microtubule-bound kinetochore	1.02E-11	14
1				Kinetochore	1.02E-11	14
1				DNA polymerase epsilon binds at the origin	6.36E-11	9
1				Orc1 removal from chromatin	4.22E-10	12
1				RPA:Cdc45:CDK:DDK:Mcm10:pre-replicative complex	1.01E-09	8
1				Unwinding of DNA	3.13E-09	6
1				Unwinding complex at replication fork	3.13E-09	6
1				intron-containing complex	5.04E-09	9
1				Formation of the Spliceosomal E complex	4.81E-08	10
1				Spliceosomal E Complex	4.81E-08	10
1				Activation of ATR in response to replication stress	4.87E-08	8

Appendix

1				<i>U4:U5:U6 trisnRNP complex</i>	6.97E-08	6
1				snRNP nuclear import and release	9.94E-08	8
1				<i>hnRNP proteins</i>	1.68E-07	6
1				<i>FA core complex</i>	2.36E-07	5
1				snRNP Assembly	4.14E-07	8
1				<i>Cross-presentation of soluble exogenous antigens (endosomes)</i>	4.97E-07	8
1				Spliceosomal A Complex	5.07E-07	10
1				<i>G0 and Early G1</i>	6.97E-07	6
1				<i>Proteolytic degradation of ubiquitinated-Cdc25A</i>	7.05E-07	8
1				<i>Cyclin A/B1 associated events during G2/M transition</i>	7.23E-07	5
1				<i>CDK-mediated phosphorylation and removal of Cdc6</i>	8.34E-07	8
1				<i>Ubiquitin Mediated Degradation of Phosphorylated Cdc25A</i>	9.83E-07	8
1				<i>Degradation of ubiquitinated p27/p21 by the 26S proteasome</i>	9.83E-07	8
1				<i>SCF(Skp2)-mediated degradation of p27/p21</i>	9.83E-07	8
1				<i>ER-Phagosome pathway</i>	9.99E-07	9
1				Degradation of multiubiquitinated cell cycle proteins	9.99E-07	9
1				APC/C:Cdh1 mediated degradation of Cdc20 and other APC/C:Cdh1 targeted proteins in late mitosis/early G1	1.14E-06	9
1				<i>U5 snRNP</i>	1.78E-06	5
1				<i>Removal of the Flap Intermediate</i>	1.78E-06	5
1				<i>U4 ATAC:U5:U6 ATAC Complex</i>	1.78E-06	5
1				<i>26S proteasome</i>	2.44E-06	7
1				<i>Monoubiquitination of FANCI by the FA ubiquitin ligase complex</i>	2.63E-06	5
1				Degradation of multiubiquitinated Cdh1	3.69E-06	8
1				Autodegradation of Cdh1 by Cdh1:APC/C	3.69E-06	8
1				<i>Monoubiquitination of FANCD2 by the FA ubiquitin ligase complex</i>	3.78E-06	5

Appendix

1				<i>Destruction of AUF1 and mRNA</i>	4.12E-06	7
1				<i>Proteasomal cleavage of exogenous antigen</i>	4.12E-06	7
1				cNAP-1 depleted centrosome	4.21E-06	8
1				Nlp-depleted centrosome	4.21E-06	8
1				Plk1-mediated phosphorylation of Nlp	4.79E-06	8
1				Loss of C-Nap-1 from centrosomes	4.79E-06	8
1				Centrosomes containing recruited CDK11p58	4.79E-06	8
1				centrosome	4.79E-06	8
1				Centrosome associated Plk1	4.79E-06	8
1				Loss of Nlp from mitotic centrosomes	4.79E-06	8
1				Dissociation of Phospho-Nlp from the centrosome	4.79E-06	8
1				centrosome containing phosphorylated Nlp	4.79E-06	8
1				<i>CDK:DDK:Mcm10:pre-replicative complex</i>	5.29E-06	5
1				APC/C:Cdh1-mediated degradation of Skp2	6.16E-06	8
1				Degradation of multiubiquitinated Securin	6.16E-06	8
1				APC/C:Cdc20 mediated degradation of Securin	6.16E-06	8
1				<i>Cdc45:CDK:DDK:Mcm10:pre-replicative complex</i>	7.23E-06	5
1				<i>Proteasomal cleavage of substrate</i>	7.73E-06	7
1				<i>26S proteasome degrades ODC holoenzyme complex</i>	7.73E-06	7
1				Cdc20:Phospho-APC/C mediated degradation of Cyclin A	8.81E-06	8
1				Degradation multiubiquitinated Cyclin A	8.81E-06	8
1				<i>Regulation of activated PAK-2p34 by proteasome mediated degradation</i>	8.96E-06	7
1				<i>Proteasome mediated degradation of COPI</i>	8.96E-06	7
1				<i>Ubiquitinated Cdc6 is degraded by the proteasome</i>	8.96E-06	7
1				<i>Ubiquitinated Orc1 is degraded by the proteasome</i>	8.96E-06	7

Appendix

1				<i>Proteasome mediated degradation of PAK-2p34</i>	8.96E-06	7
1				<i>Proteasome mediated degradation of Cyclin D1</i>	8.96E-06	7
1				<i>Ubiquitinated geminin is degraded by the proteasome</i>	8.96E-06	7
1				Activation of claspin	9.68E-06	5
1				Cdc45:CDK:DDK:Mcm10:claspin:pre-replicative complex	9.68E-06	5
1				Cdc45:CDK:DDK:Mcm10:Activated claspin:pre-replicative complex	9.68E-06	5
1				<i>Orc1 is phosphorylated by cyclin A/CDK2</i>	9.68E-06	5
1				Mature centrosomes enriched in gamma-TURC complexes	9.89E-06	8
2	<i>Taste transduction</i>	5.40E-11	10	Kinetochore	2.78E-20	18
2	<i>Homologous recombination</i>	1.83E-10	8	<i>Meiotic Recombination</i>	1.02E-15	11
2	Oocyte meiosis	9.22E-10	12	<i>Centromeric Chromatin:CENPH-I Complex</i>	2.46E-08	5
2	Progesterone-mediated oocyte maturation	1.05E-08	10	<i>Formation of Meiotic Heteroduplex</i>	1.23E-07	5
2	Ubiquitin mediated proteolysis	1.13E-08	12	<i>Centromeric Chromatin:CENPH-I:Centromeric Nucleosome:RSF Complex</i>	1.23E-07	5
2	<i>Mismatch repair</i>	8.81E-08	6	<i>Centromeric Chromatin:CENPH-I: Mis18 Complex</i>	2.81E-07	5
2	<i>Base excision repair</i>	8.88E-07	6	<i>Cdc20:phospho-APC/C:Securin complex</i>	4.06E-07	5
2	p53 signaling pathway	4.66E-06	7	Inactivation of APC/C via direct inhibition of the APC/C complex	7.83E-07	5
2	Nucleotide excision repair	5.90E-06	6	MCC:APC/C complex	7.83E-07	5
2	<i>Measles</i>	7.26E-06	9	Activation of APC/C:Cdc20 by dissociation of Cdc20:phospho-APC/C from Cdc20:phospho-APC/C:Mad2:Bub3:BubR1	7.83E-07	5
2				Activation of claspin	1.05E-06	5
2				<i>Deposition of New CENPA-containing Nucleosomes at the Centromere.</i>	1.05E-06	5
2				<i>Centromeric Chromatin: New CENPA Nucleosome:Mis18:HJURP Complex</i>	1.05E-06	5

Appendix

2				Cdc45:CDK:DDK:Mcm10:claspin:pre-replicative complex	1.05E-06	5
2				Cdc45:CDK:DDK:Mcm10:Activated claspin:pre-replicative complex	1.05E-06	5
2				Nek2A:MCC:APC/C complex	1.05E-06	5
2				Cdc2:Cyclin A:MCC:APC/C complex	1.05E-06	5
2				<i>Centromeric Chromatin:CENPH-I:Mis18:HJURP:CENPA Complex</i>	1.05E-06	5
2				Activation of ATR in response to replication stress	1.07E-06	6
2				APC/C:Cdh1-mediated degradation of Skp2	1.39E-06	5
2				<i>Ubiquitination of Securin by phospho-APC/C:Cdc20 complex</i>	1.39E-06	5
2				<i>multiubiquitinated Securin in complex with CDC20:phospho-APC/C</i>	1.39E-06	5
2				<i>cell cycle proteins:phospho-APC/C:Cdh1 complex</i>	1.39E-06	5
2				<i>Regulation of APC/C activators between G1/S and early anaphase</i>	1.39E-06	5
2				<i>multiubiquitinated Skp2:phospho-APC/C:Cdh1 complex</i>	1.39E-06	5
2				Degradation of multiubiquitinated Securin	1.39E-06	5
2				<i>RSF Complex Binds the Centromere.</i>	1.81E-06	5
2				<i>Kinesins</i>	1.81E-06	5
2				<i>Deposition of New CENPA-containing Nucleosomes at the Centromere</i>	1.81E-06	5
2				<i>Ubiquitination of Cyclin A by APC/C:Cdc20 complex</i>	2.95E-06	5
2				<i>Degradation of multiubiquitinated Nek2A</i>	2.95E-06	5
2				<i>Multiubiquitination of Nek2A</i>	2.95E-06	5
2				<i>Multiubiquitinated Nek2A</i>	2.95E-06	5
2				<i>APC-Cdc20 mediated degradation of Nek2A</i>	2.95E-06	5
2				<i>multiubiquitinated Cyclin A associated with MCC:APC/C complex</i>	2.95E-06	5
2				DNA polymerase epsilon binds at the origin	3.70E-06	5
2				Degradation of multiubiquitinated cell cycle proteins	3.70E-06	5
2				<i>Ubiquitination of cell cycle proteins targeted by the APC/C:Cdh1complex</i>	3.70E-06	5

