MAPPING INTRACELLULAR IMMUNE RESPONSES AGAINST LENTIVIRAL VECTORS IN NATURAL KILLER CELLS USING GENOME SCALE CRISPR-KNOCKOUT

by

AYDAN SARAÇ

Submitted to the Graduate School of Engineering and Natural Sciences

in partial fulfillment, of

the requirements for the degree of

Master of Science

Sabanci University

August, 2017

MAPPING INTRACELLULAR IMMUNE RESPONSES AGAINSTENTIVIRAL VECTORS
IN NATURAL KILLER CELLS USING,
GENOME SCALE CRISPR-KNOCKOUT

APPROVED BY:

Asst. Prof. Tolga Sütlü
(Thesis Supervisor) .....

Prof. Dr Batu Erman .....

Asst. Prof. Saliha Durmuş .....

DATE OF APPROVAL: 27/07/2017

ABSTRACT

MAPPING INTRACELLULAR IMMUNE RESPONSES AGAINST LENTIVIRAL VECTORS IN NATURAL KILLER CELLS USING GENOME SCALE CRISPR-KNOCKOUT

Aydan Saraç

Molecular Biology, Genetics and Bioengineering,

M.Sc. Thesis, 2017

Thesis Supervisor: Tolga Sütlü

Co-supervisor: Abdullah Karadağ

Keywords: NK Cells, lentiviral vectors, antiviral immunity, GeCKO library, CRISPR, NGS

The use of genetically modified cells for therapeutic purposes is an increasing trend that shows great promise. The major hurdle in genetic modification of human cells is the delivery of the gene-of-interest into the cell. Currently, viral vectors are most commonly used for *ex vivo* genetic modification of human cells but there's very little information about the intracellular immune response pathways triggered by viral vector entry.

With the help of advancing next generation sequencing technologies, genome-wide loss-of-function screens have the capacity to produce crucial data to enlighten several unknowns and characterize function of genes comprehensively. The CRISPR/Cas9 system was recently adapted into genome-wide screens and shows great potential in characterizing complex phenotypes. In this study, we used the efficient and high throughput genome editing ability of CRISPR/Cas9 system to discover NK cell resistance mechanisms to lentiviral gene delivery. NK cells are part of innate immune system that act as a first line of defense against viruses. Using NK cells as an immunotherapy agent is not a new idea but genetic modification of NK cells using lentiviral vectors is a very tough task due to its fully armed nature against viral agents. In order to reveal which antiviral pathways become triggered in NK cells during viral vector entry to the cell, we used Genome-scale CRISPR Knock-Out

Libraries(GeCKO). Using a controlled experiment setup; we were able to identify candidate pathways like RIG-I/MDA5 type I interferon secretion and also Toll like receptor (TLR) related pathways that may block virus entry as well as mechanisms that are used by lentiviral vectors to manipulate host cells. Also several genes, like S100A12, I BCL10 and APOB were shown significantly changed during viral vector entry, which implicate some novel pathways in NK cells against lentiviral vectors.

Mapping these pathways is the first step in the venture of overcoming the intracellular defense mechanisms against gene delivery vectors. Identification and manipulation of these pathways could lead to a dramatically increased delivery rate of the transgene, making gene therapy protocols safer and more effective. This may have broad technical applications in order to improve the efficiency of genetic modification of a wide variety of cell types.

ÖZET

CRISPR GENOM SUSTURMA KÜTÜPHANELERİ İLE NK HÜCRELERDE ANTİVİRAL SİNYAL YOLAKLARININ TESPİT EDİLMESİ

Aydan Saraç

Moleküler Biyoloji, Genetik ve Biyomühendislik

M.Sc. Tez, 2017

Tez Danışmanı: Tolga Sütlü

Eş Danışman: Abdullah Karadağ

Anahtar Kelimeler: NK Hücreleri, Lentiviral vektörler, antiviral bağışıklık, GeCKO kütüphanesi, CRISPR, NGS

Günümüzde genetik olarak programlanmış hücrelerin tedavi amaçlı kullanımıyla ilgili yapılan araştırmalarda umut verici gelişmeler yaşanmaktadır. Bu çalışmalarda yaşanan en büyük zorluklardan biri yeni genetik materyalin hedef hücreye nasıl sokulacağıdır. Klinik denemelerde ve laboratuvar modellerinde, insan hücrelerine gen aktarımı üzerine yapılan çalışmalarda viral vektörler sıklıkla kullanılmaktadır. Ancak hücrelerin viral vektör girişini tespit ederek tetikledikleri hücre içi doğal bağışıklık cevaplarını yöneten sinyal yolakları konusunda bilgimiz sınırlıdır.

Yeni Nesil Dizileme (YND) teknolojisindeki gelişmeler ile birlikte, bu hücresel değişikliklerin genom düzeyinde etkilerini saptayabilmek mümkün hale gelmiştir. CRISPR/Cas9 sistemi ile yapılan genom düzeyinde çalışmalar henüz yeni olsa da önemli bir potansiyel taşımaktadır. Bu çalışmada CRISPR/Cas9 sistemiyle genom düzeyinde değişiklik yapma stratejisi kullanılarak Doğal Öldürücü (NK) hücrelerin lentiviral vektörlere moleküler düzeyde verdiği cevapların bir haritası çıkarılmaya çalışılmıştır. NK hücreleri doğal bağışıklık sisteminin bir parçası olup virüslere karşı savunmada önemli rol üstlenen hücrelerdir. NK hücrelerini immunoterapi amacı ile kullanmak yeni bir düşünce olmamakla birlikte, bağışıklık sistemi hücresi olması sebebiyle virüslerin girişine karşı üst düzeyde

koruması olan bu hücreleri lentiviral vektörler ile genetik olarak programlamak oldukça zordur. Bu tez çalışmasında CRISPR tabanlı, genom ölçekli gen susturma kütüphanesi olan GeCKO kullanılarak NK hücrelerin lentiviral vektörlere verdiği cevaplar tespit edilmeye çalışılmıştır. Yapılan kontrollü deney kurgusu sayesinde daha önce başka hücre gruplarında RNA virüslere karşı aktif olduğu gösterilen RIG-I/MDA-5 tip1 interferon sinyal yolağının yanında TLR sinyal yolakları da öne çıkan aday mekanizmalar olarak gösterilmiştir. Bu olası mekanizmalara ek olarak S100A12, APOB ve BCL10 gibi daha önce NK hücrelerde viral vektör girişi ile ilişkilendirilmemiş genler de yeni sinyal yolaklarının olabileceğine dair ipucu vermektedir.

Etkili olan genlerden yola çıkılarak yapılan yolak analizi ile olası savunma mekanizmasının detayları ortaya çıkarılabilecektir. Tespit edilen sinyal yolaklarında hücre içinde yapılabilecek olası değişiklikler ile hücrelere gen aktarımı verimliliği arttırılabilecek, hücreleri tedavi amaçlı daha güvenli ve daha verimli programlayabilmek mümkün olabilecektir.

To my grandfather…

Canım Dedeme…

ACKNOWLEDGEMENTS

First of all, I owe my deepest gratitude to my supervisor Asst. Prof. Tolga Sütlü for believe in me from the beginning, fighting for my acceptance to grad school although I was working as a full time employee in TÜBİTAK. I could not complete this work without his encouragement and support; he was always there to help me. I think he is an inspiring person and we would like to see scientists like him in the academy more frequently.

I would like to thank my esteemed jury members; Prof. Batu Erman and Asst. Prof. Saliha Durmuş for accepting to become jury members and their support and guidance during the writing of this thesis. Especially I would like to express my gratitude to Prof. Batu Erman for being a role model for those who would like to be an immunologist and his non-stop enthusiasm for science. I learned so much from his lectures and I admire his eagerness to teach science and conduct research. I would like to also express my gratitude to Asst. Prof. Saliha Durmuş for her great contribution to the bioinformatics part of this thesis.

I am also grateful to my co-supervisor Dr. Abdullah Karadağ and our institute director Prof. Dr. Şaban Tekin from TUBİTAK MAM GMBE, for their support towards my master program and allowing me to do my thesis experiments in IGBAM lab. Also special thanks to our new bioinformaticians, Altan Kara and Kemal Şanlı for their support during data analysis, also the rest of all GMBE family members who never hesitated to help me when I was in need of help.

Having 5-year work experience in science, I should say Sabanc University was a great experience for me. I had a well-programmed education from great instructors here. Thus, I will always feel appreciated to have this M.Sc. degree from Sabanci University. At this point, I should admit that I am deeply grateful to Prof. Hikmet Budak, for taking the initiative role that allowed me to be accepted to this program.

I would like to thank all the people who, in some way or another, contributed to completion of this thesis.

I had great friends and colleagues here; I would like to thank Dr. Canan Sayitoğlu as being precious member of our lab, for her guidance and support in every circumstance also her contribution to my thesis work. We had great times with Ronay Çetin and Didem Özkazanç during our lecture study hours. Besides, I would like to thank all #sutlulab members Ayhan Parlar, Cevriye Pamukçu, Mertkaya Aras, Lolai Ikromzoda, Pegah Zahedimaram and Alp Ertunga Eyüpoğlu for their help and for being collaborative all the time.

I should say a very special thank you to Gökhan Derdiyok for being the most supportive person during my time in grad school. He has selflessly given more to me than I ever could. There were many times that I thought I could not handle; he was there to show me the light in darkest times, helped me with all his effort and tried to cheer me up with all his understanding.

I will always feel appreciated to have such a great family. Their unconditional love and support in every circumstance are my roots in life, reason of my presence. I could not achieve anything without them. I never would have been able to achieve my goals without support and belief of my mom, Evşen, courage of my dad, Aydın and my beloved sisters Ayşen and Aslıhan. I love you all to the moon and back.

I had my deepest lost during the writing of this thesis work, my grandfather, Lütfü Bayram. He was one of the most valuable people in my life. I learned really a lot from him. I could not have had such a lovely childhood without him, he was the one who always cared and loved. I am lucky since we gathered lots of those great memories together; he will always be with me…

TABLE OF CONTENTS

1. INTRODUCTION 1
   1.1. Natural Killer Cells 1
   1.2. Innate Antiviral Defense Mechanism of NK cells 2
      1.2.1. Toll Like Receptors (TLRs) 3
      1.2.2. RIG-I-Like Receptors (RLRs) 7
      1.2.3. NOD-Like Receptors (NLRs) 9
      1.2.4. C-type Lectin Receptors (CLRs) 11
      1.2.5. Cytoplasmic DNA/ RNA Sensors 13
   1.3. Genetic Manipulation Techniques 14
      1.3.1. Gene Delivery by Lentiviral Vectors 14
      1.3.2. Genome Editing by Targeted Nucleases 16
      1.3.3. CRISPR History and Mechanism 18
   1.4. Genome-wide Screening and GeCKO Library Approach 21
2. AIMS OF THIS STUDY 24
3. MATERIAL AND METHODS 26
   3.1. Materials 26
      3.1.1. Chemicals 26
      3.1.2. Equipment 27
      3.1.3. Buffers and Solutions 28
      3.1.4. Growth Media 28
      3.1.5. Commercial Kits 29
      3.1.6. Enzymes 29
      3.1.7. Bacterial Strains 29
      3.1.8. Mammalian Cell Lines 30
      3.1.9. Plasmids and Oligonucleotides 30
      3.1.10. DNA Ladders 34

3.1.11. DNA Sequencing 34
3.1.12. Software, Computer-Based Programs and Websites 34
3.2. Methods 36
3.2.1. Mammalian Cell Culture 36
3.2.2. Production of GeCKO-CRISPR Lentiviral Vectors 37
3.2.3. Transduction Workflow for GeCKO v.2 Library 38
3.2.4. sgRNA Sequencing Library Preparations 40
4. RESULTS 45
4.1. Transduction of GeCKO Libraries into NK-92 cells 45
4.2. Optimization of Library Sequencing 48
4.3. Bioinformatics Data Analysis 56
5. DISCUSSION 67
6. CONCLUSION 77
7. REFERENCES 78
APPENDIX A 92
APPENDIX B 97
APPENDIX C 106
APPENDIX D 109

LIST OF FIGURES

Figure 1.1 Pathogen recognition receptor families. 3
Figure 1.2 Recognition and signal transduction of RLRs 9
Figure 1.3 Inflamasomme complex and caspase 1 activation pathway. 11
Figure 1.4 Double strand break repairing mechanisms 17
Figure 1.5 CRISPR/Cas locus in bacteria and archaea. 19
Figure 1.6 CRISPR/Cas9 representation 20
Figure 3.1 DNA ladders 34
Figure 3.2 GeCKO v.2 lentiviral transduction workflow 39
Figure 3.3 Lentiviral v.2 backbone from Feng Zhang Lab 40
Figure 3.4 Nested PCR for GeCKO library sgRNA sequencing library. 41
Figure 3.5 Design scheme for PCR2 forward and reverse primers 41
Figure 3.6 E-Gel size selection protocol 42
Figure 3.7 Workflow for NGS 44
Figure 4.1 Workflow timeline for GeCKO sequencing libraries 47
Figure 4.2 PCR1 gradient PCR 49
Figure 4.3 PCR1 annealing temperature 50
Figure 4.4 Bioanalyzer for PCR1 annealing temperature 50
Figure 4.5 Bioanalyzer Screen for PCR 1 60°C ->PCR2 51
Figure 4.6 Gradient PCR2 optimization 51
Figure 4.7 PCR2 cycle determination 52
Figure 4.8 Bioanalyzer of libraries with high sensitivity DNA chip 53
Figure 4.9 Standard curve for qPCR 53
Figure 4.10 Amplification of determined standard and samples 54
Figure 4.11 CT values of six libraries across Standard curve 54
Figure 4.12 Melting curve analysis of libraries in qPCR 55

Figure 4.13 Phred score in HiSeq Control Software....56
Figure 4.14 Histogram of Library A base library....58
Figure 4.15 AGFP+ and AGFP- histograms, respectively....58
Figure 4.16 Histogram of Library B base library....59
Figure 4.17 Library BGFP+ and BGFP- histograms respectively.....59
Figure 4.18 Antiviral pathway analysis scheme....65
Figure 4.19 Proviral pathway analysis scheme....66
Figure 5.1 Several known antiviral pathways against RNA virus....69
Figure C.1 FastQC base library A....106
Figure C.2 FastQC of library A GFP+....106
Figure C.3. FastQC of library A GFP-....107
Figure C.4. FastQC base library B....107
Figure C.5 FastQC base library B GFP+....108
Figure C.6 FastQC base library B GFP-....108

LIST OF TABLES

Table 1-1 GeCKO Library Content 23
Table 3-1 Equipment used in this study 28
Table 3-2 Commercial kits used in the study 29
Table 3-3 Plasmid List 31
Table 3-4 Oligonucleotides list 33
Table 3-5 Software websites and programs 35
Table 4-1 Optimized PCR1 setup 48
Table 4-2 Optimized PCR2 set up 49
Table 4-3 Library concentrations and barcodes 55
Table 4-4 Count summary for library A 57
Table 4-5 Count summary for library B 57
Table 4-6 Command format of MAGeCK tool 60
Table 4-7 AGFP+ gene summary output file 61
Table 4-8 AGFP- gene summary output file 62
Table 4-9 BGFP+ gene summary output 63
Table 4-10 BGFP- gene summary output 64

LIST OF SYMBOLS AND ABBREVIATIONS

α	Alpha
β	Beta
γ	Gamma
κ	Kappa
μ	Micro
AIM 2	absent in melanoma 2
APC	antigen presenting cell
ASC	apoptosis-associated speck-like protein containing a CARD
CARD	caspase recruitment domains
Cas	CRISPR-associated protein
CD	Cluster of Differentiation
cDC	classical dendritic cells
CLR	C-type lectin receptors
CMV	cytomegalovirus
CpG	cytidine-phosphate-guanosine
CRISPR	Clustered Regularly Interspaced Short Pallindromic Repeats
crRNA	CRISPR- RNA
CTLD	C type lectin like domains
DAI	DNA-dependent activator of interferon regulatory factors
DAMPs	damage-associated molecular patterns
DCIR	DC immunoreceptor

DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethylsulfoxade
ds	Double-stranded
DSB	double-stranded breaks
E.coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EMCV	Encephalomyocarditis virus
ENU	ethylnitrosourea
Env	Envelope
FACS	Fluorescence Activated Cell Sorting
FBS	Fetal Bovine Serum
Gag	Group-specific antigen
GeCKO	Genome-scale CRISPR Knock-Out
GFP	Green Fluorescent Protein
GO	Gene Ontology
GOI	Gene-of-interest
gRNA	guide RNA
HCV	Hepatitis C
HDR	homology-directed repair
HEK	Human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodefficiency virus
HSV	herpes simplex virus
IFI16	IFN-gamma-inducible protein 16
IFN	Interferon
IkB	Inhibitor of kB Kinase
IKK	Inhibitor of kB
IPS-1	IFN-β-promoter stimulator 1

IRAK-1	IL-1 receptor-associated receptor kinase
IRF	interferon regulatory factor
LB	Luria Broth
LGP-2	Laboratory of Genetics and Physiology-2
LPS	lipopolysaccharide
LRR	leucine reach repeat
MAVS	Mitochondrial Antiviral Signaling
MDA5	melanoma differentiation-associated gene
MDP	muramyl dipeptide
MGL	Macrophage galactose C-type lectin
MHC	Major Histocompatibility Complex
miRNA	microRNA
MOI	multiplicity of infection
MSRV	multiple sclerosis-associated retrovirus
MV	measles virus
NAP1	NAK-associated protein 1
NDV	Newcastle disease virus
NF-kB	Nuclear Factor kappa B
NGS	Next Generation Sequencing
NHEJ	Nonhomologous end-joining
NK	Natural Killer cells
NLR	NOD-like receptor family, pyrin domain-containing
NLRP	Nucleotide-binding oligomerization domain-containing (NOD)-like receptors
PAM	protospacer-adjacent motif''
PAMP	Pathogen Associated Molecular Patterns
pDC	plasmacytoid dentritic cells
poly(**I:C)**	polyinosine-polycytidylic acid
PRR	Pathogen Recognition Receptor

PTM	posttranslational modification
PYD	pyrin domain
RISC	RNA-induced silencing complex
RLR	RIG like Receptors
RNAi	RNA interference
RPMI	Roswell Park Memorial Institute
RRE	Rev Response Element
RSV	respiratory syncytial virus
RVR	repeat variable residues
SBS	Sequencing by Synthesis
SCGM	Stem Cell Growth Medium
sgRNA	single gRNA
siRNA	short interfering RNAs
ssRNA	single stranded RNA molecules
STING	stimulator of interferon genes
TALEN	transcription activator-like effector nucleases
TAMPs	tumor-associated molecular patterns
TANK	Tank-binding protein
TBK	TNF receptor-associated factor
TIR	Toll–interleukin 1 receptor
TLR	Toll Like receptors
TRAF	TRAF family member-associated NF-κB activator
TRIF	TIR-domain containing adaptor
Vpr	Viral protein r
VSV	vesicular stomatitis virus
VV	Vaccina Virus
wPRE	Woodchuck hepatitis virus post-transcriptional regulatory element
ZFN	Zinc Finger Nuclease

MAGeCK	Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout
RRA	Robust Rank Aggreation

1. INTRODUCTION

1.1. Natural Killer Cells

NK cells are lymphocytes and part of the innate immune system. They constitute around 15% of circulating lymphocytes in blood and a play vital role in the fight against non-self (allogeneic) cells as well as self-cells (autologous) that are under stress factors such as viral, bacterial and parasitic infections or malignant transformation (Punt, Owen, and Caligiuri 2001).

Human NK cells originate from $CD34^{+}$ hematopoietic progenitor cells and are defined by their expression of CD56 on the cell surface along with the absence of CD3 (Robertson and Ritz 1990). Furthermore, NK cells are traditionally divided into two subgroups according to the density of CD56 expression on the cell surface, as $CD56^{dim}$ and $CD56^{bright}$ NK cells. While $CD56^{bright}$ NK cells express low levels of the Fc receptor CD16 and assume a more immunoregulatory role, $CD56^{dim}$ cells express high levels of CD16, comprise 90% of all NK cells and show high cytotoxic activity (Lanier et al. 1986).

In general terms NK cells have two modes of action; they can be cytotoxic directly without immunization and/or they can induce several chemokine and cytokine production at the site of recognition. Although they lack antigen specific response and are considered as part of innate immune system; due to their high capacity of chemokine production, NK cells play a crucial role in the communication between innate and adaptive immunity.

Unlike T and B cells, NK cells do not have antigen specific receptors; they recognize and decide the fate of their target cells thanks to a large repertoire of cell surface receptors. Most of the receptors on NK cells are common among the other hematopoietic cells as well, but apart from these common receptors, NK cells have activating and inhibitory receptors. NK cell receptors are grouped into two as activating and inhibitory receptors according to their effect on NK cell reactivity. The reaction of NK cells against a target is determined through the balance between those two groups of receptors which is a unique phenomenon of NK cells among other lymphocytes. They also have cytokine and adhesion receptors which are mostly important for NK cell maturation and migration.

1.2. Innate Antiviral Defense Mechanism of NK cells

As being a fast and efficient agent in the immune system against viral infections, NK cells potentially have evolved a strong organization to prevent invasion by viral agents. The high resistance of NK cells to viral infection is disadvantageous for studies that are based on transgene expression in NK cells by means of viral vectors. To eliminate these innate antiviral defense mechanisms, recognition and signal transduction events behind the antiviral response of NK cells against viral vectors should be well elucidated.

Pattern Recognition Receptors (PRRs) are a set of receptors which play a crucial role in the innate immune system in detecting microorganisms like viral agents, bacteria or parasites through pathogen associated molecular patterns (PAMPs). PAMPs are mostly structures of microorganisms that are vital to maintain their life such as lipopolysaccharides, phospholipids, CpG-DNA (cytidine-phosphate-guanosine-DNA), dsRNA (double stranded RNA) etc. PRRs are germline coded and they are mostly conserved among other species and kingdoms. While PRRs can be expressed on some specific innate immune cells to do their duty, there are several PRRs that can be found in most of the non-immune cells as well. Each PRR recognizes different targets/ligands (**Figure 1.1**). They can be found on the cell surface or various intracellular compartments as well as extracellular environments like plasma and other body fluids (Janeway and Medzhitov 2002; Akira, Uematsu, and Takeuchi 2006). In the following parts, membrane-bound and intracellular PRRs will be detailed further.