Appendix

2				<i>multiubiquitinated cell cycle protein:APC/C:Cdh1 complex</i>	3.70E-06	5
2				DNA polymerase alpha:primase binds at the origin	8.27E-06	5
2				Activation of the pre-replicative complex	9.91E-06	5
3	Ubiquitin mediated proteolysis	1.61E-21	26	<i>Docking of the Mature intronless derived transcript derived mRNA, TAP and Aly/Ref at the NPC</i>	5.80E-13	11
3	Aminoacyl-tRNA biosynthesis	2.35E-14	13	<i>Transport of Mature mRNA Derived from an Intronless Transcript</i>	8.72E-13	11
3	Metabolic pathways	3.28E-14	54	<i>Transport of the export-competent complex through the NPC</i>	1.88E-12	11
3	Cell cycle	4.13E-12	17	Spliceosomal active C complex with lariat containing, 5'-end cleaved pre-mRNP:CBC complex	2.22E-12	16
3	<i>Basal transcription factors</i>	5.74E-11	11	Lariat Formation and 5'-Splice Site Cleavage	2.22E-12	16
3	Oocyte meiosis	9.79E-09	13	<i>Release from the NPC and Disassembly of the mRNP</i>	2.71E-12	11
3	N-Glycan biosynthesis	3.33E-08	9	Exon Junction Complex	3.06E-12	16
3	<i>Pyrimidine metabolism</i>	3.33E-08	12	Cleavage at the 3'-Splice Site and Exon Ligation	3.06E-12	16
3	<i>Protein processing in endoplasmic reticulum</i>	1.15E-06	13	mRNA Splicing - Major Pathway	4.18E-12	16
3	<i>Inositol phosphate metabolism</i>	2.35E-06	8	Spliceosomal Intermediate C Complex	2.46E-11	15
3	<i>Phosphatidylinositol signaling system</i>	2.36E-06	9	Spliceosomal Active C Complex	2.46E-11	15
3	<i>Wnt signaling pathway</i>	2.82E-06	12	Formation of an intermediate Spliceosomal C complex	3.32E-11	15
3	<i>mRNA surveillance pathway</i>	4.04E-06	9	snRNP Assembly	4.40E-11	11
3	Progesterone-mediated oocyte maturation	4.48E-06	9	<i>Docking of the TAP:EJC Complex with the NPC</i>	9.59E-11	11
3	<i>Purine metabolism</i>	6.79E-06	12	<i>Transport of Mature mRNA derived from an Intron-Containing Transcript</i>	1.23E-10	11
3	p53 signaling pathway	7.25E-06	8	<i>Nuclear export of snRNA transcripts</i>	1.50E-10	9

Appendix

3				<i>Docking of Mature Histone mRNA complex:TAP at the NPC</i>	2.18E-10	9
3				<i>Transport of the SLBP independent Mature mRNA</i>	3.12E-10	9
3				<i>Docking of Mature Replication Dependent Histone mRNA with the NPC</i>	3.12E-10	9
3				<i>Transport of the SLBP Dependant Mature mRNA</i>	4.40E-10	9
3				<i>Nuclear Pore Complex (NPC)</i>	1.62E-09	8
3				<i>nucleoporin-associated Rev:Importin-beta:B23 complex</i>	2.35E-09	8
3				<i>Addition of Nucleotides 5 through 9 on the growing Transcript</i>	2.67E-09	9
3				<i>HIV-1 transcription complex containing 4 nucleotide long transcript</i>	2.67E-09	9
3				<i>Addition of nucleotides 5 through 9 on the growing HIV-1 transcript</i>	2.67E-09	9
3				<i>pol II transcription complex containing 4 nucleotide long transcript</i>	2.67E-09	9
3				<i>Regulation of Glucokinase by Glucokinase Regulatory Protein</i>	3.36E-09	8
3				<i>Transport of the Mature Intronless Transcript Derived Histone mRNA:TAP:Aly/Ref Complex through the NPC</i>	3.36E-09	8
3				<i>Transport of the Mature intronless transcript derived mRNA:TAP:Aly/Ref Complex through the NPC</i>	3.36E-09	8
3				<i>GCK1:GKRP [cytosol] => GCK1:GKRP [nucleoplasm]</i>	3.36E-09	8
3				<i>RNA Polymerase II Transcription Initiation</i>	3.48E-09	9
3				<i>RNA Polymerase II Transcription Pre-Initiation And Promoter Opening</i>	3.48E-09	9
3				<i>RNA Polymerase II Promoter Opening: First Transition</i>	3.48E-09	9
3				Newly Formed Phosphodiester Bond Stabilized and PPi Released	3.48E-09	9
3				<i>RNA Polymerase II Promoter Escape</i>	3.48E-09	9
3				<i>Pol II Promoter Escape Complex</i>	3.48E-09	9
3				<i>HIV-1 Transcription Initiation</i>	3.48E-09	9
3				<i>Transcription of the HIV genome</i>	3.48E-09	9
3				snRNP nuclear import and release	3.48E-09	9
3				<i>Addition of the fourth nucleotide on the Nascent Transcript: Second Transition</i>	3.48E-09	9

Appendix

3				<i>HIV-1 transcription complex containing 3 nucleotide long transcript</i>	3.48E-09	9
3				<i>HIV-1 initiation complex with phosphodiester-PPi intermediate</i>	3.48E-09	9
3				<i>HIV-1 open pre-initiation complex</i>	3.48E-09	9
3				<i>HIV-1 Promoter Escape Complex</i>	3.48E-09	9
3				<i>RNA Polymerase II HIV-1 Promoter Escape</i>	3.48E-09	9
3				<i>NTP Binds Active Site of RNA Polymerase II</i>	3.48E-09	9
3				<i>Nucleophilic Attack by 3'-hydroxyl Oxygen of nascent transcript on the Alpha Phosphate of NTP</i>	3.48E-09	9
3				<i>Addition of the third nucleotide on the nascent transcript</i>	3.48E-09	9
3				<i>HIV-1 transcription complex</i>	3.48E-09	9
3				<i>pol II transcription complex containing 3 Nucleotide long transcript</i>	3.48E-09	9
3				<i>Pol II Initiation complex with phosphodiester-PPi intermediate</i>	3.48E-09	9
3				<i>pol II closed pre-initiation complex</i>	3.48E-09	9
3				<i>Pol II initiation complex</i>	3.48E-09	9
3				<i>pol II open pre-initiation complex</i>	3.48E-09	9
3				<i>HIV-1 Promoter Opening: First Transition</i>	3.48E-09	9
3				<i>Addition of the third nucleotide on the nascent HIV-1 transcript</i>	3.48E-09	9
3				<i>Addition of the fourth nucleotide on the nascent HIV-1 transcript: Second Transition</i>	3.48E-09	9
3				<i>Fall Back to Closed Pre-initiation Complex</i>	3.48E-09	9
3				<i>Nucleophilic attack by 3'-hydroxyl oxygen of nascent HIV-1 transcript on the Alpha phosphate of NTP</i>	3.48E-09	9
3				Newly formed phosphodiester bond stabilized and PPi released	3.48E-09	9
3				<i>NTP binds active site of RNA Polymerase II in HIV-1 open pre-initiation complex</i>	3.48E-09	9
3				<i>HIV-1 closed pre-initiation complex</i>	3.48E-09	9
3				<i>pol II transcription complex</i>	3.48E-09	9
3				<i>HIV-1 initiation complex</i>	3.48E-09	9

Appendix

3				<i>Nuclear import of Rev protein</i>	4.71E-09	8
3				<i>Transport of the Mature Intronless Transcript Derived Histone mRNA:SLBP:TAP:Aly/Ref complex through the NPC</i>	4.71E-09	8
3				<i>Release of the Mature intronless derived mRNA, TAP, and Aly/Ref from the NPC</i>	4.71E-09	8
3				<i>Release of the SLBP independent Histone mRNA from the NPC</i>	4.71E-09	8
3				<i>Release of the Mature intronless transcript derived Histone mRNA:SLBP:eIF4E Complex</i>	6.50E-09	8
3				<i>RNA Polymerase II Pre-transcription Events</i>	7.99E-09	10
3				Microtubule-bound kinetochore	1.59E-08	11
3				Kinetochore	1.59E-08	11
3				Spliceosomal A Complex	3.22E-08	11
3				Formation of the Spliceosomal E complex	3.22E-08	10
3				Spliceosomal E Complex	3.22E-08	10
3				<i>COPII (Coat Protein 2) Mediated Vesicle Transport</i>	9.67E-08	5
3				Spliceosomal B Complex	1.01E-07	11
3				Formation of the Spliceosomal B Complex	1.01E-07	11
3				<i>pol II promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF:TFIIE complex</i>	2.13E-07	7
3				cNAP-1 depleted centrosome	2.61E-07	9
3				Nlp-depleted centrosome	2.61E-07	9
3				Plk1-mediated phosphorylation of Nlp	3.04E-07	9
3				Loss of C-Nap-1 from centrosomes	3.04E-07	9
3				Centrosomes containing recruited CDK11p58	3.04E-07	9
3				centrosome	3.04E-07	9
3				Centrosome associated Plk1	3.04E-07	9
3				Loss of Nlp from mitotic centrosomes	3.04E-07	9
3				Dissociation of Phospho-Nlp from the centrosome	3.04E-07	9

Appendix

3				centrosome containing phosphorylated Nlp	3.04E-07	9
3				<i>Synthesis of PIPs at the Golgi membrane</i>	5.86E-07	5
3				Cdc20:Phospho-APC/C mediated degradation of Cyclin A	6.14E-07	9
3				Degradation multiubiquitinated Cyclin A	6.14E-07	9
3				<i>mRNA 3'-end processing</i>	6.88E-07	7
3				<i>Cleavage of mRNA at the 3'-end</i>	6.88E-07	7
3				<i>Ligated exon containing complex</i>	6.88E-07	7
3				<i>Signaling by SCF-KIT</i>	6.88E-07	7
3				Mature centrosomes enriched in gamma-TURC complexes	7.01E-07	9
3				<i>Antigen processing: Ubiquitination & Proteasome degradation</i>	9.81E-07	8
3				<i>HIV-1 promoter:TFIID:TFIIA:TFIIB complex</i>	1.44E-06	5
3				<i>Biosynthesis of the N-glycan precursor (dolichol lipid-linked oligosaccharide, LLO) and transfer to a nascent protein</i>	1.44E-06	5
3				<i>Cleavage of Intronless Pre-mRNA at 3'-end</i>	1.44E-06	5
3				<i>pol II promoter:TFIID:TFIIA:TFIIB complex</i>	1.44E-06	5
3				<i>Cleavage and polyadenylation of Intronless Pre-mRNA</i>	1.44E-06	5
3				<i>intronless pre-mRNA cleavage complex</i>	1.44E-06	5
3				<i>Processing of Intronless Pre-mRNAs</i>	1.44E-06	5
3				<i>PPARG:Fatty Acid:RXRA:Mediator:Coactivator Complex</i>	1.84E-06	7
3				<i>Expression of CEBPA</i>	1.84E-06	7
3				<i>TRAP coactivator complex</i>	2.14E-06	5
3				<i>Expression of FABP4 (aP2)</i>	2.20E-06	7
3				<i>Expression of Phosphoenolpyruvate carboxykinase 1 (PEPCK-C)</i>	2.20E-06	7
3				<i>Expression of Lipoprotein lipase (LPL)</i>	2.20E-06	7
3				<i>Expression of Leptin</i>	2.20E-06	7
3				<i>Expression of Adiponectin</i>	2.20E-06	7

Appendix

3				<i>Expression of ANGPTL4</i>	2.20E-06	7
3				<i>Expression of CD36 (platelet glycoprotein IV, FAT)</i>	2.20E-06	7
3				<i>Expression of Perilipin (PLIN)</i>	2.20E-06	7
3				Degradation of multiubiquitinated Cdh1	2.69E-06	8
3				Autodegradation of Cdh1 by Cdh1:APC/C	2.69E-06	8
3				<i>HIV-1 promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF complex*</i>	2.74E-06	6
3				<i>pol II promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF complex</i>	2.74E-06	6
3				<i>PPARG:RXRA Heterodimer Binds to Fatty Acid-like Ligands</i>	3.66E-06	7
3				<i>Elongation of pre-rRNA transcript</i>	4.30E-06	5
3				<i>RNA Polymerase I promoter escape complex</i>	4.30E-06	5
3				APC/C:Cdh1-mediated degradation of Skp2	4.50E-06	8
3				Degradation of multiubiquitinated Securin	4.50E-06	8
3				APC/C:Cdc20 mediated degradation of Securin	4.50E-06	8
3				<i>RNA Polymerase I:rRNATranscript:TTF-I:Sal Box Complex</i>	5.87E-06	5
3				<i>RNA Polymerase I Transcription Initiation complex</i>	5.87E-06	5
3				Inactivation of APC/C via direct inhibition of the APC/C complex	5.87E-06	5
3				MCC:APC/C complex	5.87E-06	5
3				Activation of APC/C:Cdc20 by dissociation of Cdc20:phospho-APC/C from Cdc20:phospho-APC/C:Mad2:Bub3:BubR1	5.87E-06	5
3				<i>Loss of Rrn3 from RNA Polymerase I promoter escape complex</i>	5.87E-06	5
3				<i>Mediator Complex (consensus)</i>	6.48E-06	6
3				Degradation of multiubiquitinated cell cycle proteins	7.25E-06	8
3				<i>RNA Polymerase I Transcription Termination</i>	7.87E-06	5
3				<i>Dissociation of PTRF:Polymerase I/Nascent Pre rRNA Complex:TTF-I:Sal Box</i>	7.87E-06	5
3				Nek2A:MCC:APC/C complex	7.87E-06	5
3				Cdc2:Cyclin A:MCC:APC/C complex	7.87E-06	5

Appendix

3				<i>PTRF:Polymerase I/Nascent Pre rRNA Complex:TTF-I:Sal Box</i>	7.87E-06	5
3				<i>capped, methylated pre-mRNP:CBC complex</i>	7.88E-06	6
3				APC/C:Cdh1 mediated degradation of Cdc20 and other APC/C:Cdh1 targeted proteins in late mitosis/early G1	8.13E-06	8
4	Olfactory transduction	1.45E-38	33	Olfactory Receptor - G Protein olfactory trimer complex formation	9.28E-42	34
4				OR - G Protein Trimer Complex	9.28E-42	34
5	No match			No match		
EXPERIMENT: E-GEOD-43582, Downregulated Canonical Pathways						
Maximum stable clusters : 4						
Functional enrichment p -value <10⁻⁵						
cluster	KEGG pathway	p-value	# of proteins	Reactome pathway	p-value	# of proteins
1	Neuroactive ligand-receptor interaction	1.56E-52	60	G alpha (i) signalling events	2.41E-34	38
1	MAPK signaling pathway	4.31E-28	40	The Ligand:GPCR:Gi complex dissociates	9.30E-32	34
1	Cytokine-cytokine receptor interaction	3.55E-27	39	G alpha (q) signalling events	4.25E-20	25
1	Calcium signaling pathway	8.83E-21	28	The Ligand:GPCR:Gq complex dissociates	4.42E-20	24
1	Chemokine signaling pathway	4.29E-20	28	<i>Activation of voltage gated Potassium channels</i>	5.95E-10	10
1	Glutamatergic synapse	9.92E-20	24	<i>Octamer of Voltage gated K⁺ channels</i>	5.95E-10	10
1	Cholinergic synapse	1.42E-11	16	Collagen biosynthesis and modifying enzymes	2.02E-08	10
1	Hematopoietic cell lineage	5.94E-11	14	Association of procollagen chains	1.20E-08	9

Appendix

1	Jak-STAT signaling pathway	2.22E-10	17	The Ligand:GPCR:Gs complex dissociates	4.78E-07	9
1	<i>T cell receptor signaling pathway</i>	7.86E-10	14	G alpha (s) signalling events	1.09E-05	9
1	<i>Protein digestion and absorption</i>	1.29E-09	12	<i>Termination of O-glycan biosynthesis</i>	2.61E-08	7
1	Toll-like receptor signaling pathway	3.86E-09	13	<i>NR-MED1 Coactivator Complex</i>	1.29E-06	7
1	<i>Chagas disease (American trypanosomiasis)</i>	4.38E-09	13	<i>Activation of GIRK/Kir3 Channels</i>	3.35E-08	6
1	Pathways in cancer	5.08E-09	22	<i>GABA B receptor G-protein beta-gamma and Kir3 channel complex</i>	3.35E-08	6
1	Arachidonic acid metabolism	2.02E-08	10	<i>GABA B receptor G-protein beta-gamma complex</i>	3.35E-08	6
1	<i>Salivary secretion</i>	3.79E-08	11	<i>ST6GALNAC3/4 can add a sialic acid to the sialyl T antigen to form the disialyl T antigen</i>	8.70E-08	6
1	<i>Toxoplasmosis</i>	1.07E-07	13	<i>GalNAc alpha-2,6-sialyltransferase II can add a sialic acid to the T antigen at the alpha 6 position</i>	8.70E-08	6
1	Rheumatoid arthritis	1.17E-07	11	<i>The Ligand:GPCR:G12/13 complex dissociates</i>	4.85E-09	5
1	Vascular smooth muscle contraction	1.55E-07	12	<i>HDL-mediated lipid transport</i>	3.51E-07	5
1	<i>Fc epsilon RI signaling pathway</i>	2.14E-07	10	<i>G-protein beta-gamma subunits</i>	9.53E-07	5
1	<i>Osteoclast differentiation</i>	3.62E-07	12	<i>G-beta:G-gamma dimer</i>	1.46E-06	5
1	Amoebiasis	5.82E-07	11	<i>Glucagon:GCGR mediates GTP-GDP exchange</i>	2.17E-06	5
1	Gastric acid secretion	1.35E-06	9	<i>G-protein with G(s) alpha:GDP</i>	2.17E-06	5
1	Cardiac muscle contraction	1.93E-06	9	<i>C1GALT1 transfers galactose to the Tn antigen forming Core 1 glycoproteins (T antigens)</i>	2.17E-06	5
1	<i>Pancreatic secretion</i>	1.99E-06	10	<i>Addition of galactose to Core 6 glycoprotein</i>	2.17E-06	5
1	<i>Linoleic acid metabolism</i>	5.38E-06	6	<i>Addition of GalNAc to the Tn antigen via an alpha-1,6 linkage forms a Core 7 glycoprotein</i>	2.17E-06	5
				<i>Sialyltransferase I can add sialic acid to the T antigen at the alpha 6 position</i>	2.17E-06	5
				<i>ST3GAL1-4 can add a sialic acid to the T antigen at the alpha 3 position</i>	2.17E-06	5