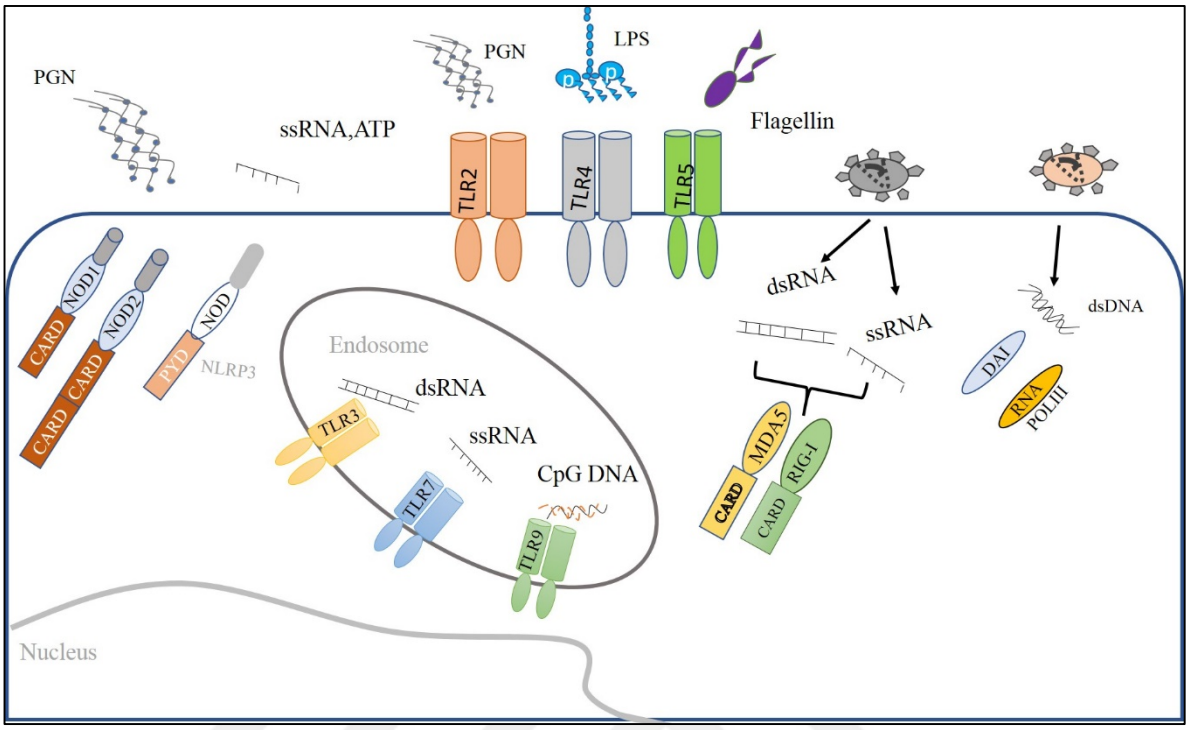

Figure 1.1 Pathogen recognition receptor families. In this figure, some of the well-known PRRs are depicted, TLRs; membrane bound receptors, mainly resides in plasma membrane and endosome. NLRs; present in cytoplasm and catch PGN, ssRNA molecules, RLRs; RIG-I and MDA-5 receptors; detect viral nucleic acid molecules in cytoplasm, DAI and RNApolII are some of dsDNA sensors in the cytoplasm.

1.2.1. Toll Like Receptors (TLRs)

One of the most well studied group of PRRs is undoubtedly the Toll Like receptors(TLRs). They are present on most of the immune cells in cellular compartments like cell membrane, endosomes and lysosomes. TLRs are trans-membrane molecules that have leucine rich repeats (LRR) outside the cell membrane (N terminus) and recognition is achieved through this part. Their cytoplasmic Toll–interleukin 1 receptor (TIR) domain regulates downstream signal transduction events of TLR triggered immune responses. Humans have 10 TLRs identified so far while mice have 13 and TLRs (1-10) are conserved between these two species. Triggering of TLRs is generally followed by production of pro-

inflammatory cytokines and type 1 interferons (IFN-I). In the case of antigen presenting cells (APC) this leads to maturation of the cell and activation of adaptive immunity (Kawai and Akira 2010).

There are two types of TLRs; one is cell surface bound while the other type resides in endosomes. TLR1, 2, 5, 6 and 10 are cell membrane bound receptors while TLR3, 7, 8 and 9 are generally found in endosomes. TLR4 can be found in both. Without any exogenous stimulation, most of these receptors reside in the endoplasmic reticulum, controlled by the Unc93b protein (Tabeta et al. 2006). Cell surface TLRs generally recognize molecules such as lipoproteins, lipids and proteins, while endosome bound TLRs mostly recognize nucleic acids from pathogens. Viral recognition through TLR molecules can be achieved through viral envelope glycoprotein and mostly viral nucleic acid molecules (Kawai and Akira 2010).

Since viruses replicate inside of the cell, their recognition on the outside of the cell is a challenge. A virus is generally taken into the cell by endosomal compartments and after degradation with endosomal enzymes their nucleic acid molecules are released. Host cells generally use these degraded envelope proteins and nucleic acids as PAMP molecules for detection of virus presence. However, nucleic acid sensing inside of cells is also a though task because while dsRNA is a unique molecule to virus and can be detected in an easy way, virus derived ssRNA is harder to differentiate from host cell RNAs (Pichlmair and Sousa 2007).

TLR2 has a broad range of recognition as it sees molecules from bacteria, viruses, fungi and parasites. It can recognize peptidoglycans, lipoproteins and lipopeptides. Several knockout studies related to TLR2 receptors revealed that it forms heterodimeric structures with TLR1, TLR6 and TLR10. In the absence of TLR1 and TLR6, the TLR2 receptor is thought to constitute homodimeric structures which cannot be observed due to the lack of relevant technology. Ligand recognition and downstream signaling events are achieved by these heterodimeric structures and this increases the capacity of recognition, as they can see different molecules as ligands depending on the dimerization partner (Ozinsky et al. 2000; Guan et al. 2010). Also in recent studies, some other accessory molecules and co-receptors have been proposed as helper structures increasing this ligand diversity. In a study related to Vaccinia Virus (VV) infection (a huge, dsDNA virus) it was shown that recognition of the

virus and immune response is regulated by both TLR2-dependent and independent pathways and more importantly, this work showed NK cell activation is achieved through direct TLR2 activation *in vivo* (Martinez, Huang, and Yang 2010). Another study on HIV showed that TLR2 recognizes p17, p24, and gp41 proteins. Although activation of NK cells was shown, the ligand for TLR2 was not found in this study while studies on TLR2 suggested the viral envelope glycoprotein as a ligand (Bieback et al. 2002).

TLR3 can recognize dsRNA. While some viral genomes are composed of double-stranded RNA, some others (positive-strand RNA, dsRNA, DNA) can produce these double-stranded RNA molecules as an intermediates during replication (Weber et al. 2006). TLR3 detection of rhinovirus (reason for common cold disease) has been studied and it was shown that inhibition of TLR3 increased viral replication rate (Hewson et al. 2005). Another example of TLR3 engagement in viral recognition and immune response was focused in a study on West Nile Virus, a ssRNA virus that causes infection in the brain. During replication of this virus dsRNA is generated and TLR3 mediated immune response is triggered (Tian Wang et al. 2004).

TLR4 was the first TLR discovered and it mainly recognizes bacterial lipopolysaccharide (LPS). Later, it was revealed that not only bacterial LPS but also fungal cell wall mannan, protozoal molecules and viral envelope glycoproteins are also recognized by TLR4. TLR4 viral recognition was shown during respiratory syncytial virus (RSV) (a single-stranded negative-sense RNA virus) infection in monocytes. Not only recognition but also replication of the virus was found to be dependent on TLR4 expression rate (Kurt-jones et al. 2000). Also human multiple sclerosis-associated retrovirus (MSRV) envelope protein has been shown to be recognized by TLR4 on endothelial cells (Duperray et al. 2015).

TLR7 and 8 can detect single stranded RNA molecules (ssRNA) and its derivatives containing U or GU repeats. In a recent study on Chinese Racoon dog, Canine distemper virus (ssRNA virus) entry to cells boosted TLR8 expression and its downstream signal transduction. TLR8 also has been in focus after the reveal of its draft sequence, its high sequence similarity among other mammals has been shown (Yong Yang et al. 2016). HIV-1 has been shown to activate TLR7 in plasmacytoid dentritic cells (pDC) (Beignon et al. 2005).

In other studies TLR7 activation after influenza virus infection has been shown (S. S. Diebold et al. 2004).

TLR9 detects unmethylated 2'-deoxyribo CpG (cytidine-phosphate-guanosine) DNA motifs that are only found in bacterial and viral genomes (Akira, Uematsu, and Takeuchi 2006). In studies with herpes simplex virus (HSV-1 and HSV-2) it has been shown that recognition of these DNA virus molecules are achieved by TLR9 in endosomal compartments of pDCs and promotes IFN-α secretion (Lund et al. 2003; Krug et al. 2004).

Downstream Signaling Events

TLR recognition and signaling can differ from cell to cell; although the first line of recognition *in vivo* is generally achieved by antigen presenting cells (APC); other cells like epithelial cells and NK cells also have these receptors. Cells secrete high levels of IFN-I when triggered by virus recognition. IFN-I includes IFN-α and IFN-β, cytokines that are secreted after viral infection and crucial both for innate and adaptive immune responses. IFN-I increases the expression of co-stimulatory molecules like CD80, CD86, and CD40 and this drives APC maturation. Viral ligands are presented on APC plasma membrane which leads to priming of $CD8^+$ and $CD4^+$ T cells. IFN-I is secreted through a set of signaling events which activates transcription factors like NF-κB, ATF2-c-Jun, interferon regulatory factor-3 (IRF3) and interferon regulatory factor-7 (IRF7). Out of these transcription factors, IRF3 and IRF7 are activated only when the virus detection happens while NF-κB and ATF2-c-Jun can also help IRF3 and 7 on the secretion of IFN-I in some cases.

Downstream signaling events in cells depend on the activated receptor. For example, pDCs express TLR7 and 9 but lack TLR3 on the endosomal compartment. On the other hand, human classical dendritic cells (cDC) lack TLR7 and 9 but viral ssRNA molecules are detected through TLR8 and they also express TLR3 on their endosomes. When TLR8 detects virus, at the end of downstream signaling events, they secrete important proinflammatory cytokines such as IL-6, IL-12 and TNFα along with low dosage of IFN-I (Rathinam and Fitzgerald 2011a).

When TLR3 detection of viral dsRNA takes place, TRIF, TIR-domain containing adaptor, recruits IRF3 and IRF3 is phosphorylated by Tank-binding protein (TBK1)/IKKε complex.

Besides, IRF7 is phosphorylated in minimum dosage and these events lead to the secretion of IL-6 and IFN-I. TLR3 detection also activates NF-κB transcription factor with the help of the same adaptor molecule, TRIF. In a previous study by our group on NK cells, BX795 (a small molecule inhibitor of TBK1/IKKε complex) was shown to increase lentiviral transduction efficiency, which also hints that NK cells might detect virus in a TLR3-dependent manner (Sutlu et al. 2012).

When TLR7 and 9 is activated through the detection of the viral genome, TIR domain of the receptors recruits MyD88 cytosolic adaptor, MyD88 interacts with IRAK1 (IL-1 receptor-associated receptor kinase) and TRAF6 (TNF receptor-associated factor-6). This combined unit activates Iκβ kinase α (IKKα) and IRF7 that triggers high dosage of IFN-I secretion.

As given the examples above, TLRs are responsible for membrane bound receptor mediated recognition of PAMPs, over cell surfaces and endosomes, while there are also some other detection events triggered by important cytoplasmic PRRs which will be explained in the next chapter.

1.2.2. RIG-I-Like Receptors (RLRs)

There are two main RNA helicases in the cytoplasm which are found to be triggered by viral RNA infection, one of them is cytoplasmic protein called retinoic acid-inducible gene (RIG-I) and the other one is melanoma differentiation-associated gene (MDA5). These are DExD/H box RNA helicases which have two caspase recruitment domains (CARD), a C-terminal domain (CTD) and one helicase unit that is responsible for ATPase activity. They were identified in 2004 (Yoneyama et al. 2004). Furthermore, there is another DExD/H box RNA helicase; Laboratory of Genetics and Physiology-2 (LGP-2) involved in the detection of viral components in the cytoplasm. This enzyme does not have CARD domains, has only an RNA helicase part. Although its mechanism has not been clarified, it is regulatory role has been shown (Sasai and Yamamoto 2013).

Both RIG-I and MDA5 proteins are stimulated by viral RNA and polyinosine-polycytidylic acid (poly(I:C)) an equivalent oligonucleotide sequence of dsRNA. RIG-I detects viral

ssRNA thanks to presence of 5' triphosphate on viral ssRNA while it cannot see 5' methylguanosine cap of host cell mRNAs (Rehwinkel et al. 2010; S. Diebold 2010). Besides, RIG-I can detect short dsRNA from viral components. On the other hand, MDA5 can distinguish viral dsRNA from host RNA by using the length of these ligand molecules since dsRNAs that are longer than 2kb can be only present in viral genome (Kato et al. 2008).

Significant roles of RLRs on innate immune defense have been understood with the help of knockout model studies in mice. Newcastle disease virus (NDV), vesicular stomatitis virus (VSV) and measles virus (MV) have been shown to be detected by both RIG-I and MDA-5 cytoplasmic proteins and MDA-5 pathway have been found as related with virus replication in the cell (Ikegame et al. 2010). Influenza A and B (Pichlmair et al. 2006) and hepatitis C virus (HCV) also were shown to be recognized by RIG-I while most of the knockout models were lethal on the embryonic growth period (Kato et al. 2006). MDA5 dependent recognition has been shown in several different viral infections; Encephalomyocarditis virus (EMCV) and VV were some of them and the result of these experiments also proved that MDA5 recognition depends not only on the length of the RNA, but also high level RNA structures which consist of both dsRNA and ssRNA (Pichlmair et al. 2009).

Apart from TLR signaling pathways; RIG-I has been shown to stimulate NF-κB and IRF3 transcription factor activation at the time of virus recognition. It has been shown that after RIG-I and MDA-5 viral detection, downstream events are regulated through posttranslational modification (PTM) events; phosphorylation and ubiquitination (**Figure 1.2**). Binding of ligand and also ATP to the cytoplasmic receptor of the helicase and CTD leads to a conformational change in the unit and is followed by CARD domain ubiquitination by TRIM25 (tripartite motif protein 25) and E3 ubiquitin ligases, which makes this complex to bind to IPS-1 (IFN-β-promoter stimulator 1 in other words; MAVS). IPS-1 is located on the outer membrane of the mitochondria and further downstream events can be achieved through this link to the mitochondria. It has a CARD domain and after activation it associates with TRAF3 (TNF receptor associated factor-3) which recruits TANK (TRAF family member-associated NF-κB activator), NAP1(NAK-associated protein 1), SINTBAD (TBK1 adaptor)

as well as TBK1/IKKε complex, which phosphorylates IRF3 and IRF7 transcription factors and IFN-I secretion occurs (Kato et al. 2006; Takeuchi and Akira 2008; Gack 2014).

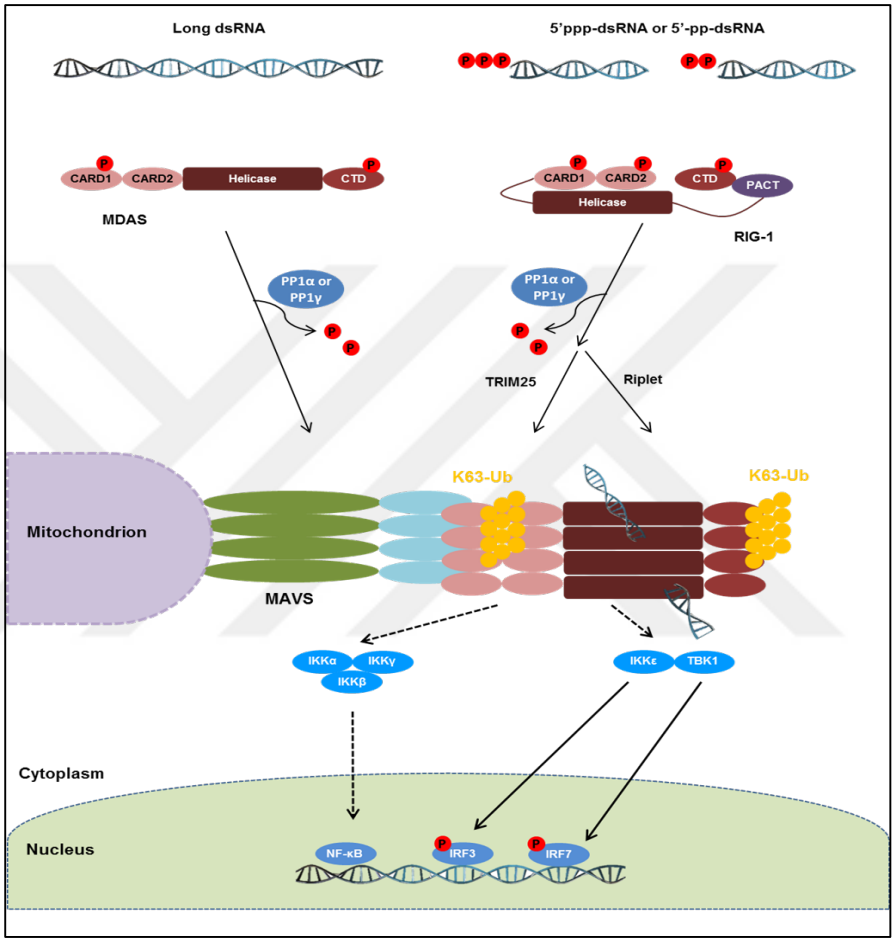

Figure 1.2 Recognition and signal transduction of RLRs After recognition of RIG-I and MDA-5 cytoplasmic PRR, downstream PTM for signal transduction events leads to IRF3, IRF7 stimulation along with NF-κB activation

1.2.3. NOD-Like Receptors (NLRs)

Nucleotide-binding oligomerization domain-containing (NOD)-like receptors or in other words; nucleotide-binding domain, leucine rich repeat-containing (NLR) proteins are mostly known due to their regulatory role in inflammatory and apoptotic responses. These

receptors are present in most cell types. Currently, 23 NLRs in humans and 34 NLR genes in mice are identified. NLRs have LRR domain on their C terminus, NOD domain at the center and mostly a CARD or a pyrin domain (PYD) at the N terminus. LRR domain serves as the sensor part which can recognize PAMPs and NOD domain is required for activation. N terminus CARD or PYD domains belong to death domain fold superfamily and this is the reason for their regulatory role in apoptosis and inflammation (Takeuchi and Akira 2010) (**Figure 1.3**).

NOD1 and NOD2 have been shown to detect some molecules which constitute peptidoglycan structures; NOD2 can detect muramyl dipeptide (MDP) while NOD1 can detect another structural component *meso*-diaminopimelic acid (*meso*-DAP). Since also TLRs can detect peptidoglycan molecules, both NLRs and TLRs work mutually in order to activate proinflammatory cytokines (Franchi et al. 2009).

NOD2 (nucleotide oligomerization domain-2) or NLRC2 (NOD-LRR (leucine-rich repeat) family with CARD 2) has been shown to serve as a cytoplasmic viral PRR and activates both IRF3 and IFN-I after Human respiratory syncytial virus (RSV) infection. This role of NOD2 is also evident in NOD2 deficient mouse models with increased viral susceptibility and pathogenesis. The role of NOD2 as a viral PRR has also been revealed both in influenza A and parainfluenza virus infections (Sabbah et al. 2009).

NLRP3 (NOD-like receptor family, pyrin domain-containing 3) forms inflammasome complexes with ASC (apoptosis-associated speck-like protein containing a CARD) and can recognize several PAMP molecules such as viral RNA, dsRNA, poly I:C and antiviral compounds such as R837 (Imiquimod) and R848 (Resiquimod) that are known to trigger antiviral activity through TLR7 and TLR8. NLRP3/ASC complexes are critical for caspase-1 activation and IL-1β production during RSV infection (Segovia et al. 2012). NLRP3 inflammasome pathway activation was also shown by influenza virus M2 protein, a proton-selective ion channel that is critical for virus action (Ichinohe, Pang, and Iwasaki 2010).

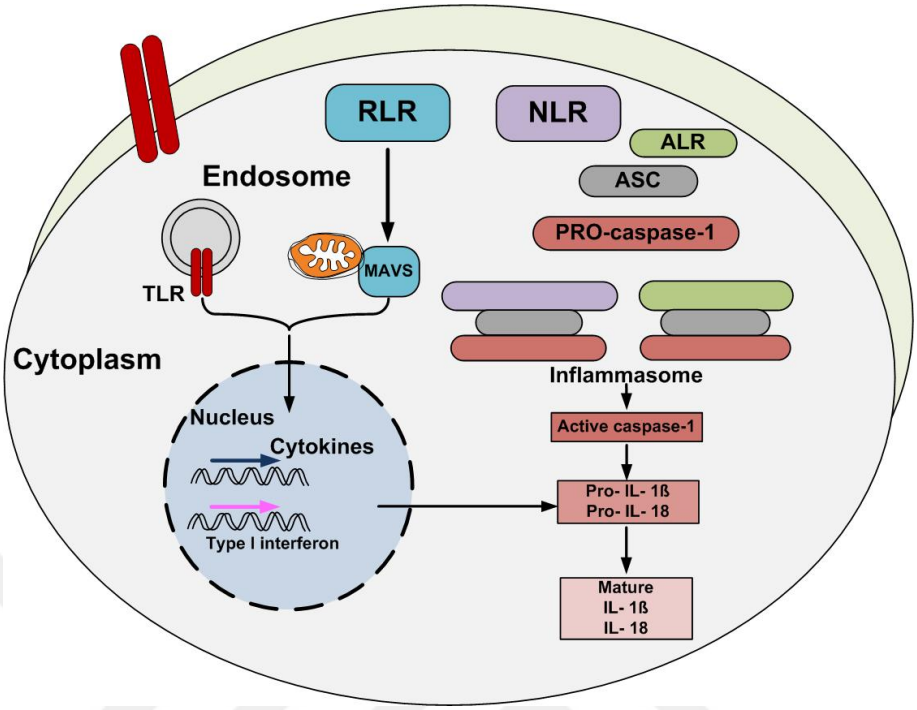

Figure 1.3 Inflamasomme complex and caspase 1 activation pathway. NLRs and ALRs are present in the cytosol. Some of virus infection result into activation of these receptors to form inflammasome with a combination of NLR, ASC, and pro-caspase-1. This complex activates Caspase-1 lead to modification in pro-IL-1β and pro-IL-18 which makes them to release in the cytoplasm (Adapted from Iwasaki, 2012).

1.2.4. C-type Lectin Receptors (CLRs)

CLRs consist of a large group of receptors which can recognize carbohydrates via their carbohydrate recognition domains. They have C type lectin like domains (CTLD), which allows them to recognize a broad range of ligands as well. Although CLRs are mostly known for recognition of glyco-compounds, they can also detect cholesterol and uric acid crystals. CLRs are one of the largest lectin family members and they are subdivided into 17 groups according to their structural and functional properties (Monteiro and Lepenies 2017).

Most of the CLRs are expressed on myeloid cells and they play a crucial role in recognition of PAMPs, damage-associated molecular patterns (DAMPs) and tumor-associated molecular

patterns (TAMPs). Since CLRs are a large receptor family with various functions, only pathogen recognition of CLRs will be summarized here.

Pathogens like bacteria, viruses, fungi and other common parasites are coated with glycans. Since viruses only replicate inside of the cell, they use host cell glycosylation pathways for their coating. Such glycosylation events are important for the stability of viral particles (Dambuza and Brown 2015). Myeloid CLRs can recognize glycolipids and glycoproteins present on both pathogens and host cells. Thus, it is an important recognition pathway while enhancing immune response to pathogen infection. Since most of the CLRs are present on APCs, they act while processing pathogen-associated antigens for presentation at the cell surface and activate adaptive immunity. CLRs on DCs and macrophages can also detect virus presence.

Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) is a transmembrane receptor which is mainly present on DCs. Several enveloped viruses; HIV-1, EBOV, HCV, dengue virus (DV), CMV and SARS-CoV have been shown to be recognized by this CLR family member (Tassaneetrithep et al. 2003). Lymph node-specific intercellular adhesion molecule-3-grabbing integrin (L-SIGN) is another transmembrane receptor of CLRs and structurally nearly identical to DC-SIGN. HIV-1, EBOV, HCV, HBV, SARS-CoV, and Marburg virus (MARV) have been shown to be recognized by this receptor (Lozach et al. 2007). Liver and lymph node sinusoidal endothelial cell C-type lectin (LSECtin) is expressed by liver, lymph node and bone marrow endothelial cells. EBOV and SARS-CoV recognition were shown for this receptor (Marzi et al. 2004). Myeloid DAP-12-associating lectin (MDL-1) is one of the transmembrane receptors of CLR family and is present on monocytes, macrophages and neutrophils. DV, Japanese encephalitis virus (JEV) and influenza viruses have been shown to be recognized by MDL-1. DCIR (DC immunoreceptor), Langerin, MMR (macrophage mannose receptor), MGL (Macrophage galactose C-type lectin) are some other CLRs that were also shown to be employed in virus recognition.

1.2.5. Cytoplasmic DNA/ RNA Sensors

Type I interferon secretion is one of the key signals for innate immunity activation via virus or any pathogenic infection. Studies of high level interferon secretion have shown that this activation is not only due to TLR dependent pathways or other receptor families mentioned above. There are also cytoplasmic sensors which facilitate immune reaction. DAI (different nomenclatures: ZBP1, DLM-1), RNA polymerase III, AIM2 and IFI16 are some examples for these sensors.

DAI (DNA-dependent activator of interferon regulatory factors(IRFs)), was one of the first identified DNA sensors in the cytosol (Takaoka et al. 2007). HSV infection, synthetic DNA and CMV (cytomegalovirus) stimulated interferon secretion have been found to be related to DAI. On the other hand, some knockdown models in mice have demonstrated that, infection with DNA viruses or synthetic DNA did not change the response of the cells which gave rise to the thought that DAI has a cell type specific response as a sensor (Rathinam and Fitzgerald 2011b; Boehme, Guerrero, and Compton 2006).

RNA polymerase III is a crucial enzyme for converting B form of AT-rich cytosolic dsDNA molecules to 5'-ppp RNA molecules that can be recognized by RIG-I to induce IFN-β secretion (Chiu, MacMillan, and Chen 2009). EBV induction of IFN-I was also related with this enzyme (Ablasser et al. 2009).

AIM 2 (absent in melanoma 2) is an interferon inducible HIN-200 family member protein and can detect dsDNA in the cytosolic compartment. Unlikely other cytoplasmic DNA sensors, AIM2 forms an inflammasomme complex. After DNA binding to its HIN200 domain, it oligomerizes and PYD domain interacts with the ASC protein to recruit procaspase 1. Critical role of AIM2 for the activation of caspase 1 and handling of IL-1β and IL-18 in APCs in mCMV and VV infection have been shown (Fernandes-Alnemri et al. 2009). Also in AIM2 deficient mice infected with mCMV, decreased serum concentrations of IL-18 and defective production of IFN-γ by NK cells was observed (Rathinam et al. 2010).

IFI16 (interferon-inducible protein 16) is a PYHIN protein family member grouped as AIM2-like receptors (ALRs) (Unterholzner et al. 2010). It contains a pyrin domain and two HIN

domains for DNA binding. IFI16 was demonstrated to recognize viral DNA, and initiate IFN-β secretion via STING (stimulator of interferon genes). STING is also an endoplasmic reticulum associated molecule that can stimulate both NF-κB and IRF3 transcription pathways to secrete IFN-I. After infection with HSV-1 virus, STING knockout models were highly lethal (Ishikawa and Barber 2008; Ishikawa, Ma, and Barber 2009).

1.3. Genetic Manipulation Techniques

Genetic manipulation is generally based on integrating a gene of interest to the host cell genome; which will lead to an expression change in the host cell for correcting or changing the wild-type genome. This is an important approach especially for the cases which the genetic background of the disease is known. With gene therapy, pathogenic mutations can be corrected via help of gene delivery techniques which can be grouped into two as non-viral gene delivery methods and viral gene delivery methods.

While viral gene delivery methods use viral vectors as agents for gene delivery, non-viral methods can be chemical and physical techniques which allow exogenous genetic material to be inserted into the cell (Mulligan 1993). Among non-viral methods, gene guns, microinjections, electroporation and lipofection can be counted. These techniques overcome the cell membrane barrier and facilitate gene delivery by either physical forces or chemical tricks (Gehl 2003; Verma et al. 2000). Although all these techniques have several advantages, the gene delivery efficiency of non-viral techniques are still not up to the standard.

Since genetic manipulation within the context of this thesis is done by using lentiviral vectors; this approach of gene delivery will be further discussed in the following section.

1.3.1. Gene Delivery by Lentiviral Vectors

Viruses are highly evolved obligatory intracellular pathogens; they can efficiently integrate to the host cell and manipulate their cellular machinery to promote replication. Lentiviruses are subgroup of retroviruses that are ssRNA viruses. After entry into the cell, their genome is reverse-transcribed and integrated to the host genome. There are several advantageous

features of lentiviruses to model as vectors that enables their use as gene therapy agents. Along with their ability of stable transgene expression, infection and integration into non-dividing cells, they have low immunogenic profile which is important to be applied as gene therapy agent. The most known lentiviruses are HIV-1 and 2, feline immunodeficiency virus (FIV) and equine infectious anemia virus (EIAV).(Naldini, Trono, and Verma 2016)

Lentiviral vectors are derived from these lentiviruses. For safety, sequences related to the replication of the virus are eliminated during vector design. This is done by discarding coding regions of the viral genome, but leaving sequences that are required for integration, and packaging of the virus intact. After deletion of these viral genes, gene of interest is inserted to the viral vector backbone. Also modified viral proteins are delivered to the cell via helper plasmids as packaging and envelope to produce an effective viral particle (Thomas, Ehrhardt, and Kay 2003)

Native forms of lentiviruses have two envelope protein groups that are covalently attached; one reside in the membrane like gp41 glycoprotein, while the other group is in the outer region like gp120 glycoprotein. In HIV-1 and these units lead to unstable characteristic of lentiviruses and effects their infectivity. However, as a therapy agent viral vectors have to be stable, thus wild type envelope genes were altered with stable G-protein of Vesicular Stomatitis Virus (VSV). Also in order to increase safety of the vectors; they were developed as self-inactivating vectors (SIN) they have deletions in the LTR (long terminal repeat) sequences that are regulatory element for viral gene expressions (Zufferey et al. 1998).

Gene therapy agents can be chosen according to intended application. sgRNA library can be inserted to cells via adenovirus, lentivirus or retroviruses. However, adenoviruses could not integrate into genome, they can be applied only for transient expression. Lentiviruses and retroviruses can be used for genome alteration since they integrate to the target genome. AAV can only enter non-dividing cells while retroviruses cannot transduce non-dividing cells. Lentiviruses can transduce both dividing and non-dividing cells and result in stable transgene expression. Because of these advantageous of lentiviruses, they are mostly employed todays gene therapy applications.(Joung et al. 2016; Naldini et al. 1998)

GeCKO v2 libraries are based on lentiviral vector and can be applied as two separate vector systems design of lentiCRISPRv2backbone; one vector system and lentiguide-PURO; dual

vector system. Both are 3rd generation lentiviral backbones (Shalem et al. 2014a). In our study one vector system was used and compatible packaging and envelope plasmids were used along with GOI plasmid (lib A and Lib B).

1.3.2. Genome Editing by Targeted Nucleases

Studying the function of genes requires to improve forward genetic methods such as inserting mutations to the target gene and observing the consequences on the phenotype. In the early days, chemical agents like ethylnitrosourea (ENU) or procarbazine which do random alterations over the genome were commonly used for such efforts. In order to investigate the function of genes or cure genetic diseases by correcting mutations in the genome, more targeted gene editing technologies have been in focus in since 1970s (Friedmann and Roblin 1972).

After DNA double-stranded breaks (DSB) were shown to effect gene function, and the finding that most of these breaks in the genome are repaired endogenously, the repair mechanisms in the genome have been thoroughly studied (Sancar 1996). Nonhomologous end-joining (NHEJ) and homology-directed repair (HDR) are two major mechanisms that are employed for repairing the DNA breaks (**Figure 1.4**). NHEJ is an error prone spontaneous action which generally results in short insertions or deletions (indels) of random nucleotides during DSB repair, while HDR mechanism is based on restoration of the break using a homologous sequence as template. Early genome editing attempts were done using this homologous recombination activity. Although this mechanism is highly specific to the target, the efficiency of the technique was quite low, as the desired modification ratio could be only 1 in 1 million or 1 billion cells (Capecchi 1989; Hsu, Lander, and Zhang 2014).

To increase the yield of these editing tools, scientists investigated other ways for editing the genome and creating these double strand breaks at the desired site, thus engineered nucleases have been developed.

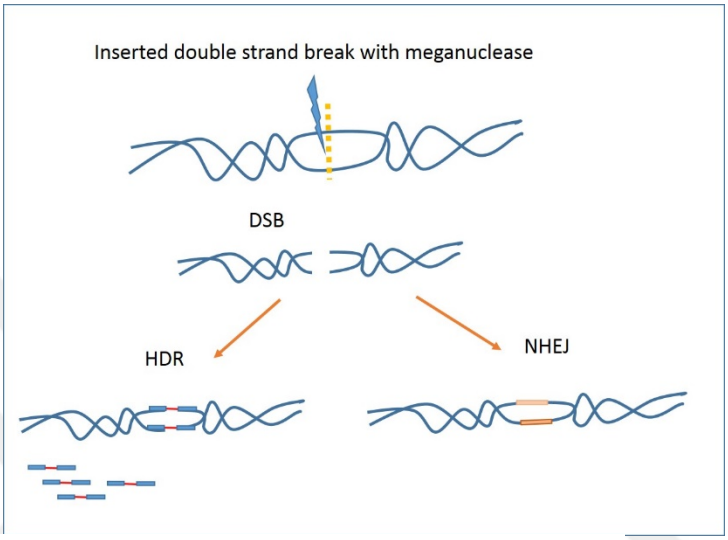

Figure 1.4 Double strand break repairing mechanisms

In 1990s, the type II restriction enzyme Fok I has been shown to have two distinct parts that work in harmony as DNA cleavage domain and DNA binding domain. By this finding the idea of changing the binding site region to alter the specificity of the enzyme was investigated (L. Li, Wu, and Chandrasegaran 1992).

Meanwhile, zinc finger transcription factors which recognize specific regions over the genome and regulate gene expression were discovered (Pavletich and Pabo 1991).They have structural repetitive units; as three identical fingers, each containing of a zinc atom, two histidine and two cysteine molecules and each unit is anchored by three nucleotides. FokI restriction enzyme DNA cleavage domain was later combined with zinc finger domain as DNA binding domain and the first engineered Zinc Finger Nuclease (ZFN) was demonstrated in 1996 (Y. G. Kim, Cha, and Chandrasegaran 1996). *Drosophila melanogaster* was the first organism whose genome was edited with ZFN in 2002 (Bibikova et al. 2002). Furthermore, ZFNs were used for viruses, bacteria, plants and mammalian cells.

ZFNs work as two monomers; each one consist of Fok1 nuclease unit anchored to a zinc finger protein and both units recognize complementary strands. There is a 5-7 bp spacer

between these two units and after dimerization of FokI nuclease units over the target DNA sequence, DSB on the target DNA occurs. Each zinc finger has a 3 bp recognition site and at least 3 zinc finger subunits are required for an active ZFN. Although dimerization from both strands increases the specificity of ZFN by requiring a longer recognition site, wild type FokI enzyme can also cleave DNA as a monomer. These results in single strand breaks on the DNA and these single nicks can result in off-target effects of ZFN editing.

Apart from ZFNs, transcription activator-like effector nucleases (TALENs), use transcription activator-like effectors (TALEs) as DNA binding/recognition domains which are derived from a plant pathogen bacteria *Xanthomonas.* Each TALE consists of tandem domain of 33-35 amino acid repeats that can recognize only one nucleotide, and specificity of each domain is achieved through 12^{th} and 13^{th} amino acids that are called repeat variable diresidues (RVDs). Each RVD recognizes only one nucleotide. Among the determined RVDs so far; guanine is recognized by Asn-Asn, adenine is recognized by Asn-Ile, cytosine is matched by His-Asp and thymine is detected by Asn-Gly. These known RVDs make design of TALENs easier than ZFNs and there are some commercial databases which help researchers to design efficient TALEN molecules. TALENs can be designed specific to any sequence on the genome even to short sequences. The only limitation of design is that there has to be a thymine residue at the 5' of the target sequence. Besides, methylated cytosine cannot be cleaved by TALENs (H. Kim and Kim 2014).

1.3.3. CRISPR History and Mechanism

The ideal genome editing tool would be one with high specificity to the target sequence, highly efficient in the desired mission and should not have any off-target effects. There have been many attempts to develop such a genome editing tool which contains all these properties only in one technique. Therefore "RNA guided endonucleases" in other words CRISPR (clustered, regularly interspaced, short palindromic repeat)/Cas9 gave hope to researchers in this area.

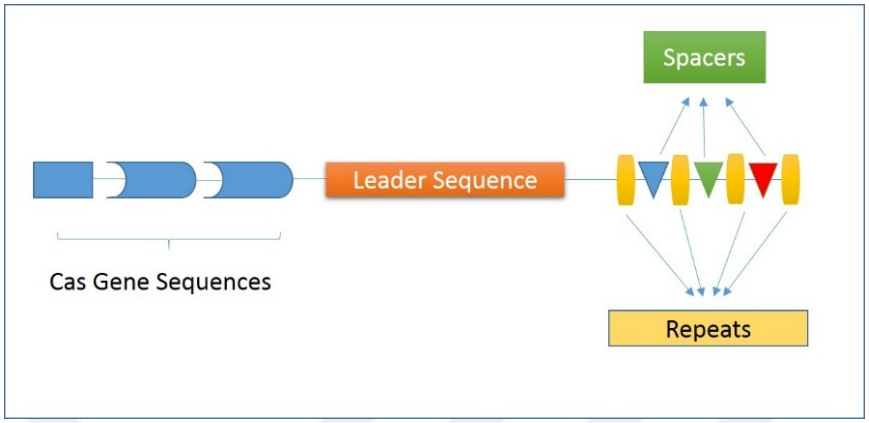

Figure 1.5 CRISPR/Cas locus in bacteria and archaea. They have adaptive like immunity; cleaved virus genome short sequences are added as spacers between repetitive units in bacteria and archaea genome, this provides cells to escape from next invasions of the same pathogens

CRISPR/Cas system is adapted from bacterial and archaeal immune system that is an acquired adaptive immune ability to fight against viruses and phages. crRNA (CRISPR-RNA) is used for DNA recognition and Cas (CRISPR-associated protein) enzyme is used for DNA cleavage. 90% of the archaeal genome constitutes of the CRISPR/Cas locus while bacterial genome has 40%. In the CRISPR locus there are repeat sequences which are interrupted by non-repeated short sequences called spacers (**Figure 1.5**). The genome of the virus or phage that invades the host cell is cleaved by Cas endonuclease and parts of this cleaved foreign DNA is integrated into the CRISPR locus of the host genome as spacer. These spacer sequences are used for later crRNA synthesis when the same virus or phage enters the cell again (Makarova et al. 2011).

As shown in previous studies there are more than 40 Cas nuclease proteins which play a crucial role in this cascade of events in bacteria and archaea. CRISPR systems are divided into three main categories due to the differences in the utilized proteins. In CRISPR II, Cas9 endonuclease is employed. Cas9 has one HNH nuclease domain and RuvC- like nuclease domain and each nuclease cleaves one opposite strand after recognition by crRNA and trans-

activating CRISPR RNA (tracrRNA) complex. These two RNA molecules should make a complex with Cas9 to become an efficient, targeted cleavage tool.

Since CRISPR II is one of the simplest ones in terms of design, this system is mostly used as a genome editing tool (F. Zhang, Wen, and Guo 2014). In contrast to the endogenous system, single chain guide RNA (sgRNA) can be designed incorporating both crRNA and tracRNA, as a single chimeric RNA molecule and utilized in CRISPR/Cas9 genome editing system (**Figure 1.6**). Once the sgRNA complexes with the Cas9 protein to guide the specificity of the enzyme, Cas9 induces DSB on the DNA and in the absence of any homologous template in the site, NHEJ occurs. Generally, NHEJ leads to random indels over the cut site and results in frameshift mutations that silence the gene.

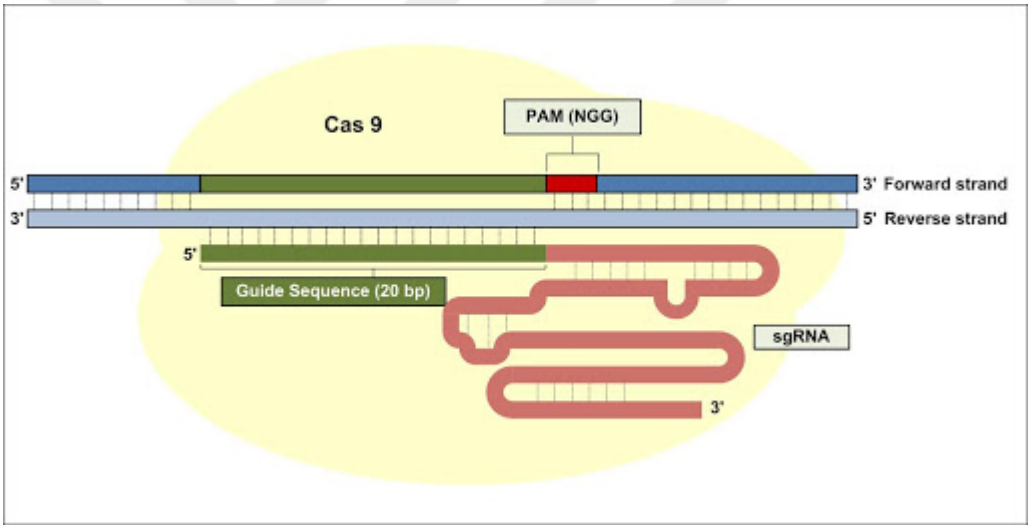

Figure 1.6 CRISPR/Cas9 representation.(Adapted from Kim & Kim, 2014)

Instead of designing a meganuclease each time for each unique target as in ZFNs and TALENs, the CRISPR/Cas9 system requires only a well-designed sgRNA. The only limitation about the target in the CRISPR/Cas9 system is that Cas9 requires a "protospacer-adjacent motif'' (PAM), a 3 bp sequence that must follow the target sequence. 5′-NGG-3′ and 5′-NAG-3' are the most frequent PAM sequences. The PAM sequence depends on which species of microorganism is used for derivation of the Cas9 gene. Requirement of PAM sequence in the target is a limiting factor as a genome editing tool. On the other hand, CRISPR/Cas9 can cleave methylated targets, which cannot be edited by ZFNs or TALENs

(Jinek et al. 2012). Easiness and efficiency of CRISPR systems help scientists to test forward genetic engineering hypotheses in large scale.

1.4. Genome-wide Screening and GeCKO Library Approach

When the use of genome editing tools became convenient, the idea of editing each cell using a library that contains different target sequences and visualizing them as a pool was put forward as a promising idea for screening the whole genomic picture behind complex phenotypes.

Forward genetic screening approaches employ a selective phenotype and only cells which abide by this phenotype are investigated genomically. One of first attempts to genomic screening was achieved through chemical mutagens which create random mutations over the genome and they were successful to identify the functions of several genes in Ras and Notch Signaling(Sundaram 2005). However, since mutations are random, it was also a taunting task to find the corresponding alterations in the genome after the cells with the desired phenotype were sorted out.

RNA interference (RNAi) is another technique which has made great contributions to genetic screening efforts. Small RNAs are 20-30 nucleotide long, noncoding double-stranded RNAs which make a ribonucleoprotein complex with Argonaute protein family members. In brief, this complex and additional proteins make up RNA-induced silencing complex (RISC). Sequence-specific interaction of RISC with mRNA leads to complete knockdown of the gene via mRNA degradation. RNAi was discovered in 1998 as microRNAs (miRNAs) in the *Caenorhabditis elegans* genome and their role in regulating gene expression in the cell was demonstrated (Fire et al. 1998). In contrast to genome derived miRNAs, short interfering RNAs (siRNA) can be endogenous or exogenously derived during viral infections. Due to their crucial regulatory role in cells, siRNAs are responsible for several important events such as growth, differentiation and reproduction in the cell(Wilson and Doudna 2013).In contrast to genomic screens which are based on random mutations, RNAi is favored because the target sequences are known and fatal mutations can be decided without any recovery process.

However, because of incomplete silencing and off-target effects, it has also some unfavorable characteristics (Boutros and Ahringer 2008).

Target specific CRISPR systems offer an alternative to genomic screening by its effective editing technique over the genome. The general workflow for the genome wide screening includes designing target specific sgRNAs which are shown computationally to recognize only one specific target sequence and oligos are produced according to this design. Generally, more than one sgRNA can be designed to the same gene in order to be sure of complete knockdown and avoid false positives coming from off-target effects. The design of oligos also has universal sequences at their ends to be amplified by known primers. Then, this oligo sequence is cloned into a vector which also has a cassette for Cas9 expression. After gene transfer, target cells can produce both Cas9 and sgRNA, eventually CRISPR/Cas9 can work.

In such screening attempts, knocking down only one gene per cell is important. During the delivery of CRISPR/Cas9 constructs into the cell, viral vector titer is kept at low multiplicity of infection (MOI) for this purpose. Since every vector sequence has its unique 20bp spacer sequence, when it is amplified by adapters that are known from vector design, it is possible to identify which sgRNA is inside the cell. Ensuring each cell receives one sgRNA, negative or positive screening can be applied to the cells. After selection, sequencing results are used to calculate which sgRNA is amplified or depleted within the pool (Sanjana 2016).

In 2014, two groups published their work on genome wide CRISPR loss-of-function screening in Science(Tim Wang et al. 2014; Shalem et al. 2014b). Feng Zhang group introduced GeCKO (genome-scale CRISPR-Cas9 knockout) libraries with this publication. Then a version update and improvements like the lentiviral packaging and choice of guide sequences was published by Sanjana *et al* (Sanjana, Shalem, and Zhang 2014).

This updated GeCKO library has more than 19,050 target genes for the human and 20,600 target genes for the mouse genome. Besides, both libraries are grouped into two as Library A and Library B where each library has 3 unique sgRNA sequences targeting one gene in the human or mouse genome, thus Lib A+ Lib B mix will have 6 sgRNA sequences per gene. Furthermore, they contain sgRNA sequences targeting miRNA genes while there are also

1000 non-targeting control sgRNAs. More detailed description of the library is given in **Table 1** (Sanjana, Shalem, and Zhang 2014; Shalem et al. 2014b).