Appendix

				<i>Addition of galactose to the Tn antigen via an alpha-1,3 linkage forms a Core 8 glycoprotein</i>	2.17E-06	5
				<i>Sialyltransferase I can add sialic acid to the Tn antigen at the alpha 6 position</i>	2.17E-06	5
				<i>Addition of GlcNAc to Core 3 forms a Core 4 glycoprotein</i>	2.17E-06	5
				<i>Addition of GlcNAc to the Tn antigen via an alpha-1,3 linkage forms a Core 5 glycoprotein</i>	2.17E-06	5
				<i>Addition of GlcNAc to the Tn antigen via a beta-1,6 linkage forms a Core 6 glycoprotein</i>	2.17E-06	5
				<i>Addition of GalNAc to mucins to form the Tn antigen</i>	2.17E-06	5
				<i>Addition of GlcNAc to the Tn antigen forms a Core 3 glycoprotein</i>	2.17E-06	5
				<i>Addition of GlcNAc to the T antigen forms a Core 2 glycoprotein</i>	2.17E-06	5
2	Neuroactive ligand-receptor interaction	1.26E-27	33	G alpha (i) signalling events	1.84E-12	16
2	MAPK signaling pathway	2.40E-25	31	G alpha (s) signalling events	1.13E-13	14
2	Cytokine-cytokine receptor interaction	4.58E-23	29	G alpha (q) signalling events	2.83E-11	14
2	Calcium signaling pathway	4.41E-13	17	The Ligand:GPCR:Gi complex dissociates	2.83E-11	14
2	Chemokine signaling pathway	1.33E-11	16	The Ligand:GPCR:Gs complex dissociates	1.60E-14	13
2	Cardiac muscle contraction	7.63E-11	11	The Ligand:GPCR:Gq complex dissociates	9.39E-11	13
2	Jak-STAT signaling pathway	1.54E-10	14			
2	Dilated cardiomyopathy	6.68E-10	11			
2	Glutamatergic synapse	1.87E-09	12			
2	Regulation of actin cytoskeleton	7.14E-09	14			
2	Hematopoietic cell lineage	7.48E-09	10			
2	Pathways in cancer	4.73E-08	16			

Appendix

2	Hypertrophic cardiomyopathy (HCM)	7.18E-08	9			
2	GnRH signaling pathway	2.79E-07	9			
2	<i>Alzheimer's disease</i>	2.66E-06	10			
2	Natural killer cell mediated cytotoxicity	2.92E-06	9			
2	<i>Measles</i>	3.32E-06	9			
2	<i>Melanoma</i>	3.33E-06	7			
2	<i>Long-term depression</i>	3.33E-06	7			
2	<i>Autoimmune thyroid disease</i>	6.36E-06	6			
3	Cytokine-cytokine receptor interaction	9.53E-21	25	G alpha (i) signalling events	3.73E-15	17
3	Calcium signaling pathway	3.96E-17	19	Collagen biosynthesis and modifying enzymes	1.76E-11	10
3	Neuroactive ligand-receptor interaction	8.66E-16	21	The Ligand:GPCR:Gi complex dissociates	2.56E-11	13
3	Cholinergic synapse	4.37E-15	15	The Ligand:GPCR:Gq complex dissociates	1.08E-10	12
3	MAPK signaling pathway	1.03E-13	19	G alpha (q) signalling events	3.90E-10	12
3	Chemokine signaling pathway	4.90E-13	16	<i>CCBP2:CCBP2 ligands</i>	2.60E-09	5
3	GnRH signaling pathway	7.66E-12	12	Association of procollagen chains	2.36E-08	7
3	<i>Fructose and mannose metabolism</i>	1.53E-10	8			
3	Hypertrophic cardiomyopathy (HCM)	5.85E-10	10			
3	Vascular smooth muscle contraction	7.35E-10	11			
3	<i>Focal adhesion</i>	2.67E-09	13			

Appendix

3	Amoebiasis	5.49E-09	10			
3	Dilated cardiomyopathy	2.30E-08	9			
3	<i>Arrhythmogenic right ventricular cardiomyopathy (ARVC)</i>	6.89E-08	8			
3	Rheumatoid arthritis	3.27E-07	8			
3	<i>Long-term potentiation</i>	6.50E-07	7			
3	Gastric acid secretion	9.69E-07	7			
3	Jak-STAT signaling pathway	2.18E-06	9			
3	<i>ECM-receptor interaction</i>	3.30E-06	7			
3	<i>Neurotrophin signaling pathway</i>	3.63E-06	8			
3	<i>African trypanosomiasis</i>	4.23E-06	5			
3	<i>Influenza A</i>	4.49E-06	9			
3	Retinol metabolism	4.94E-06	6			
3	Natural killer cell mediated cytotoxicity	5.19E-06	8			
3	<i>Glioma</i>	6.02E-06	6			
3	<i>Melanogenesis</i>	7.97E-06	7			
3	Toll-like receptor signaling pathway	9.12E-06	7			
4	Neuroactive ligand-receptor interaction	9.68E-15	17	The Ligand:GPCR:Gi complex dissociates	4.51E-11	11
4	Cytokine-cytokine receptor interaction	1.68E-12	15	G alpha (i) signalling events	2.83E-10	11
4	Retinol metabolism	3.61E-10	8	The Ligand:GPCR:Gq complex dissociates	1.81E-06	7

Appendix

4	<i>Drug metabolism - cytochrome P450</i>	1.47E-09	8	G alpha (q) signalling events	3.77E-06	7
4	<i>Metabolism of xenobiotics by cytochrome P450</i>	3.42E-08	7			
4	Calcium signaling pathway	1.35E-07	9			
4	Arachidonic acid metabolism	3.67E-07	6			
4	Chemokine signaling pathway	2.57E-06	8			
4	<i>Glycolysis / Gluconeogenesis</i>	9.77E-06	5			
EXPERIMENT: E-GEOD-57418, Upregulated Canonical Pathways						
Maximum stable clusters : 9						
Functional enrichment p -value <10⁻⁵						
cluster	KEGG pathway	p-value	# of proteins	Reactome pathway	p-value	# of proteins
1	MAPK signaling pathway	2.67E-06	8	Destabilization of mRNA by AUF1 (hnRNP D0)	6.74E-07	5
1				Activation of NF-kappaB in B Cells	9.96E-07	5
2	Spliceosome	1.17E-24	25	mRNA Splicing - Major Pathway	2.05E-22	22
2	<i>Primary immunodeficiency</i>	1.55E-15	12	Formation of an intermediate Spliceosomal C complex	2.79E-21	21
2	T cell receptor signaling pathway	6.04E-13	15	Exon Junction Complex	3.48E-21	21
2	Apoptosis	2.31E-09	11	Cleavage at the 3'-Splice Site and Exon Ligation	3.48E-21	21
2	B cell receptor signaling pathway	6.20E-09	10	Spliceosomal B Complex	4.62E-21	20
2	Pathways in cancer	1.70E-08	18	Formation of the Spliceosomal B Complex	4.62E-21	20
2	<i>RNA polymerase</i>	6.29E-07	6	Spliceosomal Intermediate C Complex	4.66E-20	20

Appendix

2	Chronic myeloid leukemia	8.77E-07	8	Spliceosomal Active C Complex	4.66E-20	20
2	<i>Osteoclast differentiation</i>	9.14E-07	10	Spliceosomal active C complex with lariat containing, 5'-end cleaved pre-mRNP:CBC complex	5.79E-20	20
2	VEGF signaling pathway	1.51E-06	8	Lariat Formation and 5'-Splice Site Cleavage	5.79E-20	20
2	Natural killer cell mediated cytotoxicity	1.53E-06	10	Spliceosomal A Complex	1.20E-17	17
2	Tuberculosis	3.49E-06	11	<i>Formation of the Spliceosomal E complex</i>	1.25E-13	13
2				<i>Spliceosomal E Complex</i>	1.25E-13	13
2				mRNA Splicing - Minor Pathway	2.82E-13	11
2				ATAC B Complex	2.82E-13	11
2				ATAC C Complex with lariat containing 5'-end cleaved mRNA	2.82E-13	11
2				ATAC C Complex	2.82E-13	11
2				Formation of AT-AC C complex	2.82E-13	11
2				ATAC spliceosome mediated 3' splice site cleavage, exon ligation	2.82E-13	11
2				<i>RNA Polymerase II CTD (phosphorylated) binds to CE</i>	5.23E-11	8
2				<i>RNA Pol II with phosphorylated CTD: CE complex with activated GT</i>	5.23E-11	8
2				<i>RNA Pol II with phosphorylated CTD: CE complex</i>	5.23E-11	8
2				<i>Transfer of GMP from the capping enzyme GT site to 5'-end of mRNA</i>	7.95E-11	8
2				<i>Methylation of GMP-cap by RNA Methyltransferase</i>	7.95E-11	8
2				<i>SPT5 subunit of Pol II binds the RNA triphosphatase (RTP)</i>	7.95E-11	8
2				<i>Capping complex (initial)</i>	7.95E-11	8
2				<i>Hydrolysis of the 5'-end of the nascent transcript by the capping enzyme</i>	7.95E-11	8
2				<i>RNA Pol II CTD phosphorylation and interaction with CE</i>	7.95E-11	8
2				<i>Formation of the CE:GMP intermediate complex</i>	7.95E-11	8
2				<i>CE:Pol II CTD:Spt5 complex</i>	7.95E-11	8