	GeCKO v1 human library	GeCKO v2 human library	GeCKO v2 mouse library
Targeting gene number	7,114	19,050	20,611
Targeting sgRNA number	10 per gene	6 per gene3 in Library A,3 in Library B)	6 per gene(3 in Library A, 3 in Library B)
Targeting miRNA number	None	1,864	1,175
Non- targeting sgRNA number	100	1000	1000
Total sgRNA sequences	73,151	123,411 (65,383 in Library A, 58,028 in Library B)	130,209 (67,405 in Library A, 62,804 in Library B)

Table 1-1 GeCKO library content (adapted from (Sanjana, Shalem, and Zhang 2014))

2. AIMS OF THIS STUDY

The use of genetically modified cells for therapeutic purposes is an increasing trend that shows great promise. The major hurdle in genetic modification of human cells is the delivery of the gene-of-interest into the cell. Currently, viral vectors are most commonly used for *ex vivo* genetic modification of human cells but there's very little information about the intracellular immune response pathways triggered by viral vector entry.

Although the literature is scarce regarding the activation of intracellular immune responses against viral vectors, it has been shown that an innate immune response against the vector can be generated by pDCs (Rossetti et al. 2011). Such responses against lentiviral gene delivery have also been documented during *in vivo* studies after systemic administration of the vector, resulting type I interferon responses and vector clearance (Brown et al. 2006).

While anti-viral responses have been studied thoroughly in wild-type virus infections, it has been mostly overlooked from a gene therapy point-of-view. It is possible that TLR or RLR mediated detection of viral vector components might activate an anti-viral response, negatively affecting the efficiency of viral gene delivery. Development of novel gene transfer protocols based on inhibition of intracellular antiviral responses opens up the possibility to enlarge the base of cell types that can efficiently be used in gene therapy protocols.

Because of the innate resistance of NK cells against viral infections, they are a good model to get detailed insights about cellular response mechanisms against viral vectors. This study aims to use NK92 cells as a model in order to map the pathways that are actively taking part in the response against lentiviral gene delivery. The identification of important pathways and genes taking part in anti-vector response is expected to help overcome the intracellular defense mechanisms against gene delivery vectors. This may have broad technical

applications in order to improve the efficiency of genetic modification of a wide variety of cell types.

3. MATERIAL AND METHODS

3.1. Materials

3.1.1. Chemicals

Chemicals and Media Components	Company
Absolute Ethanol	Sigma Aldrich, Germany
Agarose	Sigma Aldrich, Germany
Ampicillin Sodium Salt	CellGro, USA
BX795	Sigma, Germany
Chloroquine	Sigma, Germany
Distilled Water	Ambion, USA and Merck Millipore, USA
DMEM	Thermo Scientific, USA
DMSO	Sigma, Germany
DNA Gel Loading Dye 6X	Fermentas, USA
Ethidium Bromide	Sigma Aldrich, Germany
Fetal Bovine Serum	Thermo Scientific, USA
HEPES 1M Solution	Sigma, Germany
Interleukin-2	Proleukin, Novartis
Isopropanol	Sigma Aldrich, Germany
LB Agar	Sigma Aldrich, Germany

L-glutamine, 200 mM	Sigma, Germany
MEM Non-essential Amino Acid Solution	Sigma, Germany
Poly-D-Lysine	BD, USA
Puromycin	Thermo Scientific, USA
SCGM Stem Cell Growth Medium	Mediatech, USA
Sodium Hydroxide 10N	Sigma Aldrich, Germany
Sodium Pyruvate 100mM	Sigma, Germany
Tris Acetate- EDTA Buffer 10X	Sigma Aldrich, Germany

3.1.2. Equipment

Equipment	**Company**
Balance	Sartorius CPA 1003S, Germany
Bioanalyzer 2100	Agilent, USA
Centrifuge	Eppendorf, 5415D, Germany Eppendorf, 5702, Germany
Deepfreeze	-80°C Innova U725, Eppendorf, Germany , '-20°C Kirsch, Germany
E-Gel iBase Power System and E-Gel Safe Imager Transilluminator	Thermofisher Scientific, USA
Electrophoresis Apparatus	Biorad Inc., USA
Filter 0.45µm	Merck Millipore, USA
Flow cytometer	BD FACScanto, USA
Gel Documentation Systems	Gel Doc™ XR+ System, Biorad Inc., USA
Heater	Thermomixer Comfort, Eppendorf, Germany
Ice Machine	Scotsman Inc., AF20, USA

Laminar Flow	MSC 1.2, Thermo Scientific, Germany
Microliter Pipettes	Eppedorf, Germany
Microwave Oven	Bosch,Turkey
Refrigerator	Arçelik, Turkey
Sequencing Systems	Illumina HiSeq2500, USA
Shaker Incubator	New Brunswick Sci., Innova 4330, USA
TProfessional Basic Gradient 96 Thermocycler	Biometra, Germany
UV spectrophotometer	NanoDrop 2000, Thermofisher Scientific, USA
ViiA 7 Real-Time PCR System	Applied Biosystems, USA
Vortex	VELP Scientifica, Italy

Table 3-1 Equipment used in this study

3.1.3. Buffers and Solutions

Agarose Gel: For 60 ml 2% w/v gel, 1,2 g of agarose powder was dissolved in 60 ml 1X TAE buffer by heating. 0.01% (v/v) ethidium bromide was added to the solution.
Phosphate-buffered saline (PBS): For 1000 ml 1X solution, 100 ml 10X DPBS was mixed with 900 ml ddH2O.

3.1.4. Growth Media

LB Agar: Powders of 20 g LB and 15 g bacterial agar were dissolved in 1L double disstilled water for 1X LB agar solution, then it is autoclaved for 15 min in 121°C. Before pouring into sterile petri dishes, ampicillin is added to the medium as final concentration will be 100ug for 1 ml.

DMEM: 293FT cells are cultured in DMEM solution that is supplied with 10% heat inactivated fetal bovine serum, 1mM Sodium Pyruvate, 25 mM HEPES solution, 0.1mM MEM Non-essential amino acid solution and 2mM L-Glutamine.

SCGM: CellGro SCGM that were also added 20%FBS is used for culturing NK92 cells; IL-2 is provided in every two days.
Freezing Medium: Cell lines were frozen in FBS in the presence of 6% DMSO.

3.1.5. Commercial Kits

Commercial Kit	Company
Calcium Phosphate Transfection Kit	Sigma-Aldrich, USA
E-Gel Size Select Agarose Gels	Thermofisher Scientific, USA
High Sensitivity DNA Kit	Agilent, USA
Illumina PhiX Control v3	Illumina, USA
Illumina Rapid Cluster Kit	Illumina, USA
Illumina SBS Kit	Illumina, USA
KAPA library quantification kit	KAPA, USA
NuceloSpin Plasmid Midiprep Kit	Macherey-Nagel, USA
NuceloSpin Plasmid Miniprep Kit	Macherey-Nagel, USA
PureLink® HiPure Plasmid Midiprep Kit	Invitrogen, USA

Table 3-2 Commercial kits used in the study

3.1.6. Enzymes

All the restriction enzymes, polymerases and PCR reaction supplements were Fermentas, Kapa or New England Biolabs.

3.1.7. Bacterial Strains

Electrocompetent *Escherichia coli (E.coli)* DH-5α strain is used for plasmid transformation and expansion.

3.1.8. Mammalian Cell Lines

HEK293FT: Human embryonic kidney 293 cell line which contains SV40, a large protein of T antigen, that increases the expressing rate of protein carried by the vector (Invitrogen R70007).

NK-92: Human natural killer cell line isolated in the year 1992 from a non-Hodgkin's lymphoma patient (ATCC® CRL 2407™).

3.1.9. Plasmids and Oligonucleotides

	PLASMID NAME	**PURPOSE OF USE**	**SOURCE**
1	Lego-G2	Lentiviral construct for GFP expression	Prof. Boris, Fehse of University Medical, Center Hamburg-Eppendorf, Hamburg, Germany
2	lentiCRISPRv2 -GeCKO Human Library A	Lentiviral construct for CRISPR/Cas9 expression with Puromycin resistance gene – GeCKO Human knock out Library A	Addgene (#1000000048)
3	lentiCRISPRv2 -GeCKO Human Library B	Lentiviral construct for CRISPR/Cas9 expression with Puromycin resistance gene – GeCKO Human knock out Library B	Addgene (#1000000048)
4	pCMV-VSV-g	Virus production/packaging plasmid(Env)	Addgene(#8454)

5	pMDLg/pRRE	Virus production/packaging plasmid (gag/pol)	Addgene(#12251)
6	pRSV-PREV	Virus production/packaging plasmid(Rev)	Addgene(#12253)

Table 3-3 Plasmid List

	OLIGO NAME	**OLIGO NUCLEOTIDE SEQUENCE**	**SOURCE**
1	v2 Adaptor Forward	AATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCG	PCR1 of GeCKO Libraries
2	v2 Adaptor Reverse	TCTACTATTCTTTCCCCTGCACTGTtgtggcgatgtgcgctctg	PCR1 of GeCKO Libraries
3	F01	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTtAAGTAGAGtcttgtggaaaggacgaaacaccg	Illumina Sequencing indexing PCR2
4	F02	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTatACACGATCtcttgtggaaaggacgaaacaccg	Illumina Sequencing indexing PCR2
5	R02	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTtTCTACTATTCTTTCCCCTGCACTGT	Illumina Sequencing indexing PCR2
6	F03	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTgatCGCGCGGTtcttgtggaaaggacgaaacaccg	Illumina Sequencing indexing PCR2

7	F04	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTcgatCATGATCGtcttgtggaaaggacgaaacaccg	Illumina Sequencing and indexing PCR2 primers
8	F05	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTtcgatCGTTACCAtcttgtggaaaggacgaaacaccg	Illumina Sequencing and indexing PCR2 primers
9	F06	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTatcgatTCCTTGGTtcttgtggaaaggacgaaacaccg	Illumina Sequencing and indexing PCR2 primers
10	F07	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTgatcgatAACGCATTtcttgtggaaaggacgaaacaccg	Illumina Sequencing and indexing PCR2 primers
11	F08	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTcgatcgatACAGGTATtcttgtggaaaggacgaaacaccg	Illumina Sequencing and indexing PCR2 primers
12	R03	CAAGCAGAAGACGGCATACGAGATCGCGCGGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTgatTCTACTATTCTTTCCCCTGCACTGT	Illumina Sequencing and indexing PCR2 primers
13	R04	CAAGCAGAAGACGGCATACGAGATCATGATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTcgatTCTACTATTCTTTCCCCTGCACTGT	Illumina Sequencing and indexing PCR2 primers
14	R05	CAAGCAGAAGACGGCATACGAGATCGTTACCAGTGACTGGAGTTCAGACGT	Illumina Sequencing and indexing PCR2 primers

		GTGCTCTTCCGATCTtcgatTCTACTATT CTTTCCCCTGCACTGT	
15	R06	CAAGCAGAAGACGGCATACGAGATTC CTTGGTGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTatcgatTCTACTATTC TTTCCCCTGCACTGT	Illumina Sequencing and indexing PCR2 primers
16	R07	CAAGCAGAAGACGGCATACGAGATA ACGCATTGTGACTGGAGTTCAGACGT GTGCTCTTCCGATCTgatcgatTCTACTAT TCTTTCCCCTGCACTGT	Illumina Sequencing and indexing PCR2 primers
17	R08	CAAGCAGAAGACGGCATACGAGATA CAGGTATGTGACTGGAGTTCAGACGT GTGCTCTTCCGATCTcgatcgatTCTACTA TTCTTTCCCCTGCACTGT	Illumina Sequencing and indexing PCR2 primers
18	KAPA Forward primer	5'-AAT GAT ACG GCG ACC ACC GA-3'	KAPA Quantification
19	KAPA Reverse primer	5'-CAA GCA GAA GAC GGC ATA CGA-3'	KAPA Quantification

Table 3-4 Oligonucleotides list

3.1.10. DNA Ladders

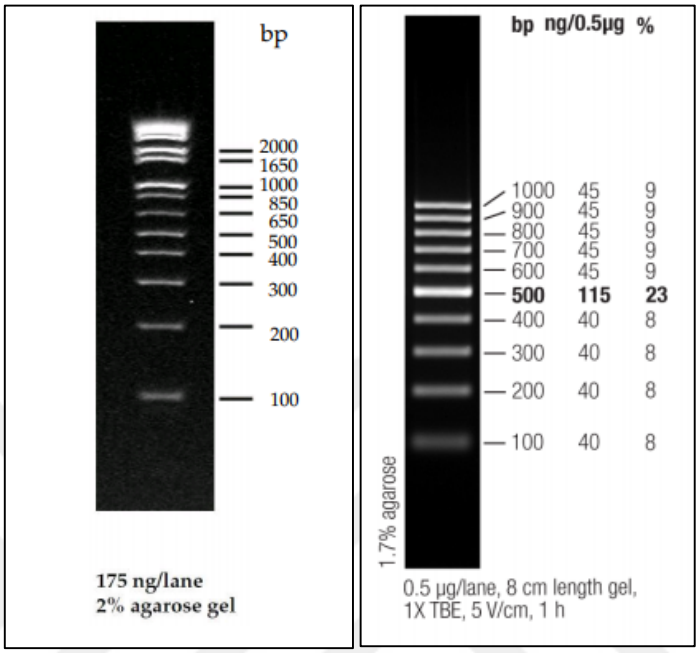

Figure 3.1 DNA ladders: E-Gel® 1 Kb Plus DNALadder on E-Gel® EX Agarose Gels and Thermo Scientific™ GeneRuler™ 100 bp DNA Ladder, respectively.

3.1.11. DNA Sequencing

All DNA Sequencing processes were done by me in laboratory of IGBAM (Advanced Genomics and Bioinformatics Research Center), in TUBITAK MAM , Kocaeli, Turkey (have been a member since 2013)

3.1.12. Software, Computer-Based Programs and Websites

SOFTWARE, PROGRAM, WEBSITES	**COMPANY/ADRESS**	**PURPOSE OF USE**
ABI ViiA7 Software	ABI, USA	Analyzing qPCR results

Addgene	https://www.addgene.org/	GeCKO library preparation, design and sequence information.
CASAVA 1.8.4	Illumina, USA	bcl to fastq conversion
FastQC	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/	Library and Sequencing Quality Check
FlowJo v10	Tree Star Inc., USA	Viewing and analyzing flow cytometer data
g:profiler	http://biit.cs.ut.ee/gprofiler/	Pathway analysis
GeCKO v.2 Library	http://www.genomeengineering.org/gecko/	GeCKO library preparation, design and sequence information.
Illumina Sequencing and Control Softwares	Illumina, USA	Control of Sequencing parameter and outcomes
MAGeCK	https://sourceforge.net/p/mageck/wiki/Home/	Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout
Python 2.7	https://www.python.org/	Trimming of adapters and MAGeCK tool analysis

Table 3-5 Software websites and programs

3.2. Methods

3.2.1. Mammalian Cell Culture

Maintenance of cell lines:

HEK293FT cells are cultured in DMEM as described in **3.1.4 Growth Media** part, they were incubated at 37°C with 5% CO2. Splitting is done when 90% confluency is reached. While splitting, cells were washed with PBS and trypsinized for 5 mins after which they were diluted in DMEM 1:3 to 1:8 ratio. Split procedure is repeated every two days, preventing cells to reach 100% confluency.

NK-92 cells were maintained in sterile tissue culture flasks using CellGro SCGM with 20% heat-inactivated fetal bovine serum (FBS) and 1000 U/ml human Interleukin-2 (IL-2) at 37°C incubator with 5% CO_2. Cell concentration was arranged as between 300,000 cells/ml to 1,000,000 cells/ml and IL-2 was freshly supplemented every 48 hours.

Cell Freezing method: Both cells (HEK293 and NK92) were split one day before freezing to 500,000 cells/ml for NK92 and to a confluency of 30-40% for HEK293FT. The day after, cells were counted and 3 million cells at least is frozen per vial. Cells where centrifuged at 300g for 5 min for adherent cells and 200g for 5 min for suspension cells. Supernatants were thrown after centrifugation and all cells were resuspended in 0,5ml FBS and incubated on ice for 15-20 min. Meanwhile, 0.5 ml FBS with 12% DMSO was freshly prepared and put on ice as well. At the end of the incubation, 500µl cells is mixed with the prepared solution of FBS and DMSO.

Final concentration DMSO becomes %6 in 1 ml mix. Afterwards, the vials were put in -80 for short-term storage and transferred to liquid nitrogen for long term storage.

Cell thawing method: Vials that are kept in liquid nitrogen were taken out and put on ice immediately. As soon as the vials thawed 5 ml FBS was added carefully in 2-3 min, being slow was important for this step. Cells were centrifuged at 300 g for 5 min supernatant was discarded. Then cells were resuspended in their complete media (DMEM for HEK, SCGM for NK) as being 500,000-700,000 cells/ml and plated.

3.2.2. Production of GeCKO-CRISPR Lentiviral Vectors

GeCKO v.2 library plasmid preparations: GeCKO v.2 library has two sub-libraries as library A and Library B, thus following protocols for transformation, expansion and plasmid prep were applied for both libraries in the same way according to the manufacturer's protocol.

Transformation of GeCKO libraries: After dilution of each of the GeCKO libraries (Lib A and Lib B separately) to 50 ng/µl in TE buffer, 2 µl of diluted libraries were added on Invitrogen® ElectroMAX DH5α cells and electroporation is done according to the manufacturer's protocol. This process is repeated 4 times for each library and at the end they were recovered with the recovery media. All the collected and recovered cells were pooled and mixed in a total of 8ml and plated on large scale pre-warmed bacteria bioassay plates (24,5 cm^2) that contain ampicillin. 4 ml cultures were dissipated onto the two agar plates per sub-library and incubated for 14 hour at 32°C. After incubation, 5-10 ml liquid LB is added onto the bioassay plates and with help of cell scraper, colonies were collected. 35 ml bacterial suspension is collected and after centrifugation at 3000g for 30 min, bacteria is suspended and continued with maxiprep for plasmid purification following manufacturer's protocol.

GeCKO CRISPR lentiviral vector production: 14×10^6 HEK293FT cells were seeded in poly-L-lysine-coated 15-cm cell culture plates overnight in DMEM, 10% FBS. Next morning, cells were transfected with 30µl gene of interest (GOI) plasmid

(**Figure 3.2**) (library A or library B) along with 15ug pMDLg/pRRE (Gag/Pol), 10μg pRSV-Rev (Rev) and 5μg phCMV-VSV-G (Env) using Calcium Phosphate Transfection kit in the presence of 25 μM Chloroquine. After 8-10 h incubation, medium of transfection plates are replaced with complete growth media. Then, virus containing supernatant is collected at 24h and 36h time points and filtered with 0,45 μm filters. All virus containing supernatants were stored at -80°C.

3.2.3. Transduction Workflow for GeCKO v.2 Library

Virus titration was determined through transduction of cells with a serial dilution of viral supernatant for 6 hours. 24 hours after transduction, cells were split into two and were cultured in the presence or absence of 1μg/ml Puromycin. Cell counts were monitored every two days. When the count of viable cells with puromycin^{+} wells is divided by the cell counts of cells grown without puromycin, the percentage of cells that initially received the vector were calculated. Based on this calculation and in parallel to the protocol paper (Sanjana, Shalem, and Zhang 2014), expected transduction efficiency was determined as 20% in order to be sure that only one sgRNA will enter one cell during transduction. General transduction workflow can be seen in **Figure 3.2** below.

Transduction of NK-92 cells with GeCKO v.2 Library A or Library B: 10×10^{6} NK92 cells were transduced in the presence of 3 μM BX795, with GeCKO v.2 Library A or Library B lentiviral virus for 6 hours. 24hr post-transduction, puromycin selection was initiated and cell count was recorded every two days. After the selection process of two weeks, cells reached the expected count as 50×10^{6} cells, which is also the required amount for 300X cell representation of the library. Genomic DNA is isolated from cells and measured via Nanodrop Spectrophotometer. Nested PCR (PCR1 and PCR2) is applied to both libraries.

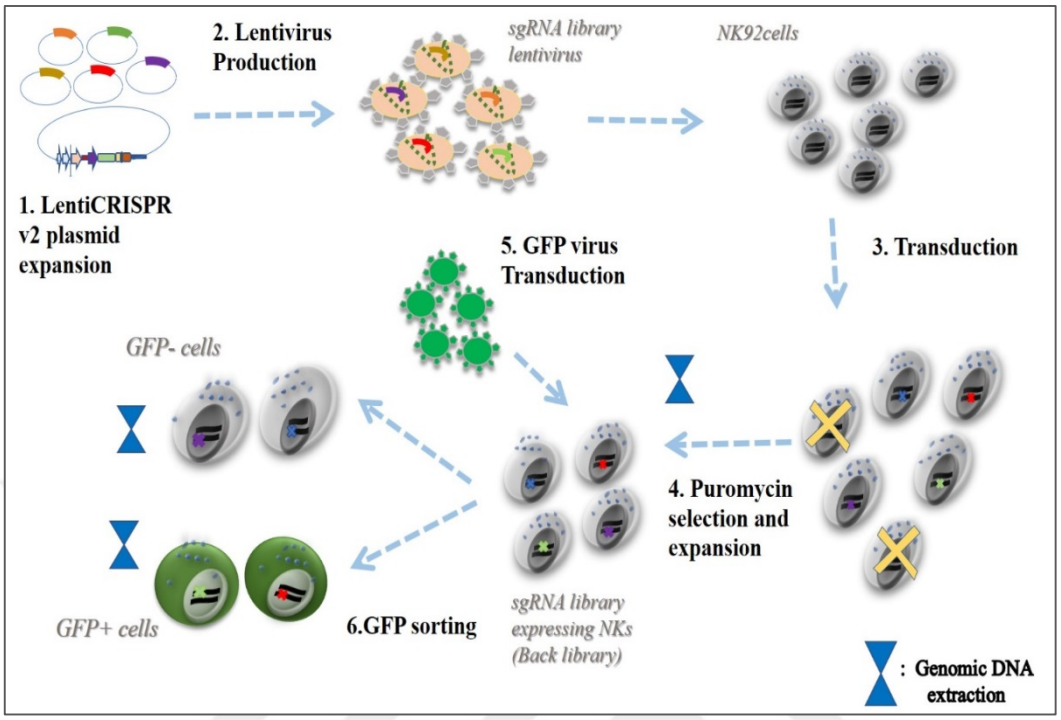

Figure 3.2 GeCKO v.2 lentiviral transduction workflow

Transduction of GeCKO^{+} NK92 cells with GFP expressing vector: 10×10^{6} of GeCKO lentiCRISPR transduced cells for each library were transduced again with LeGO-G2 GFP expressing vector for 6 hour at MOI=10 in the absence of inhibitor (BX795). 3 days after transduction, cells were analyzed on FACS and both GFP positive and negative cells were sorted. Then these cell populations were expanded and when reached at least 40 million cells (approximately 14 days) their genomic DNA was isolated. See timeline for all GeCKO library transduction events in **Figure 4.1**

3.2.4. sgRNA Sequencing Library Preparations

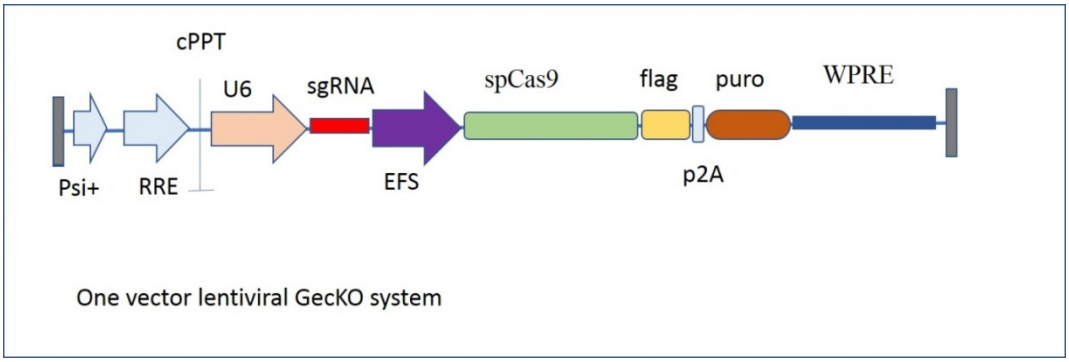

Figure 3.3 Lentiviral v.2 backbone from Feng Zhang Lab (taken from addgene.org)

Nested PCR for NGS: Lentiviral vectors integrate randomly into the genome almost typically into active gene loci. (Biffi et al. 2011). In our experiments, we adjusted transduction efficiencies such that at most only one virus would enter one cell. To reveal the identity of the sgRNA sequence incorporated into individual cells, we performed nested PCR from genomic DNA. The identity of sgRNAs indicate the identity of the genes that are silenced because sgRNAs together with the Cas9 enzyme expressed by the lentivirus are presumed to mutate and silence specific genes with sequences homologous to that of the sgRNA. Nested PCR protocol was given in **Figure 3.4** in detail.