Appendix

2				<i>Capping complex (intermediate)</i>	7.95E-11	8
2				<i>Capping complex (hydrolyzed)</i>	7.95E-11	8
2				<i>Covalent CE:GMP intermediate complex</i>	7.95E-11	8
2				<i>Capping complex (with freed 5'- GMP)</i>	7.95E-11	8
2				<i>Capping complex (GpppN..)</i>	7.95E-11	8
2				<i>mRNA Capping</i>	1.72E-10	8
2				<i>Internal Methylation of mRNA</i>	1.84E-10	7
2				<i>intron-containing complex</i>	1.89E-10	9
2				<i>post exon ligation complex</i>	7.89E-10	7
2				<i>ATAC A Complex</i>	1.19E-09	8
2				<i>capped, methylated pre-mRNP: CBC complex</i>	1.57E-09	8
2				<i>Addition of nucleotides between position +11 and +30</i>	1.79E-09	7
2				<i>Pol II transcription complex with (ser5) phosphorylated CTD containing extruded transcript to +30</i>	1.79E-09	7
2				<i>Pol II transcription complex containing extruded transcript to +30</i>	1.79E-09	7
2				<i>pol II transcription complex containing 11 nucleotide long transcript</i>	1.79E-09	7
2				<i>pol II transcription complex containing 9 nucleotide long transcript</i>	1.79E-09	7
2				<i>Phosphorylation (Ser5) of RNA pol II CTD</i>	1.79E-09	7
2				<i>Addition of nucleotides 10 and 11 on the growing transcript: Third Transition</i>	1.79E-09	7
2				<i>Addition of nucleotides between position +11 and +30 on HIV-1 transcript</i>	1.79E-09	7
2				<i>RNA Polymearse II: NTP:TFIIF complex</i>	1.79E-09	7
2				<i>Pol II transcription complex containing transcript to +30</i>	1.79E-09	7
2				<i>HIV-1 transcription complex containing 9 nucleotide long transcript</i>	1.79E-09	7
2				<i>HIV-1 transcription complex containing 11 nucleotide long transcript</i>	1.79E-09	7
2				<i>Addition of nucleotides 10 and 11 on the growing HIV-1 transcript: Third Transition</i>	1.79E-09	7

Appendix

2				<i>HIV-1 transcription complex containing extruded transcript to +30</i>	1.79E-09	7
2				<i>HIV-1 transcription complex containing transcript to +30</i>	1.79E-09	7
2				<i>HIV-1 transcription complex with (ser5) phosphorylated CTD containing extruded transcript to +30</i>	1.79E-09	7
2				<i>RNA Polymerase II (unphosphorylated):TFIIF complex</i>	2.63E-09	6
2				<i>RNA Polymerase II holoenzyme complex (hyperphosphorylated):TFIIF complex</i>	2.63E-09	6
2				<i>RNA Polymerase II holoenzyme complex (hypophosphorylated):TFIIF complex</i>	2.63E-09	6
2				<i>HIV-1 Polymerase II (phosphorylated):TFIIF:capped pre-mRNA</i>	2.63E-09	6
2				<i>RNA polymerase II (phosphorylated):TFIIF complex</i>	2.63E-09	6
2				<i>RNA Polymerase II (phosphorylated):TFIIF:capped pre-mRNA</i>	2.63E-09	6
2				<i>pol II transcription complex containing 4-9 nucleotide long transcript</i>	3.70E-09	7
2				<i>HIV-1 transcription complex containing 4-9 nucleotide long transcript</i>	3.70E-09	7
2				<i>RNA Polymerase II Transcription Elongation</i>	4.34E-09	8
2				<i>Recruitment of elongation factors to form HIV-1 elongation complex</i>	4.34E-09	8
2				<i>Addition of nucleotides leads to transcript elongation</i>	4.34E-09	8
2				<i>Elongation complex</i>	4.34E-09	8
2				<i>HIV-1 elongation complex</i>	4.34E-09	8
2				<i>Hypophosphorylation of RNA Pol II CTD by FCP1P protein</i>	5.17E-09	7
2				<i>Formation of RNA Pol II elongation complex</i>	6.88E-09	8
2				<i>Formation of HIV-1 elongation complex in the absence of HIV-1 Tat</i>	6.88E-09	8
2				<i>Antigen Activates B Cell Receptor Leading to Generation of Second Messengers</i>	7.12E-09	7
2				<i>capped, methylated pre-mRNA:CBC Complex</i>	7.54E-09	6
2				<i>HIV-1 capped pre-mRNA:CBC:RNA Pol II (phosphorylated) complex</i>	7.54E-09	6
2				<i>capped pre-mRNA:CBC:RNA Pol II (phosphorylated) complex</i>	7.54E-09	6
2				<i>RNA Pol II (hypophosphorylated):capped pre-mRNA complex</i>	1.20E-08	6

Appendix

2				<i>HIV-1 paused processive elongation complex</i>	1.70E-08	7
2				<i>2-4 nt.backtracking of Pol II complex on the template leading to elongation pausing</i>	1.70E-08	7
2				<i>Resumption of elongation after recovery from pausing</i>	1.70E-08	7
2				<i>Formation of the Early Elongation Complex</i>	1.70E-08	7
2				<i>Abortive termination of HIV-1 elongation after arrest</i>	1.70E-08	7
2				<i>Elongation arrest and recovery</i>	1.70E-08	7
2				<i>HIV-1 elongation arrest and recovery</i>	1.70E-08	7
2				<i>Pausing and recovery of HIV-1 elongation</i>	1.70E-08	7
2				<i>Formation of the HIV-1 Early Elongation Complex</i>	1.70E-08	7
2				<i>Processive elongation complex</i>	1.70E-08	7
2				<i>Arrested processive elongation complex</i>	1.70E-08	7
2				<i>Paused processive elongation complex</i>	1.70E-08	7
2				<i>HIV-1 processive elongation complex</i>	1.70E-08	7
2				<i>Resumption of elongation of HIV-1 transcript after recovery from pausing</i>	1.70E-08	7
2				<i>Aborted elongation complex after arrest</i>	1.70E-08	7
2				<i>Elongation complex prior to separation</i>	1.70E-08	7
2				<i>Elongation complex with separated and uncleaved transcript</i>	1.70E-08	7
2				<i>2-4 nt.backtracking of Pol II complex on the HIV-1 template leading to elongation pausing</i>	1.70E-08	7
2				<i>HIV-1 arrested processive elongation complex</i>	1.70E-08	7
2				<i>HIV-1 aborted elongation complex after arrest</i>	1.70E-08	7
2				<i>Abortive termination of elongation after arrest</i>	1.70E-08	7
2				<i>Elongating transcript encounters a lesion in the template</i>	1.70E-08	7
2				<i>RNA Pol II (hypophosphorylated) complex bound to DSIF protein</i>	1.83E-08	6
2				<i>Docking of the TAP:EJC Complex with the NPC</i>	4.06E-08	8

Appendix

2				<i>Transport of Mature mRNA derived from an Intron-Containing Transcript</i>	4.83E-08	8
2				<i>DSIF:NELF:early elongation complex</i>	5.58E-08	6
2				<i>Abortive elongation of HIV-1 transcript in the absence of Tat</i>	5.58E-08	6
2				<i>DSIF:NELF:early elongation complex after limited nucleotide addition</i>	5.58E-08	6
2				<i>Early elongation complex with separated aborted transcript</i>	5.58E-08	6
2				<i>Aborted early elongation complex</i>	5.58E-08	6
2				<i>Formation of DSIF:NELF:HIV-1 early elongation complex</i>	5.58E-08	6
2				<i>Aborted HIV-1 early elongation complex</i>	5.58E-08	6
2				<i>RNA Polymerase II holoenzyme complex (unphosphorylated)</i>	6.67E-08	5
2				<i>RNA Polymerase II holoenzyme complex (generic)</i>	6.67E-08	5
2				<i>RNA Polymerase II holoenzyme complex (phosphorylated)</i>	6.67E-08	5
2				<i>Active Pol II complex with repaired DNA template:mRNA hybrid</i>	6.67E-08	5
2				<i>Stalled Pol II in TC-NER</i>	6.67E-08	5
2				<i>Stalled Pol II complex with damaged DNA hybrid</i>	6.67E-08	5
2				<i>Active Pol II transcription complex with damaged DNA hybrid</i>	6.67E-08	5
2				<i>RNA Polymerase II holoenzyme complex (hyperphosphorylated)</i>	6.67E-08	5
2				<i>Early elongation complex with hyperphosphorylated Pol II CTD</i>	7.74E-08	6
2				<i>HIV-1 early elongation complex with hyperphosphorylated Pol II CTD</i>	7.74E-08	6
2				<i>Addition of Nucleotides 5 through 9 on the growing Transcript</i>	1.12E-07	7
2				<i>HIV-1 transcription complex containing 4 nucleotide long transcript</i>	1.12E-07	7
2				<i>Addition of nucleotides 5 through 9 on the growing HIV-1 transcript</i>	1.12E-07	7
2				<i>pol II transcription complex containing 4 nucleotide long transcript</i>	1.12E-07	7
2				<i>RNA Polymerase II Pre-transcription Events</i>	1.26E-07	8
2				<i>RNA Polymerase II Transcription Initiation</i>	1.36E-07	7
2				<i>RNA Polymerase II Transcription Pre-Initiation And Promoter Opening</i>	1.36E-07	7