The sequence of the integrated lenti-CRISPR vector was already known **(Figure 3.3)**. PCR1 primers called v2 adaptors; amplified the vector from an outer region of the sgRNA sequence; from U6 promoter to a part of sgRNA frame. PCR2 were mainly done to add Illumina universal adaptors to the first PCR1 products which is used for flowcell binding. In the design of PCR2 primers there was also Illumina sequencing primers, different indexes for per library barcoding and a variable stagger sequence for increasing diversity of libraries. Detailed design scheme for PCR2 primers can be seen in **Figure 3.5**.

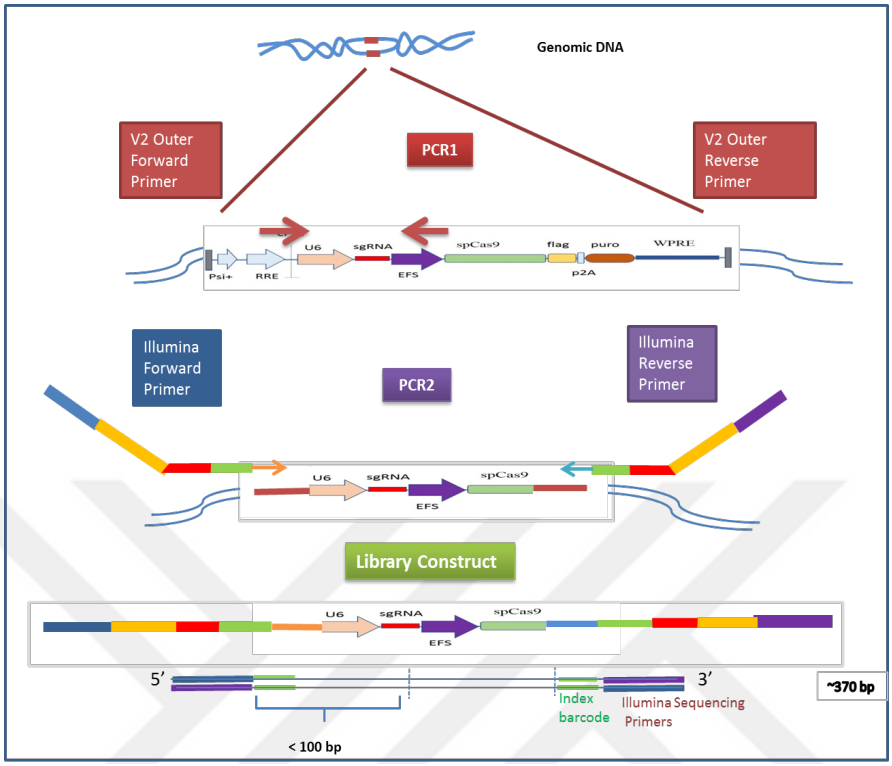

Figure 3.4 Nested PCR for GeCKO library sgRNA sequencing library. PCR1 primers targets for relatively small sequence from the integrated vector; from U6 promoter to a part of sgRNA sequence which includes target sequence, PCR2 primers (detailed design given in Fig.3.4) uses PCR1 adapters as priming site and adds Illumina sequencing adapter. At the end library length is around ~370bp

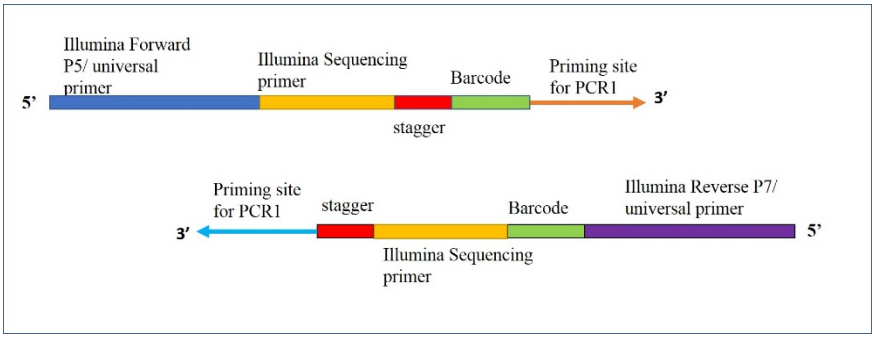

Figure 3.5 Design scheme for PCR2 forward and reverse primers

E- Gel Size selection: Gels are commercially available agarose cassettes, they are easy to handle and there are a series of cassettes that can be selected for different

purposes. E-Gel Size Select %2 Agarose Gels are specialized for precise size selection which are mostly developed for NGS applications. Also, concentration of gel is chosen according to target band size and resolution in these experiments. E-Gel Cassettes are prepared ethidium bromide free gels and allow nucleic acids to be dyed with “proprietary” florescent dye stain while running on the E-Gel® iBase™ Power System. Furthermore, thanks to E- Gel® Safe imager, real time running of the template is visualized while running. Loading, running and collecting the desired band is done according to manufacturer’s protocol.

Before inserting the iBase, package of agarose gel cassette is opened carefully and covers of wells (top wells and bottom wells) are removed, excess buffer from the well is cleaned with KimWipe. Bottom wells are filled with 25 ul nuclease free water, while the small well in the middle filled with 15ul. Templates are loaded to the top wells; any empty wells are filled with 20ul nuclease free water. M signed ladder lane, in the middle of top wells is filled with 5-10ul E-Gel ladder. Proper protocol for the gel concentration and purpose is picked up from the menu and with “go” command run is started. Corresponding lanes will run at the same speed with the ladder. Thus, when the desired band size reaches the bottom well, the run is stopped and the liquid inside the well and gel pieces are collected into an Eppendorf tube. In order to prevent cross contamination different samples are recommended to load onto different gels. Detailed photo can be seen in **Figure 3.6.**

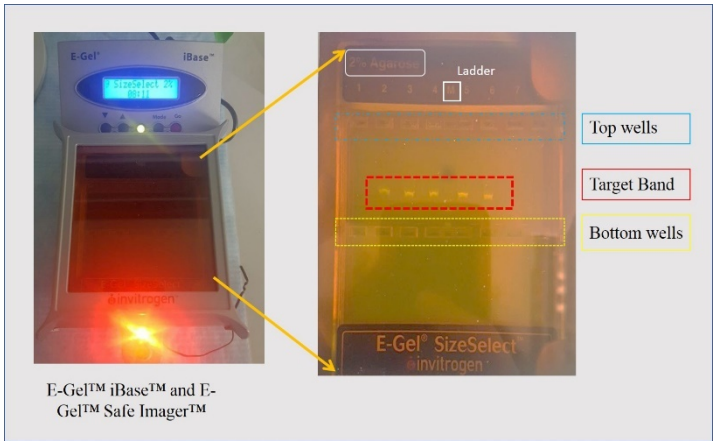

Figure 3.6 E-Gel size selection protocol

Gel Purification: Next generation sequencing machines require really pure templates, otherwise, amplification steps may be inhibited and required coverage amount cannot be reached. In order to guarantee the purity of the PCR products, gel extraction protocol is applied to the samples which were size selected via E-Gel Size Select protocol. For gel extraction, column based Min Elute Gel Extraction Kit is applied according to manufacturer's protocol and to quantify DNA amount all samples were measured with Nanodrop.

Bioanalyzer Quality/Quantification Check (QT/QC check): Bioanalyzer is a quite sensitive device to evaluate size, purity and quantity of DNA, RNA and protein preparations. The idea of Bioanalyzer is doing automated gel electrophoresis in microvolumes for micro amounts. It has several ready-to-use assays for different size amounts and one of them can be chosen accordingly your template concentration. For instance, since the amount of our template was in low amounts, High Sensitivity DNA kit was chosen, are several kits, can be used through the purpose, in our experiment we used High Sensitivity DNA kit, in order to check length of our library and quantitate the amount. High Sensitivity DNA kit was applied for every library which will be loaded for sequencing.

Library Quantitation via Realtime PCR: Cluster formation is a crucial process for obtaining enough intensity and also deepness of libraries on HiSeq. There are several parameters which effect cluster formation and should be carefully assessed; like purity, length and concentration of the template. Thus, every library should be quantified carefully. One of the most sensitive ways, real-time-PCR is used for this purpose. KAPA library quantification kit for Illumina, contains Illumina universal primers, pre-diluted known concentration standards and KAPA SYBR FAST qPCR Master Mix. This kit is used according manufacturer's instructions, after real-time PCR, accurate concentration of adaptor ligated library can be calculated. Before setting up PCR reaction a serial dilution plate is prepared, 1/4000 and 1/8000 diluted template is used for RT-PCR reaction.

Pooling and Denaturation of Libraries: Six libraries; Library A back, Library A GFP+, Library A GFP-, Library B back, Library B GFP+, Library B GFP- were

prepared accordingly with the protocols which were explained above, also timeline for these processes is given in **Figure 4.1.**

Also steps for next generation sequencing can be seen in **Figure 3.7.** After calculation of qPCR results; denaturation process is done according to Illumina's suggestions. Templates which are diluted to 2nM based on RT-PCR were mixed with 0,1 N NaOH in same volume for denaturation to single strand DNA. Then, desired loading concentration libraries were diluted in HT1 solution, provided in the Cluster Kit.

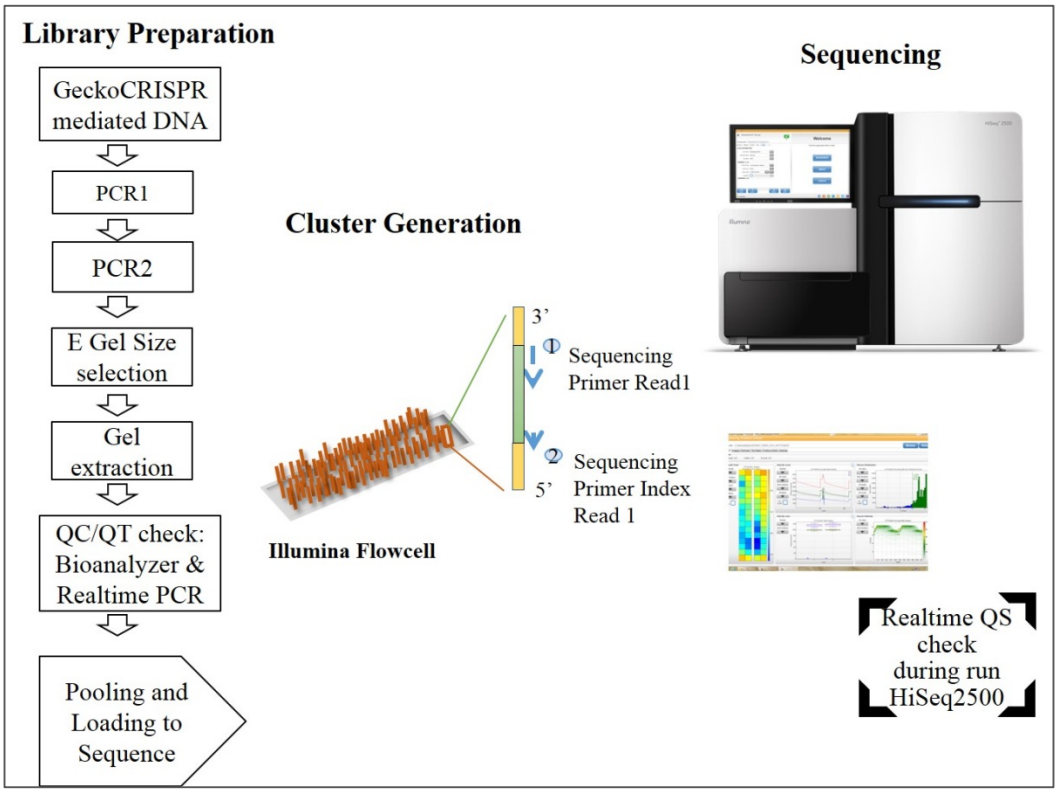

Figure 3.7 Workflow for NGS

4. RESULTS

4.1. Transduction of GeCKO Libraries into NK-92 cells

Genome wide loss-of-function screens are crucial experiments to enlighten several unknowns in terms of functional genome analysis. The set-up of those series of experiments (all experiments for GeCKO libraries can be seenin **Figure 4.1**) is designed to reveal antiviral pathways of NK cells via CRISPR knockout screen. In order to test CRISPR-mediated loss of function screen in NK cells, we have used the NK92 cell line as a model system. GeCKO libraries were designed and prepared by Zhang Lab from MIT and they were commercially available as plasmid pools. For human, GeCKO has two subgroup libraries, and when they are used together, one gene is targeted by six sgRNAs, and all of these sgRNAs are designed in such a way that they are specific to their target sites. For more detailed information papers of Zhang lab can be visited (Sanjana, Shalem, and Zhang 2014).

Both libraries were initially transduced separately into NK92 cell line in the presence of BX795, a chemical which was shown to increased NK cell transduction efficiency. Virus titer was carefully adjusted to get 20% transduction efficiency which ensures one cell is hit by only one sgRNA carrying viral particle. Preparation of GeCKO library viral particles and transduction of NK92 cells were carried out by our former group member, Dr. Canan Sayitoglu (Sayitoglu 2017).

Following library transduction, NK92 cells were selected for fourteen days with puromycin, to ensure that the cells that did not receive any viral particles were eliminated from the culture. When the selection process was completed, cells were

expanded and genomic DNA of two base libraries were extracted. Their DNA concentrations were measured as 220ng/ul.

For the second transduction, LEGO-G2 lentiviral vectors were used. This vector has GFP expression cassette which makes it possible to sort out the cells that received the virus by GFP positivity. During the second transduction, BX795 was not used in order to make sure that only the cells that easily accepted the virus (presumably the cells that were hit at an antiviral gene) were transduced. After LeGO-G2 transduction, cells were sorted for GFP positivity by FACS and expanded to reach sufficient numbers for genomic DNA extraction.

Genome editing of NK cells via lentiviral vectors is a very hard task due to their immune nature. The experiment design here was built to reveal virus dependent pathways when an NK cell confronts with the viral vector. Expectedly, base libraries from A and B sgRNA libraries, will show the gRNAs that achieved to integrate into the genome of the cells and hopefully lead to loss of function in their respective target genes. After the second transduction, sorted GFP+ cells included the cells which accepted the virus without any problems while GFP- cells were resistant to virus infection even after a random gene is silenced by CRISPR/Cas9. By comparing frequencies of the sgRNA counts of GFP+ or GFP- cells with the base libraries we were able to identify whether the knockout of a certain gene was associated with easier or harder integration of the GFP coding virus into the NK cell genome.

Since there are many constructs in GeCKO libraries (around 65000 in A and 55000 in B), it had critical importance to represent all the gRNAs adequately in the final library. To ensure a sequencing coverage of at least 300X, following calculation has been done: For having 300X minimum coverage for library representation for per sgRNA, each sgRNA should be read from 300 cells, since there is around 65000 constructs in GeCKO Library A, this requires about 20 million (650000 x 300) cells for Library A. Genomic DNA amount of a cell is accepted as around 6,6 pg/cell, thus genomic DNA from 1 million cell is approximately 6,6 µg. Therefore, at least 130 µg DNA was required for 300X representation of library A. Same calculation was done for library B as well and 115 ug genomic DNA was required for Library B.

Thus, to guarantee enough representation, 40 million cells were used for each library. All the required steps for this experiment is given below at **Figure 4.1** Workflow timeline for GeCKO sequencing libraries.

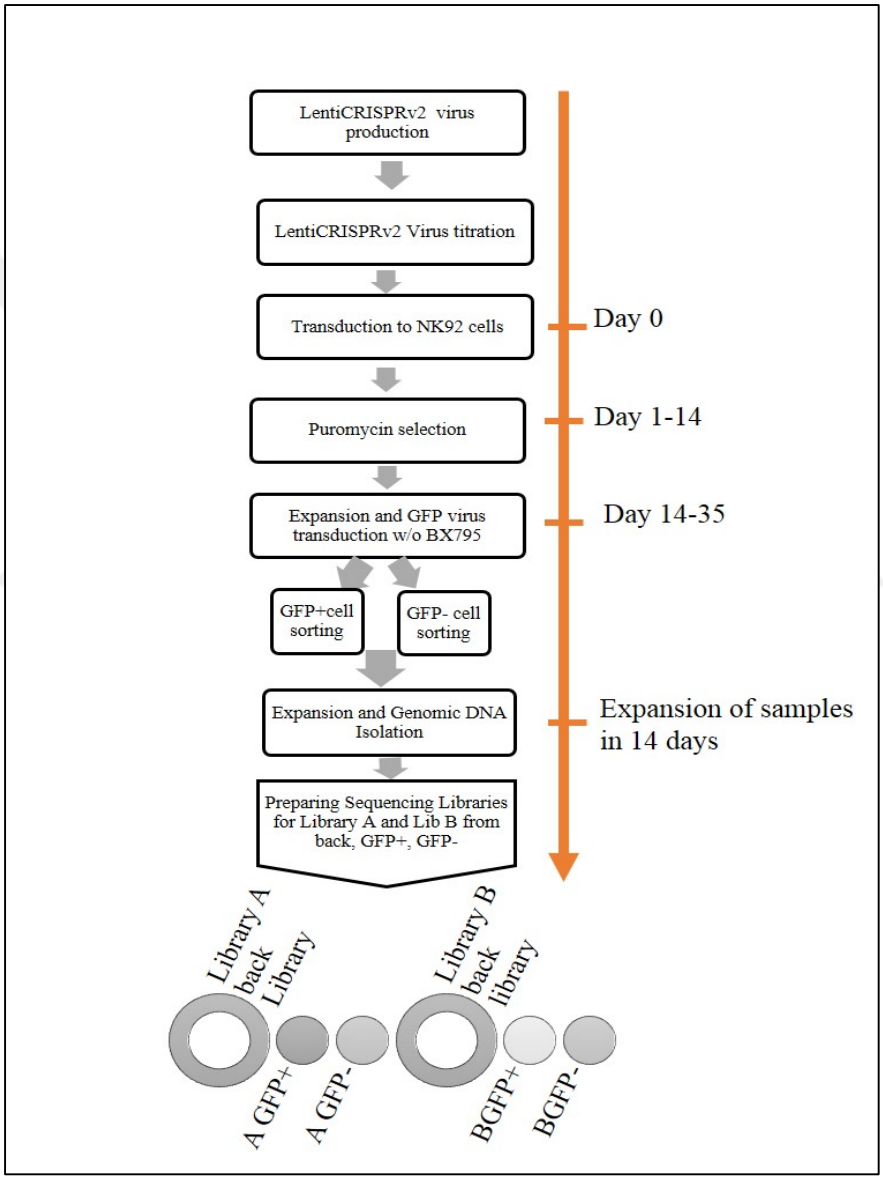

Figure 4.1 Workflow timeline for GeCKO sequencing libraries

4.2. Optimization of Library Sequencing

Two sets of oligos were designed for nested PCR as can be seen in **Table 3-4** Oligonucleotides list. All PCR reactions for amplification of target sgRNA library was done with NEB Phusion® High Fidelity polymerase. Details of the optimization steps are given below.

Based on the known sequence of lentiCRISPRv.2 vector and primers, expected band length for both PCRs were known. While PCR1 product is supposed to be ~280 bp, PCR2 would be 360-380 bp according to the stagger and barcode combination. As given in **Figure 3.5**, each of PCR2 primers has stagger sequences besides the barcode sequence. While barcode, in other words indexes, are necessary to multiplex different libraries in the same lane of HiSeq2500, in an amplicon based library, it is possible to have "monotemplate issue" that increases error rate of reading the clusters. In order to prevent this, stagger sequences are added to the design of PCR2 primers.

Although, it is not possible to see PCR1 band on agarose gel in low cycle numbers, it could be checked either following PCR2 reaction or increasing the cycle number for PCR1 in order to visualize the efficiency of PCR on an agarose gel. Gradient PCR1 is also applied to find best annealing temperature as can be seen in **Figure 4.2**

Optimized PCR1 set up was as in **Table 4-1**

PCR 1 setup	
DNA template(1,5 ug)	
5X Phusion GC Buffer	10ul
10 mM dNTP	1ul
10uM Forward	2,5ul
10uM Reverse	2,5 ul
Phusion HF Polymerase	0,5 ul
ddH2O	
Total	50 ul

Table 4-1 Optimized PCR1 setup

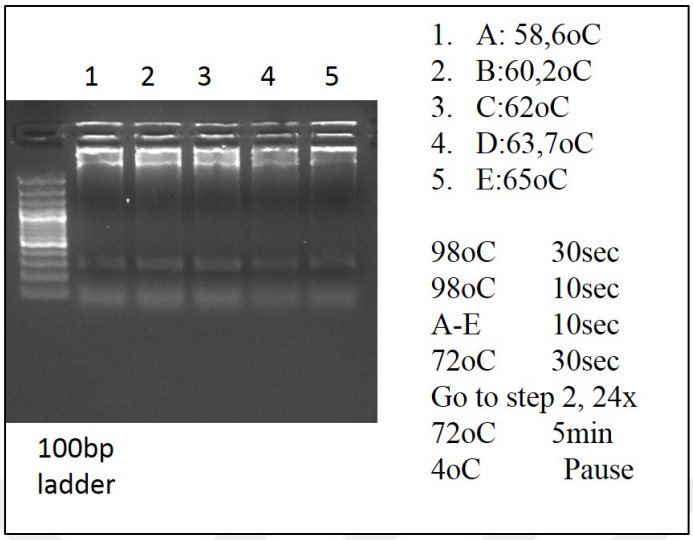

Figure 4.2 PCR1 gradient PCR

Lane 2 and Lane 3 were almost same in the agarose gel in **Figure 4.2**, to see its effect on the PCR efficiency after PCR2; two different conditions were tried respectively with the same PCR2 condition given in Optimized PCR2 . Since the PCR2 band that used the second lane of PCR1 (60°C annealing temperature) is slightly brighter than the third one; annealing temperature for PCR1 was defined as 60°C. PCR2 setup is optimized according to enzyme protocol as follows:

PCR 2 set up	
PCR1 product	1
5X Phusion GC Buffer	4 ul
10 mM dNTP	0,4 ul
10uM Forward F(03-08)	1 ul
10uM Reverse R(03-08)	1 ul
Phusion HF Polymerase	0,2 ul
ddH2O	12,4 ul
Total	20 ul

Table 4-2 Optimized PCR2 set up

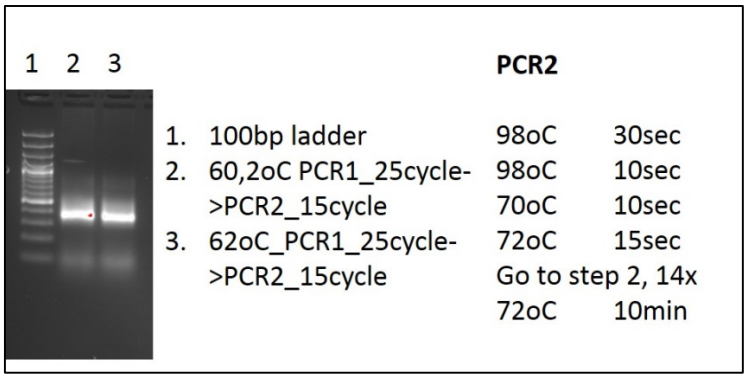

Figure 4.3 PCR1 annealing temperature

Also in order to see whether this band is exactly what we expect; PCR1 products were run on Bioanalyzer, via High Sensitivity DNA Kit. Although there are unexpected background noises since PCR product was not purified before loading to gel, expected band size was seen in this setting.