Appendix

2				<i>RNA Polymerase II Promoter Opening: First Transition</i>	1.36E-07	7
2				Newly Formed Phosphodiester Bond Stabilized and PPi Released	1.36E-07	7
2				<i>RNA Polymerase II Promoter Escape</i>	1.36E-07	7
2				<i>Pol II Promoter Escape Complex</i>	1.36E-07	7
2				<i>HIV-1 Transcription Initiation</i>	1.36E-07	7
2				<i>Transcription of the HIV genome</i>	1.36E-07	7
2				<i>Addition of the fourth nucleotide on the Nascent Transcript: Second Transition</i>	1.36E-07	7
2				<i>HIV-1 transcription complex containing 3 nucleotide long transcript</i>	1.36E-07	7
2				<i>HIV-1 initiation complex with phosphodiester-PPi intermediate</i>	1.36E-07	7
2				<i>HIV-1 open pre-initiation complex</i>	1.36E-07	7
2				<i>HIV-1 Promoter Escape Complex</i>	1.36E-07	7
2				<i>RNA Polymerase II HIV-1 Promoter Escape</i>	1.36E-07	7
2				<i>NTP Binds Active Site of RNA Polymerase II</i>	1.36E-07	7
2				<i>Nucleophilic Attack by 3'-hydroxyl Oxygen of nascent transcript on the Alpha Phosphate of NTP</i>	1.36E-07	7
2				<i>Addition of the third nucleotide on the nascent transcript</i>	1.36E-07	7
2				<i>HIV-1 transcription complex</i>	1.36E-07	7
2				<i>pol II transcription complex containing 3 Nucleotide long transcript</i>	1.36E-07	7
2				<i>Pol II Initiation complex with phosphodiester-PPi intermediate</i>	1.36E-07	7
2				<i>pol II closed pre-initiation complex</i>	1.36E-07	7
2				<i>Pol II initiation complex</i>	1.36E-07	7
2				<i>pol II open pre-initiation complex</i>	1.36E-07	7
2				<i>HIV-1 Promoter Opening: First Transition</i>	1.36E-07	7
2				<i>Addition of the third nucleotide on the nascent HIV-1 transcript</i>	1.36E-07	7
2				<i>Addition of the fourth nucleotide on the nascent HIV-1 transcript: Second Transition</i>	1.36E-07	7

Appendix

2				<i>Fall Back to Closed Pre-initiation Complex</i>	1.36E-07	7
2				<i>Nucleophilic attack by 3'-hydroxyl oxygen of nascent HIV-1 transcript on the Alpha phosphate of NTP</i>	1.36E-07	7
2				Newly formed phosphodiester bond stabilized and PPi released	1.36E-07	7
2				<i>NTP binds active site of RNA Polymerase II in HIV-1 open pre-initiation complex</i>	1.36E-07	7
2				<i>HIV-1 closed pre-initiation complex</i>	1.36E-07	7
2				<i>pol II transcription complex</i>	1.36E-07	7
2				<i>HIV-1 initiation complex</i>	1.36E-07	7
2				<i>Activation of Gene Expression by SREBP (SREBF)</i>	1.36E-07	7
2				<i>Hyperphosphorylation (Ser2) of RNA Pol II CTD by P-TEFb complex</i>	1.41E-07	6
2				<i>Formation of transcription-coupled NER (TC-NER) repair complex</i>	3.14E-07	6
2				<i>Dual incision reaction in TC-NER</i>	3.14E-07	6
2				<i>Assembly of repair proteins at the site of Pol II blockage</i>	3.14E-07	6
2				<i>Displacement of stalled Pol II from the lesion site</i>	3.14E-07	6
2				<i>HIV-1 promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF complex*</i>	4.00E-07	6
2				<i>pol II promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF complex</i>	4.00E-07	6
2				<i>pol II promoter:TFIID:TFIIA:TFIIB:Pol II:TFIIF:TFIIE complex</i>	6.29E-07	6
2				<i>Downregulation of SMAD2/3:SMAD4 transcriptional activity</i>	1.16E-06	5
2				<i>Cholesterol biosynthesis</i>	1.16E-06	5
2				<i>Transport of the export-competent complex through the NPC</i>	1.70E-06	6
2				<i>Release from the NPC and Disassembly of the mRNP</i>	2.04E-06	6
2				<i>Downstream TCR signaling</i>	2.86E-06	6
2				<i>MicroRNA (miRNA) Biogenesis</i>	3.44E-06	5
2				<i>PPARG:Fatty Acid:RXRA:Mediator:Coactivator Complex</i>	3.94E-06	6
2				<i>Expression of CEBPA</i>	3.94E-06	6
2				<i>Expression of FABP4 (aP2)</i>	4.60E-06	6

Appendix

2				<i>Expression of Phosphoenolpyruvate carboxykinase 1 (PEPCK-C)</i>	4.60E-06	6
2				<i>Expression of Lipoprotein lipase (LPL)</i>	4.60E-06	6
2				<i>Expression of Leptin</i>	4.60E-06	6
2				<i>Expression of Adiponectin</i>	4.60E-06	6
2				<i>Expression of ANGPTL4</i>	4.60E-06	6
2				<i>Expression of CD36 (platelet glycoprotein IV, FAT)</i>	4.60E-06	6
2				<i>Expression of Perilipin (PLIN)</i>	4.60E-06	6
2				<i>PPARG:RXRA Heterodimer Binds to Fatty Acid-like Ligands</i>	7.10E-06	6
2				<i>Transport of the Mature Intronless Transcript Derived Histone mRNA:TAP:Aly/Ref Complex through the NPC</i>	8.30E-06	5
2				<i>Transport of the Mature intronless transcript derived mRNA:TAP:Aly/Ref Complex through the NPC</i>	8.30E-06	5
3	Pathways in cancer	2.14E-13	19	Exon Junction Complex	9.74E-10	10
3	Chemokine signaling pathway	1.39E-16	18	Cleavage at the 3'-Splice Site and Exon Ligation	9.74E-10	10
3	Insulin signaling pathway	1.92E-16	16	mRNA Splicing - Major Pathway	1.18E-09	10
3	Neurotrophin signaling pathway	1.99E-15	15	intron-containing complex	3.31E-09	7
3	T cell receptor signaling pathway	1.45E-13	13	Spliceosomal Intermediate C Complex	1.35E-08	9
3	Focal adhesion	4.28E-10	13	Spliceosomal Active C Complex	1.35E-08	9
3	ErbB signaling pathway	1.92E-13	12	Spliceosomal active C complex with lariat containing, 5'-end cleaved pre-mRNP:CBC complex	1.47E-08	9
3	Natural killer cell mediated cytotoxicity	6.37E-10	11	Lariat Formation and 5'-Splice Site Cleavage	1.47E-08	9
3	MAPK signaling pathway	1.06E-06	11	Formation of an intermediate Spliceosomal C complex	1.61E-08	9
3	Glioma	5.68E-12	10	Spliceosomal B Complex	9.59E-08	8
3	Renal cell carcinoma	1.76E-11	10	Formation of the Spliceosomal B Complex	9.59E-08	8

Appendix

3	Chronic myeloid leukemia	2.05E-11	10	mRNA Splicing - Minor Pathway	4.82E-06	5
3	Prostate cancer	2.19E-10	10	ATAC B Complex	4.82E-06	5
3	<i>Chagas disease (American trypanosomiasis)</i>	8.83E-10	10	ATAC C Complex with lariat containing 5'-end cleaved mRNA	4.82E-06	5
3	<i>Progesterone-mediated oocyte maturation</i>	2.74E-09	9	ATAC C Complex	4.82E-06	5
3	Influenza A	1.32E-06	9	Formation of AT-AC C complex	4.82E-06	5
3	<i>Endometrial cancer</i>	1.16E-09	8	ATAC spliceosome mediated 3' splice site cleavage, exon ligation	4.82E-06	5
	<i>Non-small cell lung cancer</i>	1.36E-09	8			
	<i>Acute myeloid leukemia</i>	2.17E-09	8			
	<i>Adipocytokine signaling pathway</i>	9.64E-09	8			
	<i>Melanoma</i>	1.38E-08	8			
	B cell receptor signaling pathway	1.95E-08	8			
	VEGF signaling pathway	2.42E-08	8			
	<i>Fc gamma R-mediated phagocytosis</i>	1.14E-07	8			
	<i>Oocyte meiosis</i>	4.39E-07	8			
	Spliceosome	1.35E-06	8			
	<i>Prion diseases</i>	2.18E-09	7			
	<i>mTOR signaling pathway</i>	3.45E-08	7			
	<i>Colorectal cancer</i>	7.63E-08	7			
	Pancreatic cancer	1.73E-07	7			
	<i>Long-term potentiation</i>	2.38E-07	7			