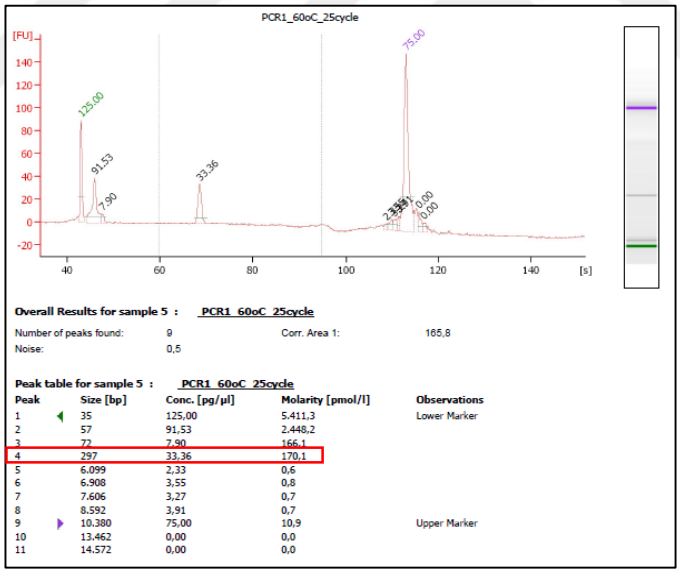

Figure 4.4 Bioanalyzer for PCR1 annealing temperature

Also product of trial PCR2 is also loaded to the bioanalyzer chip after PCR2;

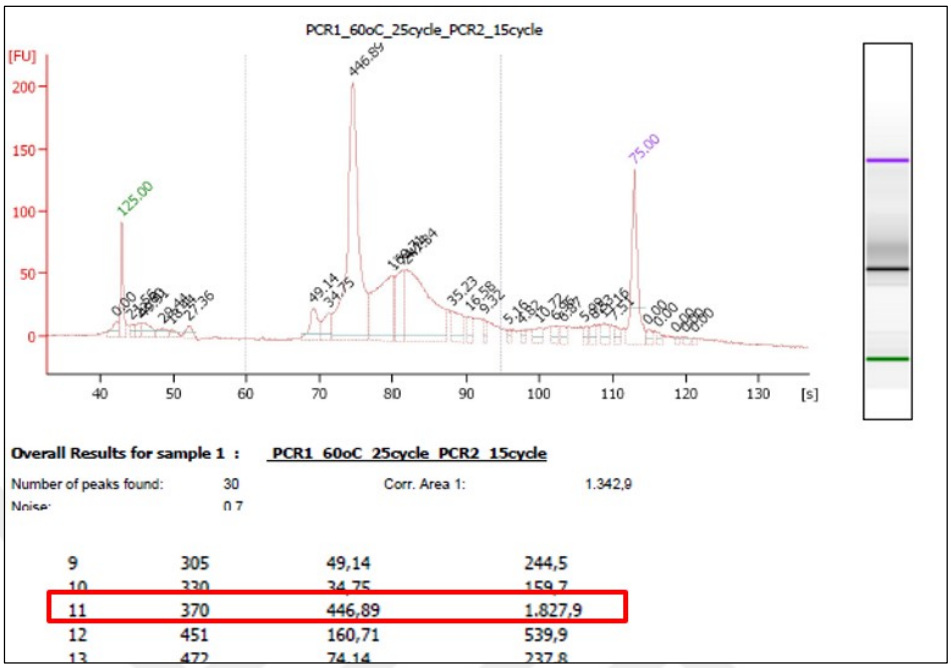

Figure 4.5 Bioanalyzer Screen for PCR 1 60°C ->PCR2

After seeing a specific significant band both in PCR1 and PCR2 reactions, more precise bands after PCR2 is required, thus optimization experiments were continued with adjusting the annealing temperature of PCR2 reaction.

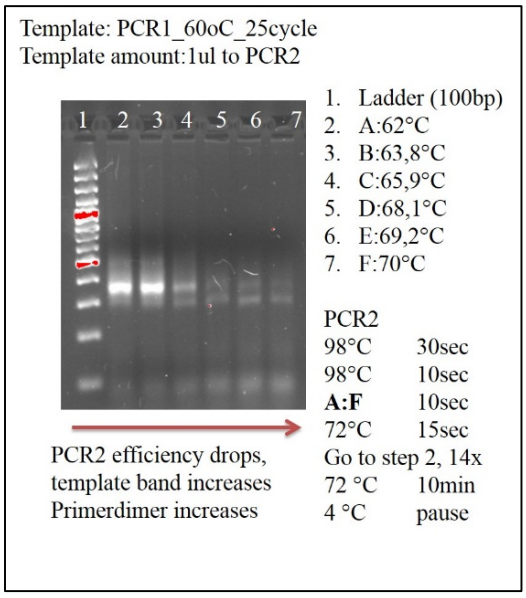

Figure 4.6 Gradient PCR2 optimization

As can be seen in **Figure 4.6**, 62°C annealing temperature PCR efficiency was highest, thus it was set for annealing.

PCR cycle optimization was a crucial step to minimize undesired effects and PCR-based bias from the libraries. The number of cycles were decreased with a series of experiments. For this purpose, different cycles of PCR2 in optimized conditions were applied and as it can be seen **Figure 4.7**, 15 cycle PCR2 was assessed to be enough for continuing to E-Gel size select step.

Amount of input templates were calculated according to the instructions in previous paragraph: 130 µg Library A and 155 µg for Library B in PCR1 step. Since the optimized PCR1 set up has 1,5 ug DNA template, 87 tubes of PCR1 were required for Library A, while 77 tubes of PCR1 is done for Library B.

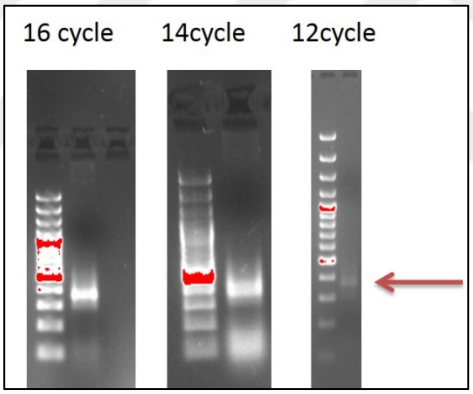

Figure 4.7 PCR2 cycle determination(100bp ladder)

4.3. Pre-Sequencing Library Controls

After PCR1 reactions of each library were completed, all reactions from PCR1 wells for each library were pooled in a single tube and mixed well. As recommended, one PCR2 was set for 10K constructs. Since there is 65000 constructs for library A and 57000 construct for library B; 7 PCR2 reactions for A libraries and 6 for B libraries were done. Then, these wells were loaded to the E-Gel Size Select gels and size selection executed as described in methods part. After gel extraction and elution with EB buffer in 20 ul, Bioanalyzer was run via High Sensitivity DNA Kit for libraries. Bioanalyzer results can be seen in **Figure 4.8**.

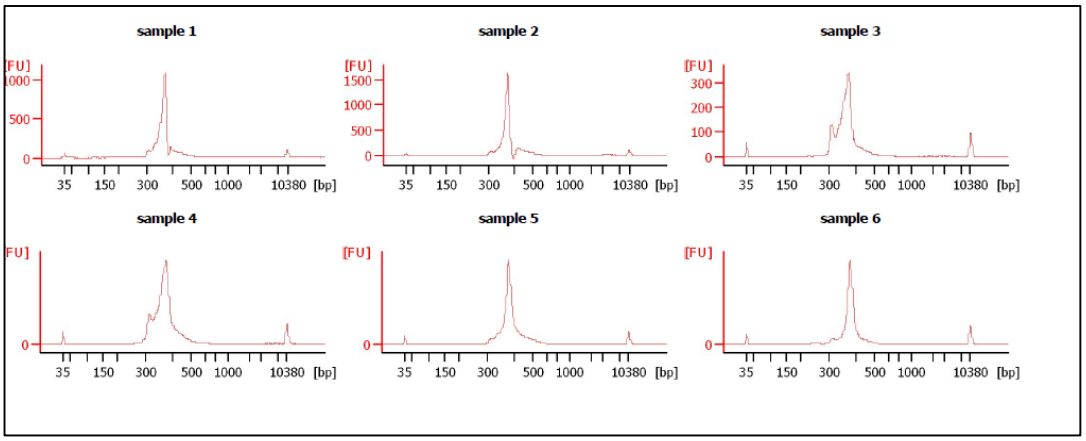

Figure 4.8 Bioanalyzer of libraries with high sensitivity DNA chip

It was important to have a clear band in Bioanalyzer for following pre-sequencing steps. In order to have proper reading in HiSeq2500 all the input fragments should be approximately in same length and also any adapter dimer contamination can make a huge data waste which prevents enough depth on reads. Real-time PCR was applied for measuring concentration of adaptor ligated libraries which have Illumina universal adapters.

Real-time quantification kit had 6 adaptor ligated standards, control libraries, that are already diluted to known concentrations so as using their amplification data PCR efficiency curve is drawn in **Figure 4.9.** Also PCR efficiency was calculated as 96% through the equation below.

$$\textbf{Efficiency} = \left[\left(\frac{-1}{slope}\right) - 1\right] \times 100\ \%$$

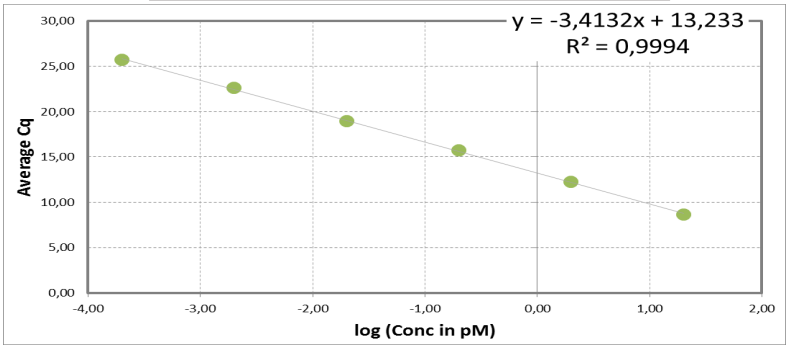

Figure 4.9 Standard curve for qPCR

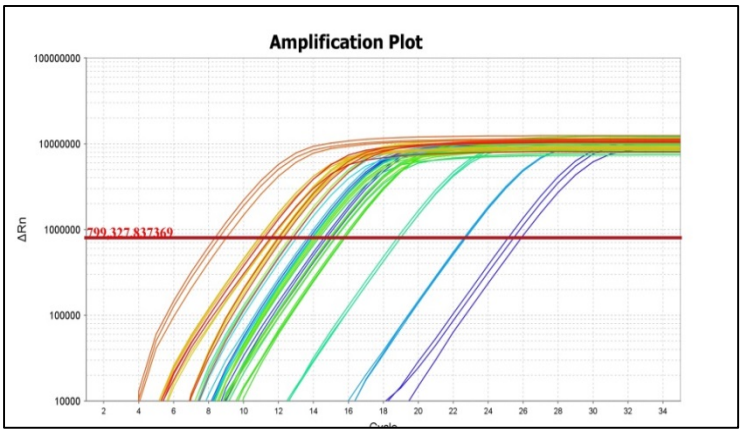

Figure 4.10 Amplification of determined standard and samples

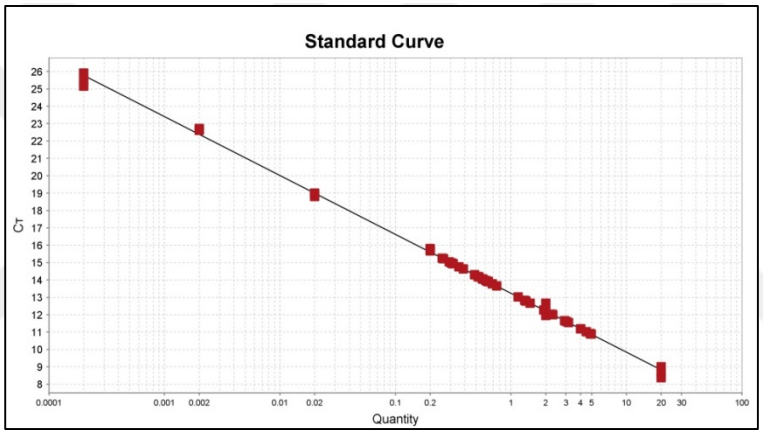

Figure 4.11 CT values of six libraries across Standard curve

Samples had been diluted prior to use as real-time templates in EB buffer as 1/4000 and 1/8000 and they were loaded as triplicates. All PCR setup is done according to kit's protocol. CT (threshold cycle) values of the libraries should be between these six standards to have an accurate measurement of concentrations and standard deviations > 0,5 is omitted. Sample and standards amplification plots were given in **Figure 4.10** and while CT values can be seen in **Figure 4.11.**

$$Library\ Stock\ Concentration = \mathrm{A} \times \frac{452}{Length} \times Dilution$$

A: Conc. Calculated by qPCR
Length : Average fragment Length

Based on the equation above concentration of libraries is calculated as follows in **Table 4-3**

#	Barcode Sequence	Sample name	Concentration of Undiluted Library (nM)
1	CGCGCGGT	Base Library A	17,72
2	CATGATCG	Base Library B	24,62
3	CGTTACCA	A GFP+	2,76
4	TCCTTGGT	A GFP-	2,78
5	AACGCATT	B GFP+	3,28
6	ACAGGTAT	B GFP-	3,03

Table 4-3 Library concentrations and barcodes

Furthermore, Melting Curve analysis was done during qPCR and the result plot is shown below; all the libraries have one length amplified fragment as can be seen in **Figure 4.12.**

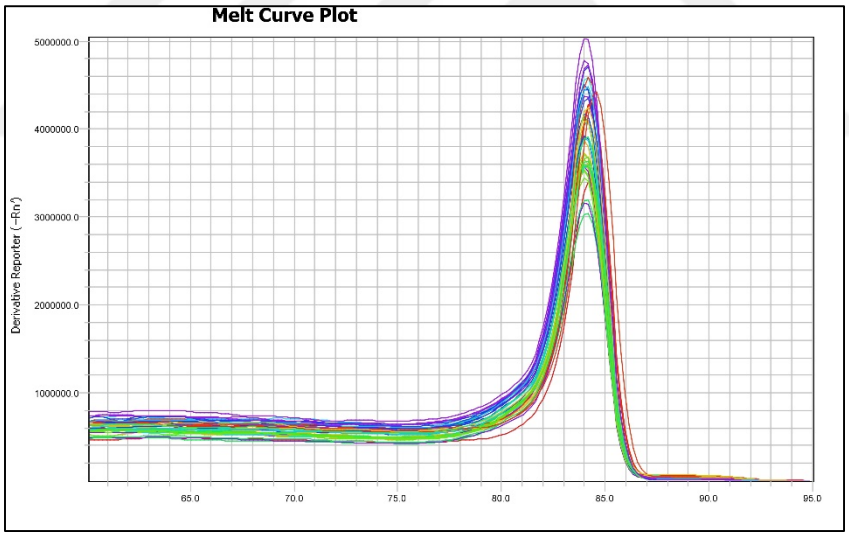

Figure 4.12 Melting curve analysis of libraries in qPCR

Dilution was done according to qPCR results, then each library was denatured separately and all of them were diluted to 10 pM before pooling equal volumes. Also 5% pHix control library is added to the pool, in order to eliminate risk of mono-template issue on HiSeq and also reveal an accurate error and phasing rate for sequencing.

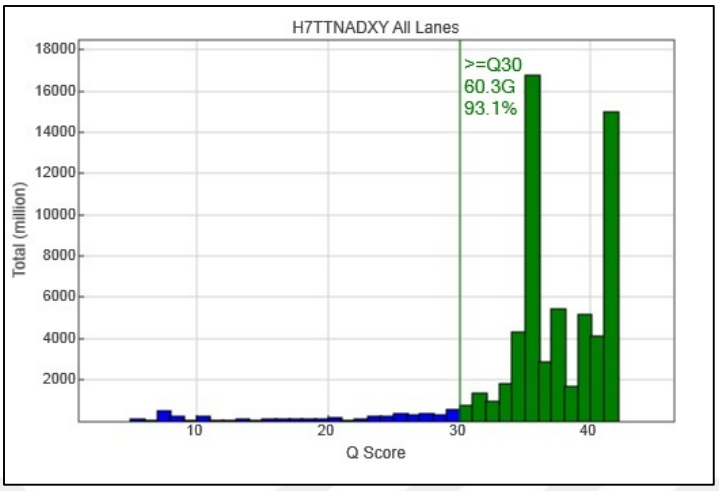

Figure 4.13 Phred score in HiSeq Control Software

One of the results screen of HiSeq2500 Sequencing Software was given in for this run is **Figure 4.13,** this plot shows the Phred score for read that is calculated by the HiSeq machine. Phred score calculates logarithmic base calling error probabilities via the equation given below. When Phred score is given as Q30 like in the plot above it means the probability of an incorrect base calling is only 1 in 1000.

4.4. Bioinformatics Data Analysis

When sequencing was accomplished, there were approximately 30Gb data generated for 6 of GeCKO libraries. Illumina sequencers produce data as ".bcl". It was converted into ".fastq" format via the CASAVA v1.8.4 program and multiplexed according to indices for further bioinformatic analysis. (Illumina Inc. [CASAVA v1.8.4], 2013)

After conversion, all reads of library were subjected to FastQC software which is a quality control step before going into sequencing analyses (Andrews 2017). It tests both the sequencer and the sequencing library quality. Since forward reads were necessary to be used in the following analyses in this experimental setup, forward reads were tested and all the following analyses were based on forward reads. 6 libraries were checked via FastQC, the report of the software was almost the same

for all libraries. All FastQC per base quality images of libraries are in APPENDIX C. As can be seen in the figures the reads for libraries were all in green region which shows high quality of the read- generally 20 is accepted for quality limit-.

Following FastQC analysis, reads were trimmed via in house Python scripts to the sgRNA target sequence from 5'. Stagger and forward indices were trimmed from the library, also the remaining sequences were cut after 20bp from the gRNA, by this way only target sequence of sgRNAs were left from the reads. Since the rest of the reads are identical, it was not necessary to take them into account for further steps.

Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout (MAGeCK) was used for further analysis, this tool is designed for CRISPR loss-of-function genomic screens mainly.(W. Li et al. 2014) Firstly, count command was applied via fastq reads.

Count summaries of the libraries can be seen in the **Table 4-4** and **Table 4-5.** As expected there were at least 13 M read for both A and B libraries, that means at least 200X coverage were taken on reads (65K×200±13M)). Almost 70% of the reads were mapped to the reference library that was the outcome of no mismatch alignment search via Bowtie alignment in the count command of MAGeCK tool. Only approximately 5K sgRNA were not read in the base/control libraries which were a sign of first transduction with BX795 were successful and most of the sgRNAs had integrated to NK cell successfully. This 5K sgRNA loss on control library can be due to these sgRNA were targeting lethal genes in NK cells so they were not present.

Label	Reads	Mapped	Percentage	TotalsgRNAs	Zerocounts
AGFP+	18692069	13323294	0.7128	63950	11271
LIBA	15861549	11957933	0.7539	63950	5257
AGFP-	22021710	15685532	0.7123	63950	40549

Table 4-4 Count summary for library A

Label	Reads	Mapped	Percentage	TotalsgRNAs	Zerocounts
BGFP+	17845355	13490403	0.756	56869	29693
LIB-B	14622097	11441343	0.7825	56869	4136
BGFP-	16816592	12922187	0.7684	56869	27699

Table 4-5 Count summary for library B

In order to reveal distributions of the sgRNA read counts among libraries following histograms were drawn. As can be observed, distribution of sgRNA read count among base libraries were consistent with the expected ratios and most of sgRNAs on these libraries were sequenced more than 10 times. On the other hand, second transduction via GFP virus seemed to manipulate sgRNA counts in both GFP+ and GFP- libraries according to their histograms. This can be also validated in count summary tables (**Table 4-4** and **Table 4-5**) while some sgRNAs were read in high numbers in these GFP libraries a considerable count of sgRNA were diminished, for example 40K sgRNA were absent in library A GFP-, also half of the sgRNA (30K) were absent in these sub-libraries of Library B.