Appendix

	<i>Fc epsilon RI signaling pathway</i>	5.21E-07	7			
	<i>Gap junction</i>	1.23E-06	7			
	Apoptosis	1.23E-06	7			
	Toll-like receptor signaling pathway	3.44E-06	7			
	<i>Cholinergic synapse</i>	7.36E-06	7			
	<i>Vascular smooth muscle contraction</i>	7.81E-06	7			
	<i>Bladder cancer</i>	1.97E-07	6			
	<i>Type II diabetes mellitus</i>	4.08E-07	6			
4	Cytokine-cytokine receptor interaction	1.55E-21	22	None		
	<i>Systemic lupus erythematosus</i>	2.50E-17	14			
	Chemokine signaling pathway	1.93E-13	14			
	Pathways in cancer	3.97E-09	13			
	Focal adhesion	3.34E-08	10			
	<i>Malaria</i>	1.04E-07	6			
	<i>ECM-receptor interaction</i>	1.63E-07	7			
	<i>Rheumatoid arthritis</i>	2.24E-07	7			
	<i>Amoebiasis</i>	6.52E-07	7			
	<i>Pertussis</i>	9.38E-07	6			
	Tuberculosis	1.76E-06	8			
	Toxoplasmosis	2.93E-06	7			
	Apoptosis	3.22E-06	6			

Appendix

	MAPK signaling pathway	4.66E-06	9			
5	Pathways in cancer	7.35E-08	9	<i>26S proteasome</i>	3.90E-08	5
5	Measles	9.61E-10	8	<i>Destruction of AUF1 and mRNA</i>	5.72E-08	5
5	Influenza A	6.99E-09	8	<i>Proteasomal cleavage of exogenous antigen</i>	5.72E-08	5
5	Toll-like receptor signaling pathway	4.36E-09	7	<i>Cross-presentation of soluble exogenous antigens (endosomes)</i>	8.18E-08	5
5	RIG-I-like receptor signaling pathway	1.71E-08	6	<i>26S proteasome degrades ODC holoenzyme complex</i>	9.16E-08	5
5	Hepatitis C	7.23E-07	6	<i>Proteasomal cleavage of substrate</i>	9.16E-08	5
5	Toxoplasmosis	7.57E-07	6	<i>Proteasome mediated degradation of Cyclin D1</i>	1.02E-07	5
5	<i>Proteasome</i>	6.46E-08	5	<i>Ubiquitinated geminin is degraded by the proteasome</i>	1.02E-07	5
5	<i>NOD-like receptor signaling pathway</i>	2.50E-07	5	<i>Ubiquitinated Cdc6 is degraded by the proteasome</i>	1.02E-07	5
5	<i>Cytosolic DNA-sensing pathway</i>	3.26E-07	5	<i>Proteolytic degradation of ubiquitinated-Cdc25A</i>	1.02E-07	5
5				<i>Ubiquitinated Orc1 is degraded by the proteasome</i>	1.02E-07	5
5				<i>Proteasome mediated degradation of PAK-2p34</i>	1.02E-07	5
5				<i>Regulation of activated PAK-2p34 by proteasome mediated degradation</i>	1.02E-07	5
5				<i>Proteasome mediated degradation of COP1</i>	1.02E-07	5
5				<i>Ubiquitin-dependent degradation of Cyclin D1</i>	1.14E-07	5
5				<i>CDK-mediated phosphorylation and removal of Cdc6</i>	1.14E-07	5
5				<i>Autodegradation of the E3 ubiquitin ligase COP1</i>	1.27E-07	5
5				<i>Ubiquitin Mediated Degradation of Phosphorylated Cdc25A</i>	1.27E-07	5
5				Destabilization of mRNA by AUF1 (hnRNP D0)	1.40E-07	5
5				<i>Antigen processing: Ubiquitination & Proteasome degradation</i>	1.55E-07	5
5				<i>SCF-mediated degradation of Emi1</i>	1.55E-07	5
5				<i>SCF-beta-TrCP mediated degradation of Emi1</i>	1.55E-07	5

Appendix

5				<i>Degradation of beta-catenin by the destruction complex</i>	1.72E-07	5
5				<i>Degradation of ubiquitinated -beta catenin by the proteasome</i>	1.72E-07	5
5				Activation of NF-kappaB in B Cells	2.08E-07	5
5				<i>Interferon gamma signaling</i>	2.29E-07	5
5				<i>CDT1 association with the CDC6:ORC:origin complex</i>	2.29E-07	5
5				<i>Degradation of multiubiquitinated Cdh1</i>	2.99E-07	5
5				<i>Autodegradation of Cdh1 by Cdh1:APC/C</i>	2.99E-07	5
5				<i>APC/C:Cdh1-mediated degradation of Skp2</i>	4.20E-07	5
5				<i>Degradation of multiubiquitinated Securin</i>	4.20E-07	5
5				<i>APC/C:Cdc20 mediated degradation of Securin</i>	4.20E-07	5
5				<i>Degradation multiubiquitinated Cyclin A</i>	5.33E-07	5
5				<i>Cdc20:Phospho-APC/C mediated degradation of Cyclin A</i>	5.33E-07	5
5				<i>Degradation of multiubiquitinated cell cycle proteins</i>	5.75E-07	5
5				<i>ER-Phagosome pathway</i>	5.75E-07	5
5				<i>APC/C:Cdh1 mediated degradation of Cdc20 and other APC/C:Cdh1 targeted proteins in late mitosis/early G1</i>	6.20E-07	5
5				<i>SCF(Skp2)-mediated degradation of p27/p21</i>	2.06E-09	6
5				<i>Degradation of ubiquitinated p27/p21 by the 26S proteasome</i>	2.06E-09	6
5				<i>Orc1 removal from chromatin</i>	1.18E-08	6
6	Pathways in cancer	3.49E-11	13	Spliceosomal Intermediate C Complex	8.48E-07	6
6	Jak-STAT signaling pathway	9.53E-14	12	Spliceosomal Active C Complex	8.48E-07	6
6	Measles	1.52E-11	10	Spliceosomal active C complex with lariat containing, 5'-end cleaved pre-mRNP:CBC complex	9.00E-07	6
6	Cytokine-cytokine receptor interaction	1.33E-08	10	Lariat Formation and 5'-Splice Site Cleavage	9.00E-07	6
6	Cell cycle	1.37E-07	7	Formation of an intermediate Spliceosomal C complex	9.55E-07	6

Appendix

6	Neurotrophin signaling pathway	1.53E-07	7	Exon Junction Complex	1.01E-06	6
6	Spliceosome	1.71E-07	7	Cleavage at the 3'-Splice Site and Exon Ligation	1.01E-06	6
6	Hepatitis C	2.25E-07	7	mRNA Splicing - Major Pathway	1.14E-06	6
6	Chemokine signaling pathway	2.29E-06	7	Spliceosomal A Complex	6.23E-06	5
	Regulation of actin cytoskeleton	5.02E-06	7			
	Prostate cancer	4.19E-07	6			
	Influenza A	1.85E-05	6			
	Focal adhesion	4.56E-05	6			
	Pancreatic cancer	2.19E-06	5			
	Small cell lung cancer	8.92E-06	5			
7				None		
8				None		
9				None		
EXPERIMENT: E-GEOD-57418, Downregulated Canonical Pathways						
Maximum stable clusters : 9						
Functional enrichment p -value <10⁻⁵						
cluster	KEGG pathway	p-value	# of proteins	Reactome pathway	p-value	# of proteins
1	none			none		

Appendix

2	Parkinson's disease	2.43E-67	46	Respiratory electron transport	3.15E-37	24
2	Huntington's disease	1.78E-64	49	<i>Complex I - NADH:Ubiquinone oxidoreductase</i>	4.13E-26	16
2	Alzheimer's disease	2.21E-61	46	<i>NADH enters the respiratory chain at Complex I</i>	4.13E-26	16
2	Oxidative phosphorylation	5.13E-59	42	<i>HP subcomplex</i>	1.25E-15	10
2	Metabolic pathways	5.76E-50	72	<i>ATPase-ATP complex</i>	8.76E-11	6
2	<i>Propanoate metabolism</i>	1.31E-14	10	<i>ATPase-ADP and Pi complex</i>	8.76E-11	6
2	Valine, leucine and isoleucine degradation	1.45E-14	11	<i>Enzyme-bound ATP is released</i>	8.76E-11	6
2	Cardiac muscle contraction	6.54E-12	11	<i>ATP is synthesized from ADP and Pi by ATPase</i>	8.76E-11	6
2	<i>Cholinergic synapse</i>	3.36E-11	12	<i>Formation of ATP by chemiosmotic coupling</i>	8.76E-11	6
2	<i>Calcium signaling pathway</i>	4.62E-10	13	<i>ADP and Pi bind to ATPase</i>	8.76E-11	6
2	<i>Citrate cycle (TCA cycle)</i>	1.04E-09	7	<i>ATPase complex</i>	8.76E-11	6
2	Melanogenesis	2.46E-09	10	<i>IP sub-complex</i>	7.55E-11	5
2	Gastric acid secretion	2.46E-09	9	<i>ATPase CF(0)</i>	4.49E-10	5
2	Lysosome	1.97E-08	10			
2	<i>Glutamatergic synapse</i>	2.91E-08	10			
2	<i>Pancreatic secretion</i>	3.35E-08	9			
2	<i>GnRH signaling pathway</i>	5.82E-07	8			
2	<i>Endocrine and other factor-regulated calcium reabsorption</i>	1.04E-06	6			
2	<i>beta-Alanine metabolism</i>	1.41E-06	5			
2	<i>Vascular smooth muscle contraction</i>	1.74E-06	8			