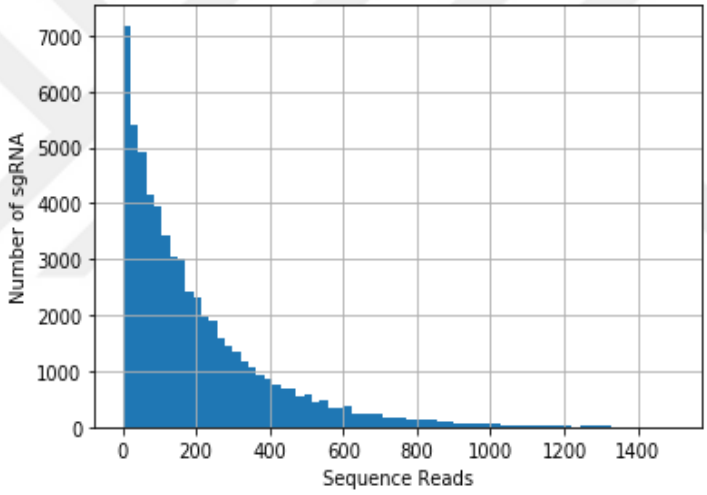

Figure 4.14 Histogram of Library A base library

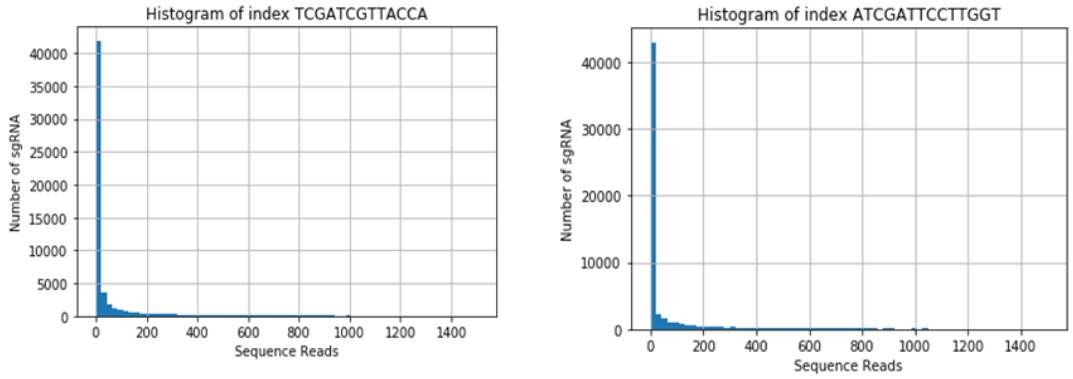

Figure 4.15 AGFP+ and AGFP- histograms, respectively

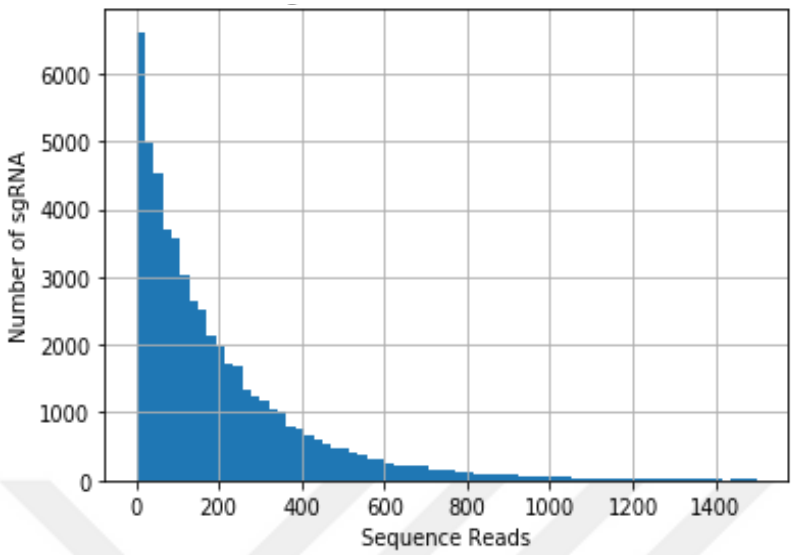

Figure 4.16 Histogram of Library B base library

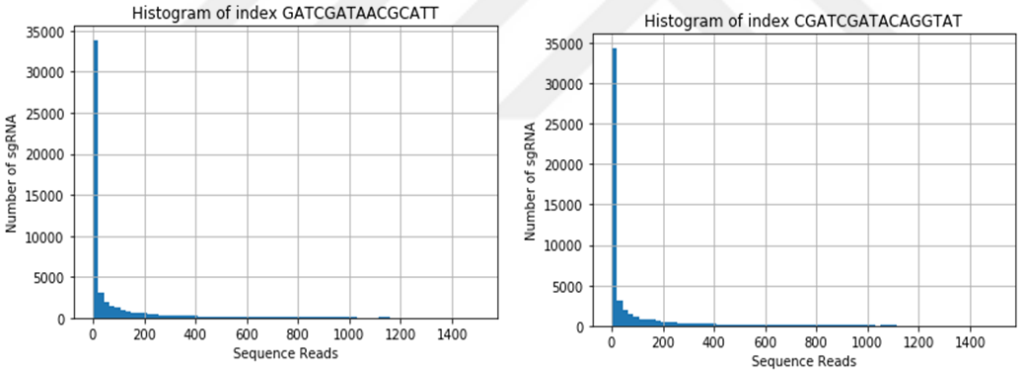

Figure 4.17 Library BGFP+ and BGFP- histograms respectively.

Then output files that includes count table of all libraries were used for "test" command in MAGeCK tool. Example of a count and test command format can be seen in **Table 4-6.**

>$mageck count -l Referencelibrary.csv -n Outputfoldername --sample-label_R01, R02 R03" --fastq InputFastQFiles (index_R1, index_R02, index_R03) >$mageck test –k count_table.txt -t R01 -c R02 -n Ref-Cntrl.txt

Table 4-6 Command format of MAGeCK tool; count command and test command respectively.

This "test" command compares two count tables as reference library and observed library. It operates several algorithms like α-RRA (robust ranking aggregation) that considers all sgRNAs hits to that gene and makes a ranking list for both negative and positive selection. α-RRA is mainly based on the assumption that if a gene has no role in selection, sgRNAs targeting that gene should be regularly distributed across control and virally infected libraries. It calculates p-values of the genes accordingly (Kolde et al. 2012; W. Li et al. 2014). Target gene ranking are listed in the gene summary output format. First 20 genes of all libraries were shared in the tables below. (**Table 4-5,6,7,8**) Only positive selection result columns for GFP+ and negative selection result columns for GFP- libraries were given due to space considerations.

Id	#	Score	p-value	Fdr	Rank	Good sgrna	Lfc
Dummy guide	1K	2,60E-09	2,40E-03	0.00495	1	262	0.32696
TTC40	3	2,35E-01	7,59E-01	0.784653	2	2	20.402
hsa-mir-557	4	5,00E-01	0.00017413	0.981921	3	4	29.974
HSPB9	3	7,04E-01	0.0002182	0.981921	4	2	52.338
LRRC3B	3	0.00010918	0.00033891	0.981921	5	3	62.139
TMEM255B	3	0.00011727	0.00035855	0.981921	6	2	12.136
TSPAN7	3	0.00013499	0.00040933	0.981921	7	3	65.714
CLYBL	3	0.00015616	0.00047591	0.981921	8	3	89.384
DNMT1	3	0.00016418	0.00049364	0.981921	9	2	39.754
ZNHIT2	3	0.00017996	0.00053483	0.981921	10	3	57.819
CDS1	3	0.0001828	0.00054202	0.981921	11	2	52.748
STOM	3	0.00022102	0.00065507	0.981921	12	3	40.696
BCL10	3	0.0002344	0.00069339	0.981921	13	3	43.609
PADI2	3	0.00026789	0.00078728	0.981921	14	3	59.121
MDGA1	3	0.00028238	0.00082656	0.981921	15	3	56.479
ZNF317	3	0.00030489	0.00088213	0.981921	16	1	32.712
hsa-mir-5690	4	0.00034397	0.0012749	0.981921	17	1	-0.32559
SMTNL2	3	0.00034901	0.001034	0.981921	18	3	58.155
NPAT	3	0.0003518	0.0010426	0.981921	19	1	-0.069936
ACTL9	3	0.00035227	0.0010445	0.981921	20	2	0,0631

Table 4-7 AGFP+ gene summary output file
Column explanations for this gene summary table is: ID column gives gene IDs, num; is the number of sgRNAs targeted for the gene, the rest column is estimated both for positive or negative; score, RRA algorithm lo value for this gene, p- value, calculated p-value via permutation, fdr, false discovery rate for the gene, rank, ranking of the gene according to selection, goodRNA is number of sgRNAs whose ranking is below the alpha cutoff in negative selection, lfc; log fold change of this gene.

Id	#	Score	p-value	Fdr	rank	Good sgrna	Lfc
hsa-mir-3065	4	4,04E-02	1,13E-01	0.200495	1	4	-72.853
hsa-mir-214	4	6,68E-02	1,94E-01	0.200495	2	4	-77.791
SULT2B1	3	2,35E-01	7,59E-01	0.392327	3	1	0.28788
SUPT20H	2	2,84E-02	5,48E-01	0.377888	4	2	-95.146
PRODH2	3	4,36E-01	0.00014012	0.56082	5	2	-55.458
SYT11	3	5,96E-01	0.00018658	0.56082	6	2	-95.463
HR	3	6,10E-01	0.00018993	0.56082	7	3	-90.062
MDH2	3	7,04E-01	0.0002182	0.563738	8	2	-64.352
ENTHD2	3	8,06E-01	0.0002455	0.563806	9	2	-77.966
DVL1	3	0.00011727	0.00035855	0.741089	10	2	-6.746
LTC4S	3	0.00018609	0.0005492	0.787248	11	2	-66.405
ADI1	3	0.00018708	0.00055112	0.787248	12	3	-86.007
SUCLG1	3	0.00021109	0.0006201	0.787248	13	1	-38.944
PGF	3	0.00021388	0.00063016	0.787248	14	3	-84.942
hsa-mir-8060	4	0.0002189	0.00081937	0.787248	15	1	-49.299
hsa-mir-3670-1	4	0.0002499	0.00092955	0.787248	16	3	-71.757
DOPEY2	3	0.00025799	0.00075854	0.787248	17	3	-75.024
C16orf52	3	0.00026692	0.00078297	0.787248	18	3	-84.563
AHRR	3	0.00027201	0.00079638	0.787248	19	3	-86.842
TSC1	3	0.00029002	0.00084333	0.787248	20	3	-81.203

Table 4-8 AGFP- gene summary output file

ID	#	Score	p-value	Fdr	rank	good sgrna	lfc
dummy guide	1K	2,60E-11	2,63E-03	0.002475	1	575	12.54
TIAF1	3	1,24E-03	2,63E-03	0.002475	2	3	80.42
RHBDF2	3	2,64E-01	8,49E-01	0.483911	3	3	46.62
CBWD2	2	5,44E-01	0.00010277	0.483911	4	2	91.66
GSTM3	3	7,91E-01	0.00024417	0.789829	5	2	21.58
KDR	3	0.00010779	0.00033564	0.789829	6	3	69.27
RRM2B	3	0.00011349	0.00035036	0.789829	7	2	97.66
HTR2C	3	0.00013188	0.0004024	0.789829	8	2	53.58
CH25H	3	0.00013586	0.00041186	0.789829	9	3	69.78
TTC9B	3	0.00014883	0.00045917	0.789829	10	2	98.55
TMEM208	3	0.00014983	0.00046128	0.789829	11	3	68.36
DLST	3	0.00018462	0.0005601	0.839418	12	3	38.38
TXNDC16	3	0.00019396	0.00058428	0.839418	13	2	87.31
COX7C	3	0.00022734	0.00068363	0.839418	14	2	84.25
SCYL1	3	0.00023737	0.00071202	0.839418	15	3	76.64
EBP	3	0.00023805	0.00071307	0.839418	16	3	82.41
UCHL5	3	0.00029011	0.00086499	0.922167	17	2	3.93
FAM180A	3	0.00029676	0.00088129	0.922167	18	2	62.78
RFESD	3	0.00034326	0.0010285	0.960526	19	3	69.60
FFAR3	3	0.00036922	0.0011057	0.960526	20	2	24.67

Table 4-9 BGFP+ gene summary output

ID	#	Score	p-value	Fdr	rank	good sgrna	Lfc
LRRN4CL	3	1,89E-01	6,23E-01	0.198727	1	2	-10.473
LAPTM4B	3	1,89E-01	6,23E-01	0.198727	2	2	-10.473
HAUS7	3	1,89E-01	6,23E-01	0.198727	3	2	-10.473
LYG2	3	1,89E-01	6,23E-01	0.198727	4	2	-10.473
TMEM233	3	1,89E-01	6,23E-01	0.198727	5	2	-10.473
PSMB3	3	2,28E-01	7,39E-01	0.198727	6	3	-12.224
RPS10	3	2,28E-01	7,39E-01	0.198727	7	3	-0.848
ARHGAP21	3	3,67E-02	0.0001164	0.274134	8	3	-10.473
TP53RK	3	0.0001028	0.0003198	0.669113	9	3	-16.477
ECE2	3	0.0001344	0.0004097	0.669113	10	3	-19.758
C6orf136	3	0.0001693	0.0005191	0.669113	11	3	-20.704
RNMT	3	0.0001714	0.0005259	0.669113	12	2	-0.848
ZNF560	3	0.0001714	0.0005259	0.669113	13	3	-0.848
RBBP8	3	0.0001714	0.0005259	0.669113	14	3	-10.473
APOBEC1	3	0.0002563	0.0007667	0.669113	15	3	-22.429
MRPL39	3	0.0004881	0.0014432	0.669113	16	2	-12.224
HOXA1	3	0.0004881	0.0014432	0.669113	17	2	-10.473
SLC2A13	3	0.0004881	0.0014432	0.669113	18	3	-12.224
WSB2	3	0.0004881	0.0014432	0.669113	19	2	-10.473
HINFP	3	0.0004881	0.0014432	0.669113	20	2	-0.848

Table 4-10 BGFP- gene summary output

According to calculated p-values there were around 1000 genes for each library that were significantly selected ($p<0,05$) both in library A and library B. To estimate essential genes for revealing antiviral and proviral pathways; genes that were shown to be significantly changed in both library A and library B is taken to the pathway analysis.

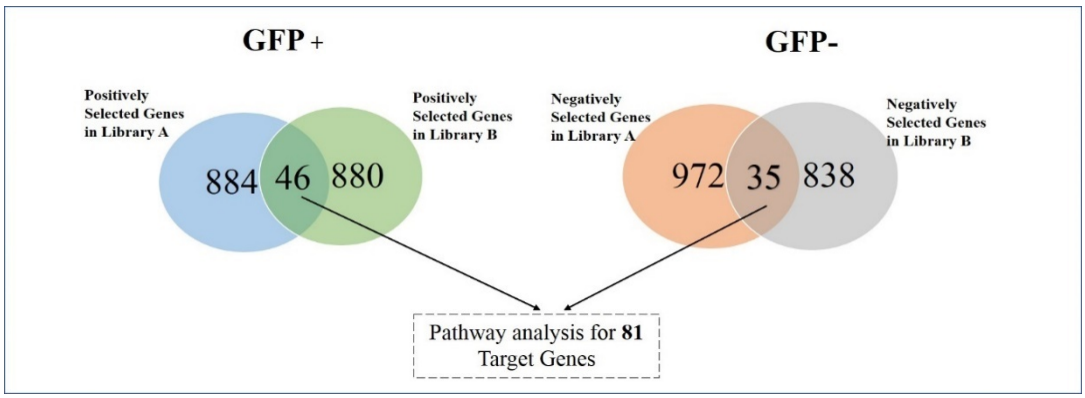

Figure 4.18 Antiviral pathway analysis scheme

In antiviral pathway analysis GFP+ libraries were ranked as positively selected genes while GFP- libraries were ranked according to negatively selected. Because when a sgRNA count is increased in GFP+ libraries in contrast to its base library, this means the gene were selected during the second transduction. Since this gene silencing allowed a second transduction with GFP virus, this candidate gene would work against viral vector entry which means belong to antiviral pathways. In contrast, GFP- libraries were not susceptible to a second transduction which means that silencing the corresponding gene leads to a stronger attitude against infection: Thus, negative selection of GFP- library derived gene output would give antiviral mechanisms as well.

In order to use both representation from library A and library B common genes of GFP+ and GFP- libraries that were shown to be significantly changed (p value<0,05) were searched. Then the common genes were merged in a table which were 81 candidate genes for antiviral pathways. (**Figure 4.18**)

Pathway analysis were done with REACTOME database (Fabregat et al. 2017) via the web tool g:profiler. (Reimand et al. 2016). The outcome of pathway analysis is given in **APPENDIX A.** Only pathways that have p-values are below 0,005 are given and they were taken into account during discussion part.

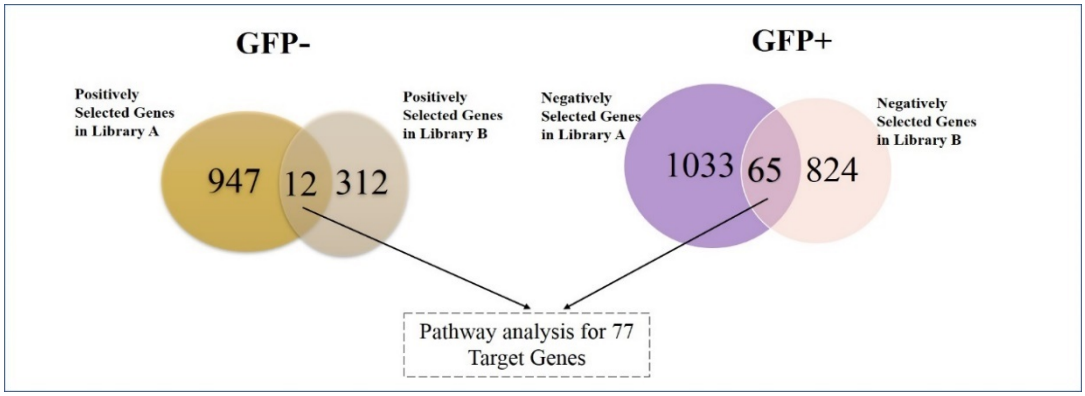

Figure 4.19 Proviral pathway analysis scheme

For revealing candidate proviral pathways; in other words, virus or vector hijacking mechanisms over host cell was found as merging common genes from both libraries (A and B) as negatively ranked GFP+ genes and positively ranked GFP- genes. When a gene silencing results with a host cell that does not allow viral vector entry, this may mean that normally this gene would be carrying a helper role for virus to come in. So any proviral gene would be positively selected in GFP- libraries and negatively selected in GFP+ libraries. Again, in order to use both representation from Library A and B, we searched common genes among them and 77 gene were employed in pathway analysis as given in **Figure 4.19**. The output table of pathways that possess p-values lower than 0,05 is given in the **APPENDIX B**.

Furthermore, these candidate gene sets for both antiviral and proviral pathways were analyzed with GO (Gene Ontology) Cellular Component Analysis via g:profiler web tool , which gives the expression places of the genes on the cells. This analyses were also important to reveal the physical places of mechanisms against or in response to viral vector entry or infection. Analyze output tables were given in **APPENDIX D.**

5. DISCUSSION

NK cells function as part of the innate immune system. Therefore, they have several features for detecting and eliminating pathogens. They are attractive candidates for immunotherapy but because they are fully armed against viral agents, their genetic modification is a daunting task.

In order to reveal which mechanisms, become active during viral vector entry to the cell and to identify downstream events that are manipulated by viral vector we performed a genetic screen. For this screen, we used GeCKO CRISPR knockout libraries to mutate target genes in NK cells and assessed the enrichment or depletion of cells from populations of lentivirally infected cells. The NK92 cell line which is derived from a malignant non-Hodgkin's lymphoma used in these experiments for revealing NK cell antiviral and proviral pathways.

Transduction efficiency of CRISPR library was adjusted such that only one sgRNA would target one cell and lead to a DSB only at a single gene. After transduction with CRISPR libraries, cells were selected with puromycin and expanded. Then, a second transduction was carried out with LEGO-G2 GFP expressing viruses and cells were sorted to separate GFP expressing cells from non-GFP expressing ones. Each population was expanded to 40 million cells, genomic DNA was extracted and PCR was performed to identify the incorporated sgRNA gene. In this study, amplicon based sequencing was applied and for this purpose the suggested barcode and stagger designs from Addgene were used for amplification of targeted sgRNA sequences. Since loaded libraries were identical except 20 bp target sequences, it would create a monotemplate issue during the HiSeq2500 run; a problem of Illumina reads when the

same signals are received from many clusters at the same time. In order to circumvent this problem and to increase the quality of our sequencing data, samples were run with other transcriptome libraries to increase the diversity between clusters. Also each sample was mixed with an Illumina pHIX control library (%5) to monitor phasing and percentage of the error during read. Libraries were read as paired end, but only the forward reads were used to analyze amplicon sequences. This precaution was taken because there was a risk of failed multiplexing of in-house barcode indices with the TruSeq adapters Illumina warns its customers when using custom adapters, In this case, forward barcodes would be used for further analyses and custom multiplexing. We found that our in-house barcodes created a problem during multiplexing indeed, the CASAVA program failed to separate libraries through their indices. Therefore, forward reads were used for distinguishing libraries through indices. After demultiplexing, all the index sequences and stagger sequences were trimmed from the data and MAGeCK was used for further analyses

CRISPR loss-of-function genomic screen data is not easy to analyze and produce a meaningful output because it requires a replicate model in order to estimate expected read counts for sgRNA and most of the studies were not applied with replicates due to limitation about setup. Besides, sgRNA counts are extremely variable between sample groups and it is difficult to calculate correlations in this over-dispersed scheme. Another limitation for our analysis stems from the fact that is there are multiple, for example 6 sgRNAs for each gene and their effectivity and specificity to target gene is variable, also their counts are variable, thus the functional significance of each sgRNA can only be estimated and needs to be reconfirmed with individual sgRNA targeting.

Amplicon sequencing analyses usually have been done via transcriptome sequencing tools. Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout (MAGeCK) was developed only for these CRISPR/Cas9 loss of function screens and in order to eliminate unwanted outcomes of both CRISPR and amplicon sequencing, it has several advantages to use for GeCKO library sequencing. This tool starts with normalization of sgRNA counts via median of each library read count. Next, negative

binomial (NB) model with RRA is applied to calculate p- values of counts and show whether the sgRNA count change between groups is statistically significant or not. Then it makes a ranking according to p-values of the genes that are positively and negatively selected. As we compared the positively selected and negatively selected genes, two lists were prepared according to common genes between enriched samples of cells infected with the A and B viral libraries. This analysis scheme is shown in more detail in **Figure 4.18**and **Figure 4.19.** These genes were searched in the REACTOME databases and pathway lists were prepared accordingly. We focused on the known antiviral innate immune response genes to analyze the output of our screen.