Appendix

2	<i>Amyotrophic lateral sclerosis (ALS)</i>	1.93E-06	6			
2	<i>Salivary secretion</i>	2.02E-06	7			
2	<i>Butanoate metabolism</i>	2.44E-06	5			
3	Lysosome	3.34E-16	13	Respiratory electron transport	8.92E-09	7
3	Oxidative phosphorylation	9.70E-13	11	<i>Transferrin endocytosis and recycling</i>	3.09E-08	5
3	Alzheimer's disease	1.29E-11	11	<i>Rho GTPase cycle</i>	2.75E-07	6
3	Huntington's disease	8.40E-10	10			
3	Cardiac muscle contraction	2.54E-07	6			
3	Parkinson's disease	3.28E-07	7			
3	<i>Tuberculosis</i>	2.77E-06	7			
4	Lysosome	7.81E-11	10	None		
4	Valine, leucine and isoleucine degradation	2.79E-08	6			
4	<i>Hematopoietic cell lineage</i>	7.83E-08	7			
4	Melanogenesis	3.27E-06	6			
5	Metabolic pathways	4.63E-06	10	MHC class II antigen presentation	2.90E-08	5
5	Lysosome	1.13E-06	5			
5	<i>Phagosome</i>	2.68E-06	5			
6	<i>Pertussis</i>	7.51E-09	5	None		

Appendix

6	<i>Complement and coagulation cascades</i>	7.51E-09	5			
6	<i>Chagas disease (American trypanosomiasis)</i>	4.93E-08	5			
7	none			None		
8	Lysosome	2.40E-08	6	None		
9	None			None		

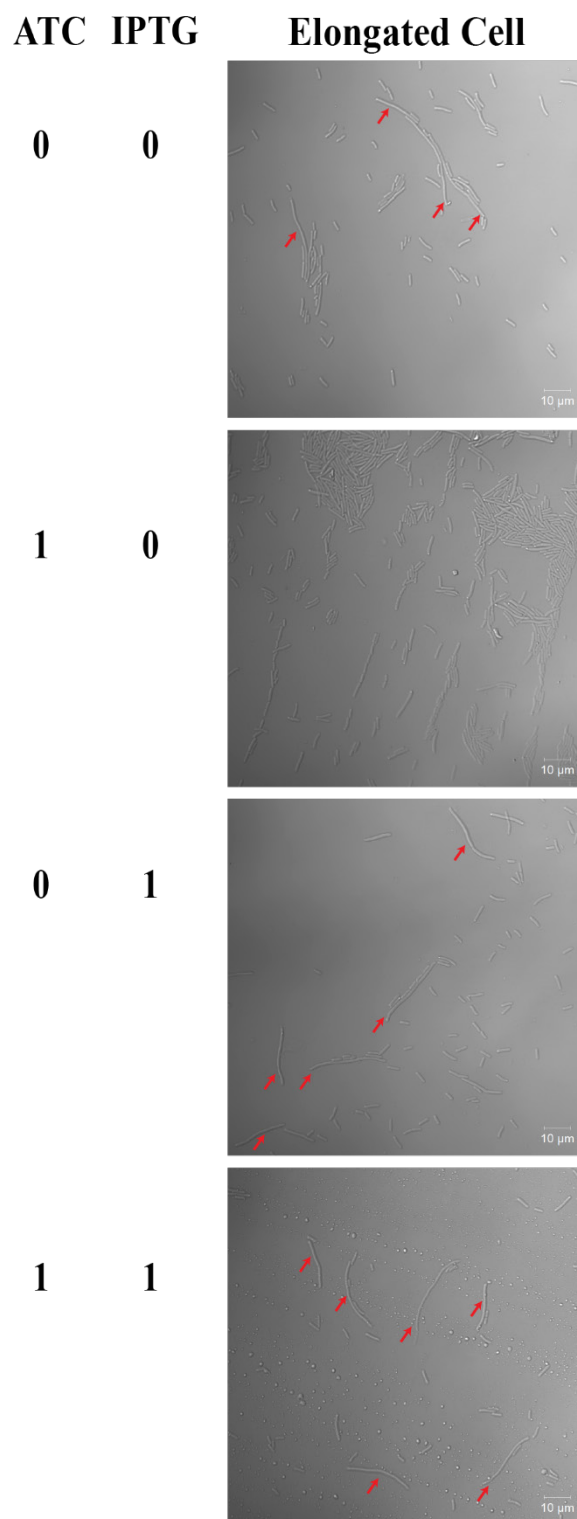


Figure A.1. Remaining dataset of experimental behaviour of the engineered cell in ImPLY circuit. The red arrows are pointing towards elongated cells.

Appendix

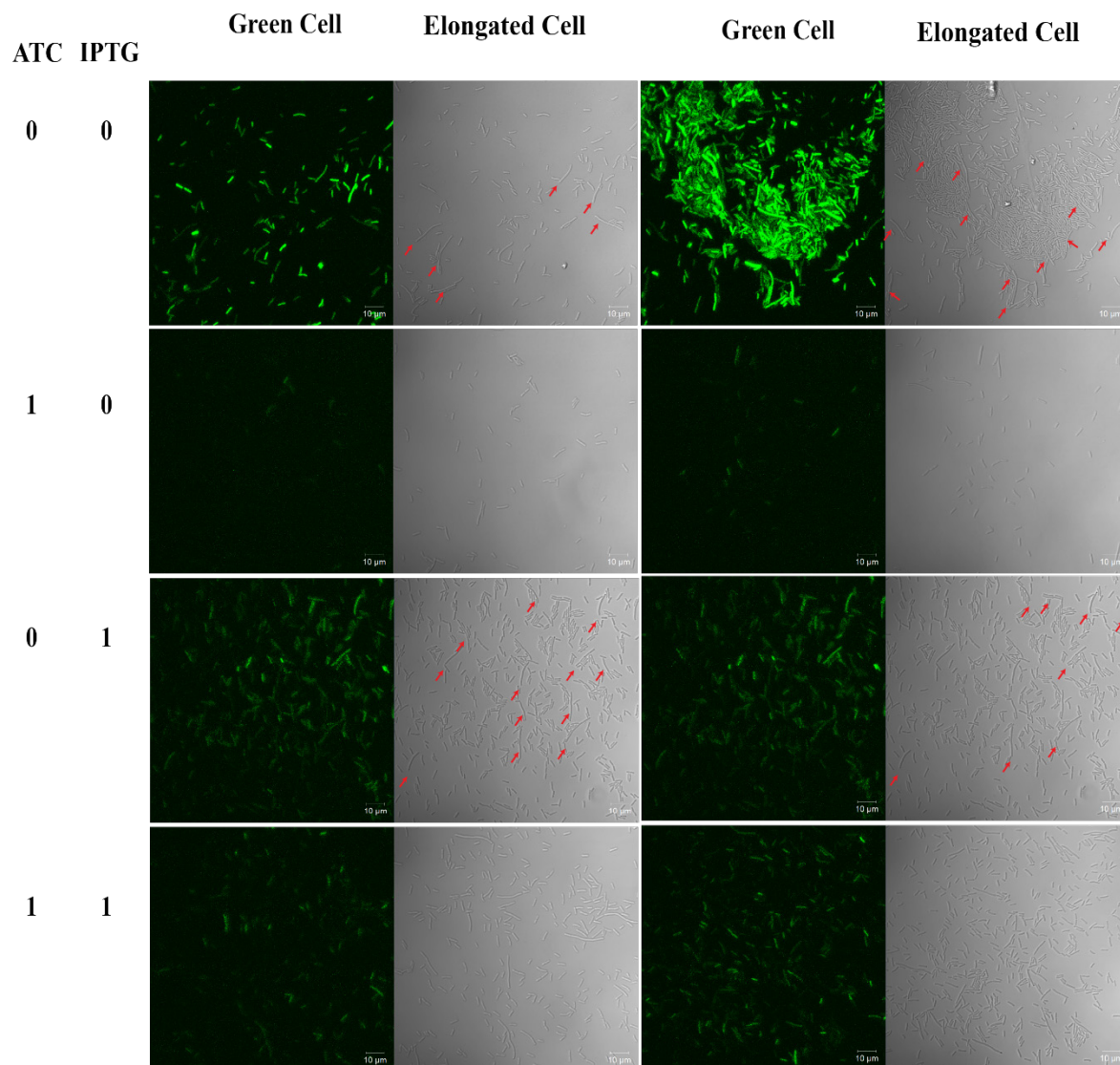


Figure A.2. Remaining dataset of experimental behaviour of the engineered cell in two-input-two-output integrated circuit. The red arrows are pointing towards elongated cells.