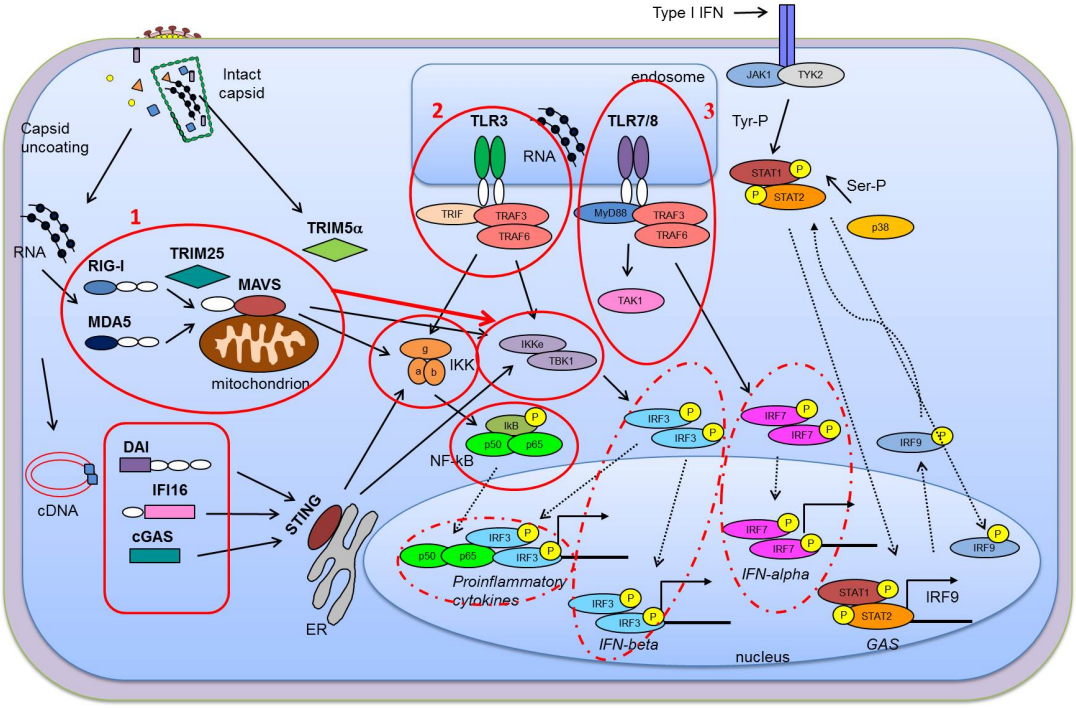

Figure 5.1 Several known antiviral pathways against RNA virusIn the pathways analysis of antiviral candidate genes; RIG-I/MDA-5 dependent Type I interferon secretion is occurred as top in the list. (Highlighted in red in figure- 1) Downstream complexes were also in the pathway list like IKK, TBK1/IKKε complexes. Toll like receptor cascades, Especially TLR3 and TLR7 also Myd88 dependent downstream signaling events were also in pathway list. (2 and 3) Also cytoplasmic sensor pathway were in the list. Along with Type I interferon secretion and NFκB activation were prominent as well in the list.(adapted from (Sayitoglu 2017))

RIG-I/MDA5 mediated induction of the IFN-α and β pathway was one of the highly significant results of pathway analysis which was parallel with previous findings. RLRs are intracellular sensors for foreign RNA molecules in the cytoplasm. S100A12, IRF2, IFNB1, NKIRAS2 were all identified to be enriched in the GFP+ libraries when both the A and B libraries were analyzed.

As described in the introduction part, RIG-I and MDA5 dependent IFN I (α and β) stimulation after viral infection was shown in previous studies and found to be regulated by the transcription factors IRF3 and NF-κB (Yoneyama et al. 2004). In the mouse innate immune system, TLR and RIG-I activation were shown to increase type I interferon and inflammatory cytokine secretion *in vivo* during influenza infection (K. Zhou et al. 2016) In a previous study by our group, BX795, an inhibitor molecule of the TBK1/IKKε complex, which resides in the RIG-I/MDA5 and TLR3 pathways was demonstrated to increased NK cell viral transduction efficiency (Sutlu et al. 2012).The IFNB1 gene encodes for IFN β which is a type I interferon and one of the primary cytokine was found to be related with crucial innate immune responses against viruses, mainly RIG-I/MDA5 mediated pathway, Toll Like receptor pathways were demonstrated to have IFNB1 regulation in their downstream signaling transduction events as well (Koyama et al. 2008).

Since most of the genes of PRR and pathogen recognition pathways share common features it is not easy to say which mechanism was most prominent in lentiviral vector recognition and which antiviral signals belongs to which mechanism in NK cells. The positively selected genes proposed that Toll-Like receptors cascades were also prominent. Since our lentiviral vector is an RNA virus, as parallel with the previous knowledge apart from RIG-I/MDA5 pathways, TLR3 and TLR7 expression increase would be expected (Breckpot et al. 2010; Pisegna et al. 2017; Beignon et al. 2005).

UBC was another promising gene that can be related with several innate immune system pathways and as being crucial for other pathways in the cells, already found critical for RIG-I/MDA5 pathway as well. Ubiquitination is a post translational modification event when a covalent attachment of ubiquitin protein to target protein

takes place. Since viruses are restricted intracellular parasites, they use these ubiquitination events to manipulate host cell antiviral mechanisms. RIG-I and MDA5 CARD domains have K63 polyubiquitin chain binding motifs and any disruption on this part were found to be impaired this pathway activation additionally IRF3 and IFNβ secretion was eliminated (Jiang et al. 2012). There is also UBC gene which was positively selected for antiviral pathways in our gene list and presence of several other genes within RIG-I/MDA5 pathway, which can show RIG-I mediated response of NK cells. PSMD1 (Proteasome 26S Subunit, Non-ATPase 1) is another gene also responsible for ATP-dependent degradation of ubiquitinated proteins were positively selected in GFP+ libraries which is thought to be part of innate immunity pathways (Heaton, Borg, and Dixit 2016).

BCL10 (B-cell CLL/Lymphoma 10) is a protein coding gene which was found to be crucial for NF-κB stimulation via B and T cell receptors. Bcl10 knockout mice models could not respond to antigen receptor activation and could not proliferate (Ruland et al. 2001). In later studies, overexpression of Bcl10 gene could stimulate NF-κB with the need the important regulatory subunit of IKB kinase complex, IKKγ (NEMO) was able to be demonstrated (Lucas et al. 2001). In a paper from 2003, they were able to reveal that Bcl10 makes ubiquitination on IKKγ and when an altered form of IKKγ which cannot be ubiquitinated was used, there was no stimulation of NF-κB expression. Furthermore, UBC13 and paracaspase enzyme were required for this ubiquitin addition (H. Zhou et al. 2003). Another study on NF-κB activation regulation claimed that Bcl10 activates IκB kinase complex and result into the phosphorylation and degradation of the IκB inhibitors of NF-κB. In the paper, they demonstrated IκB kinase complex phosphorylates Bcl10 after antigen recognition and Bcl10 exposed to proteolysis by the beta-TrCP ubiquitin ligase/proteasome downstream events. Besides they demonstrated Bcl10 that is altered from according phosphorylation site failed to degrade and accumulated in nucleus and result into Il-2 production. Thus, Bcl10 plays a regulatory role on NF-κB activation (Lobry et al. 2007).

NF-κB stays in inactive form in the cytoplasm as bound to IκB. Stimulation by cytokines leads to phosphorylation and degradation of IκB and NF-κB turns into its active form. There is only one study from the year 2000, that NKIRAS2 (NFKB Inhibitor Interacting Ras Like 2) was shown to interact with IκB protein and makes its degradation slower, which will also delay NF-κB activation (Fenwick et al. 2000). As the previously mentioned Bcl10 papers, NKIRAS2 may compete for with Bcl10 on this duty. As NKIRAS2 was also a candidate for antiviral pathway of NK cells according to our results there should be some other mechanisms of NKIRAS2, because knockout would increase the NF-κB secretion in the cell which makes it resistant to the lentiviral transduction, which is not true in this case.

IRF2 (Interferon regulatory factor 2) was also estimated as one of important genes of antiviral pathways. IRF2 is a transcription factor and it competitively inhibits IRF1 mediated interferon α and β secretion. In a cancer study, it was demonstrated that IRF1 presence activates NK cells against tumor cells (Ksienzyk et al. 2011). Also another study revealed that IRF2 was required for stimulation of the TLR3 pathway (Ren et al. 2015). In a study with hepatitis C virus infection, RIG-I and IRF-2 expression in B cells were demonstrated to be upregulated, then in vitro studies confirmed IRF2 stimulatory role on RIG-I promoter (Masumi et al. 2010) All these findings may suggest IRF2 dependent activation of RIG-I antiviral pathways can be present in NK cells as well, but it is still need to be clarified how these regulatory factors affect downstream signaling events for innate immune stimulation *in vivo*.

The S100A12 gene encodes a calcium binding protein and has been shown to be present in mostly innate immune cells like neutrophils and macrophages (Jackson et al. 2017). This gene was investigated in bacterial infections (Helicobacter pylori), and its stimulation was shown during infection, in vitro studies verified its inhibitory role to bacterial growth. Also another study on mycobacteria showed that they were capable of killing *Mycobacterium leprae* in infected macrophages and necessary for TLR2/1L and IFN-γ pathway against this intracellular pathogen (Haley et al. 2015). Although it was not reported before, S100A12 can be a candidate gene to antiviral pathways in NK cells as well (Realegeno et al. 2016).

In recent studies on HCV, it was shown that virus was present in lipoviral particles inside the cell which has a highly lipidated layer and this layer is consists of apoB and apoE proteins (Merz et al. 2011). Another study focused on this finding and APOB was silenced in human hepatoma cells via the TALEN gene editing technique. Although there was no change in virus entry and replication, a remarkable decrease in HCV RNA and protein levels were observed after infection. Also, inhibitor drug of the protein, apoB confirmed the affectivity of APOB protein against viral infection in vitro (Schaefer et al. 2016). Presence of APOB on our antiviral candidate list can also indicate that using the same inhibitor drug may increase lentiviral transduction.

SerpineI expression has never been shown in NK cells before, it codes for serine proteinase inhibitor (serpin) superfamily. In a study on Zebrafish it was shown that all the serine protease inhibitors stimulation was increased during viral haemorrhagic septicemia virus (VHSV) infection except Serpine1. In the same study, they demonstrated annic acid, tiplaxtinin, EGCG drugs, which were designed to target Serpine1 gene prevented viral infection not only for this virus but also the others. Although, the mechanism of this Serpine1 dependent protection still should be clarified, this makes the gene a promising candidate antiviral agent as well (Estepa and Coll 2015). SFTPD (Surfactant Protein D) has been shown to be incorporated with PAMPs like bacterial LPS, oligosaccharides and fatty acids. Although these genes are mostly active in type II lung cells, they are also present in epithelial cells as well (Leth-Larsen et al. 2005).

DAP12 is an accessory molecule in NK cells and responsible for initiating signal transduction when the cell is under stress like virus infection or tumor formation (Campbell and Colonna 1999). FGF10, TREM2, UBC, PSMD1 genes were shown to be related with DAP12 signaling and all of them were present on the antiviral pathway list. Although DAP12 can induce cytokine secretion, it also has inhibitory role that leads to competitive inhibition with TLR pathways. TREM2 (Triggering Receptor Expressed on Myeloid cells 2) was found to be subjected to DAP12-mediated inhibition in a study on macrophages. When TREM-2 was bound by its ligand it results in decreased inflammatory response (Hamerman et al. 2006).

Cytokine-mediated signaling pathways also demonstrated an increase via loss of function genomic screen with the genes TREM2, UBC, IRF2, LIMS1, IFNB1, PSMD1 coming up as significant in our analysis. Apart from previously mentioned genes LIMS1 (LIM Zinc Finger Domain Containing 1) also known as PINCH is responsible for cell to cell adhesion and found in most of the cells (Yanwu Yang et al. 2009). Till now, several cytokines have been reported as NK cell cytokines; tumor necrosis factor-α (TNF-α) and interferon γ (IFN-γ) are two prominent ones. Also some other immune regulatory factors were revealed to secreted by NK cells as interleukin family members; IL-5, IL-10, IL-13(Beignon et al. 2005).

On the other hand, GFP+ libraries were ranked according to negatively selected ones and GFP- libraries were ordered according to positively selected; genes that p-values are below 0,05 was searched for common. This list would include proviral pathways that are used or manipulated in the host cell by virus to allow more efficient infectivity or pathogenesis.

On the top of the pro-viral pathway list cell cycle pathways were noticed. Cell cycle arrests were reported in primary HIV-I (Groschel and Bushman 2005; G. Li et al. 2010).Although mechanism of this arrest is not clear still viral accessory proteins were shown to be related with G2/S cell cycle infection (S. Zhang et al. 2006). Previous studies have demonstrated that the stimulation of cells prior to lentiviral genetic modification presents a chance to boost transduction efficiencies (Costello et al., 2000). Although lentiviral vectors do not require active cell division for successful integration of the viral genome, ongoing cellular activity is critical for efficient genetic modification. In a study on T cell transduction with HIV-1 derived lentiviral vector; it was shown that HIV infection on these cells were stuck on reverse transcription level, lentiviral genome was not integrated in G_0 phase so transduction did not completed, on the other hand when the cells were induced with cytokines transferred cells into G_{1b} stage and infection achieved with completion of integration. This study showed that cells that are G2/S phase increased the transduction rate.(Sutton et al. 1999).This observation is well represented in our data which puts forward cell-cycle related pathways as having a pro-viral effect.

Since viruses are obligatory intracellular parasites and could not carry its own translation machinery, they have to use host cells mechanism for survival, spread and evolution. Eukaryotic translation mainly includes following steps; after transcription in the cytoplasm, mature mRNAs moves from nucleus to cytoplasm and 40S small ribosomal subunits looks for start codon to start translation. When the codon is found 60s ribosomal units is engaged and amino acid synthesis is initiated. As using key steps on this mechanism viruses manipulates host cell translational mechanisms. So as expected; host cell translational pathways were increased after viral infection, 60s unit expression, 40s unit expression, also viral mRNA translation was stimulated according to our proviral pathway list as well which was concordant with previous finding (M Gale et al. 2000; Walsh, Mathews, and Mohr 2013; Aranda and Maule 1998).Studies on HIV-1 translational mechanisms have shown that DDX3 alone and also with eIF4G and PABP proteins can bind to 5' of HIV viruses and initiates translation (Guerrero et al. 2015). DDX19B was present in our candidate proviral gene list instead of DDX3 which is another RNA-dependent ATPase which may have a role in nuclear import and export in lentiviral genome translations (Soto-Rifo, Rubilar, and Ohlmann 2013; Rajakylä et al. 2015).

Nonsense-Mediated mRNA Decay is a defense mechanism against viruses as it eliminates default transcripts or virus derived transcripts in the cytoplasm. Parallel with this knowledge, in a recent study with RNA interference, silencing of Upf1, SMG5, and SMG7 Nonsense-Mediated mRNA Decay Pathway related genes inhibited virus replication (Balistreri et al. 2014). However, in our case RPL gene family members (RPL18,10,13,2) which were associated with NMD as well, seemed to have a pro-viral effect. This situation can be due to another role for NMD as increasing the activation threshold of innate signal transduction events and downregulating RLRs on cytoplasm (Eckard et al. 2014; Gloggnitzer et al. 2014; Rigby and Rehwinkel 2015).

As in the list IL-3, IL-5 and IL7 were found to be related as increasing viral delivery before (Williams et al. 1990). Also there are several influenza virus related pathways in the candidate proviral pathway list. Although, they are both different RNA viruses,

probably they share common mechanisms which have not been validated for lentiviruses yet (Michael Gale and Katze 1997).

Nerve growth factor(NGF) presence in the proviral list were conspicuous as well. Although the role of NGF has not been revealed yet, NGF was shown to be induced in monocytes which has TLR dependent activation. By binding to its TrKA (tyrosine kinase) receptor, it prevents inflammatory cytokine production (Prencipe et al. 2014). Besides, secretion of NGF on mouse NK cell were shown in aprevious study (Ralainirina et al. 2010). Thus presence of this signaling event in our candidate NK cell proviral gene list is in line with these findings and probably NGF was another mechanism of viruses which used for manipulation of the NK cell as well.

6. CONCLUSION

This study is an example of CRISPR library Knockout Genomic Screen to reveal NK cell intracellular innate immune pathways against lentiviral vector-mediated gene delivery. As being a fast agent of the innate immune system, characterization NK cell response to viral vector and comparison to the responses observed against wild type viruses will enlighten downstream pathways that interfere with gene therapy protocols. With this study we could show that CRISPR library knockout screens can be used effectively to reveal mechanisms *in vitro* for hematopoietic derived cells even for cells like NK cells with the help of next generation sequencing.

Also we showed, as expected, pattern recognition receptors and pathways are active in NK cells and especially RIG-I/MDA5 mediated pathway was prominently used for virus recognition. Besides, several antiviral candidate genes and pathways are determined which can potentially be used for further research to increase NK cell or any other lymphoid cells gene delivery efficiency. These findings especially increased the hopes for manipulation of NK cells for using them as immunotherapy agents. Furthermore, candidate proviral agents were decided for lentiviral viruses and most of the findings were parallel with the previous knowledge and also with primate viruses. Using this output, it may be possible to modify lentiviral gene delivery processes in order to increase their efficiency of infection.

7. REFERENCES

Ablasser, Andrea, Franz Bauernfeind, Gunther Hartmann, Eicke Latz, Katherine A Fitzgerald, and Veit Hornung. 2009. "RIG-I-Dependent Sensing of poly(dA:dT) through the Induction of an RNA Polymerase III–transcribed RNA Intermediate." *Nature Immunology* 10 (10): 1065–72. doi:10.1038/ni.1779.

Akira, Shizuo, Satoshi Uematsu, and Osamu Takeuchi. 2006. "Pathogen Recognition and Innate Immunity." *Cell* 124 (4): 783–801. doi:10.1016/j.cell.2006.02.015.

Andrews, Simon. 2017. "Babraham Bioinformatics - FastQC A Quality Control Tool for High Throughput Sequence Data." Accessed July 28. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Andzelm, Milena M., Xi Chen, Konrad Krzewski, Jordan S. Orange, and Jack L. Strominger. 2007. "Myosin IIA Is Required for Cytolytic Granule Exocytosis in Human NK Cells." *Journal of Experimental Medicine* 204 (10). http://jem.rupress.org/content/204/10/2285.short.

Aranda, Miguel, and Andrew Maule. 1998. "Virus-Induced Host Gene Shutoff in Animals and Plants." *Virology* 243 (2): 261–67. doi:10.1006/viro.1998.9032.

Balistreri, Giuseppe, Peter Horvath, Christoph Schweingruber, David Zünd, Gerald McInerney, Andres Merits, Oliver Mühlemann, Claus Azzalin, and Ari Helenius. 2014. "The Host Nonsense-Mediated mRNA Decay Pathway Restricts Mammalian RNA Virus Replication." *Cell Host & Microbe* 16 (3): 403–11. doi:10.1016/j.chom.2014.08.007.

Beignon, Anne-Sophie, Kelli McKenna, Mojca Skoberne, Olivier Manches, Ida DaSilva, Daniel G Kavanagh, Marie Larsson, Robert J Gorelick, Jeffrey D Lifson, and Nina Bhardwaj. 2005. "Endocytosis of HIV-1 Activates Plasmacytoid Dendritic Cells via Toll-like Receptor-Viral RNA Interactions." *The Journal of Clinical Investigation* 115 (11). American Society for Clinical Investigation: 3265–75. doi:10.1172/JCI26032.

Bibikova, Marina, Mary Golic, Kent G Golic, and Dana Carroll. 2002. "Targeted Chromosomal Cleavage and Mutagenesis in Drosophila Using Zinc-Finger Nucleases." *Genetics* 161 (3): 1169–75.

http://www.ncbi.nlm.nih.gov/pubmed/12136019.

Bieback, Karen, Egil Lien, Ingo M Klagge, Elita Avota, Jürgen Schneider-Schaulies, W Paul Duprex, Herrmann Wagner, Carsten J Kirschning, Volker Ter Meulen, and Sibylle Schneider-Schaulies. 2002. "Hemagglutinin Protein of Wild-Type Measles Virus Activates Toll-like Receptor 2 Signaling." *Journal of Virology* 76 (17). American Society for Microbiology (ASM): 8729–36. doi:10.1128/jvi.76.17.8729-8736.2002.

Biffi, Alessandra, Cynthia C. Bartolomae, Daniela Cesana, Natalie Cartier, Patrik Aubourg, Marco Ranzani, Martina Cesani, et al. 2011. "Lentiviral Vector Common Integration Sites in Preclinical Models and a Clinical Trial Reflect a Benign Integration Bias and Not Oncogenic Selection." *Blood* 117 (20). http://www.bloodjournal.org/content/117/20/5332.short?sso-checked=true.

Boehme, Karl W, Mario Guerrero, and Teresa Compton. 2006. "Human Cytomegalovirus Envelope Glycoproteins B and H Are Necessary for TLR2 Activation in Permissive Cells." *Journal of Immunology (Baltimore, Md. : 1950)* 177 (10): 7094–7102. http://www.ncbi.nlm.nih.gov/pubmed/17082626.

Boutros, Michael, and Julie Ahringer. 2008. "The Art and Design of Genetic Screens: RNA Interference." *Nature Reviews Genetics* 9 (7). Nature Publishing Group: 554–66. doi:10.1038/nrg2364.

Breckpot, Karine, David Escors, Frederick Arce, Lucienne Lopes, Katarzyna Karwacz, Sandra Van Lint, Marleen Keyaerts, and Mary Collins. 2010. "HIV-1 Lentiviral Vector Immunogenicity Is Mediated by Toll-like Receptor 3 (TLR3) and TLR7." *Journal of Virology* 84 (11): 5627–36. doi:10.1128/JVI.00014-10.

Brown, B. D, G. Sitia, A. Annoni, E. Hauben, L. Sergi Sergi, A. Zingale, M. G. Roncarolo, L. G Guidotti, and L. Naldini. 2006. "In Vivo Administration of Lentiviral Vectors Triggers a Type I Interferon Response That Restricts Hepatocyte Gene Transfer and Promotes Vector Clearance." *Blood* 109 (7): 2797–2805. doi:10.1182/blood-2006-10-049312.

Campbell, Kerry S., and Marco Colonna. 1999. "DAP12: A Key Accessory Protein for Relaying Signals by Natural Killer Cell Receptors." *The International Journal of Biochemistry & Cell Biology* 31 (6): 631–36. doi:10.1016/S1357-2725(99)00022-9.

Capecchi, M R. 1989. "Altering the Genome by Homologous Recombination." *Science (New York, N.Y.)* 244 (4910): 1288–92. http://www.ncbi.nlm.nih.gov/pubmed/2660260.

Chiu, Yu-Hsin, John B. MacMillan, and Zhijian J. Chen. 2009. "RNA Polymerase III Detects Cytosolic DNA and Induces Type I Interferons through the RIG-I Pathway." *Cell* 138 (3): 576–91. doi:10.1016/j.cell.2009.06.015.

Dambuza, Ivy M, and Gordon D Brown. 2015. "C-Type Lectins in Immunity: Recent Developments." *Current Opinion in Immunology* 32 (February). Elsevier: 21–

27. doi:10.1016/j.coi.2014.12.002.

Diebold, S. S., Tsuneyasu Kaisho, Hiroaki Hemmi, Shizuo Akira, and Caetano Reis e Sousa. 2004. “Innate Antiviral Responses by Means of TLR7-Mediated Recognition of Single-Stranded RNA.” *Science* 303 (5663): 1529–31. doi:10.1126/science.1093616.

Diebold, Sandra. 2010. “Innate Recognition of Viruses.” *Immunology Letters* 128 (1): 17–20. doi:10.1016/j.imlet.2009.09.010.

Duperray, Alain, Delphin Barbe, Gilda Raguenez, Babette B. Weksler, Ignacio A. Romero, Pierre-Olivier Couraud, Hervé Perron, and Patrice N. Marche. 2015. “Inflammatory Response of Endothelial Cells to a Human Endogenous Retrovirus Associated with Multiple Sclerosis Is Mediated by TLR4.” *International Immunology* 27 (11): 545–53. doi:10.1093/intimm/dxv025.

Eckard, Sterling C, Gillian I Rice, Alexandre Fabre, Catherine Badens, Elizabeth E Gray, Jane L Hartley, Yanick J Crow, and Daniel B Stetson. 2014. “The SKIV2L RNA Exosome Limits Activation of the RIG-I-like Receptors.” *Nature Immunology* 15 (9). NIH Public Access: 839–45. doi:10.1038/ni.2948.

Estepa, Amparo, and Julio Coll. 2015. “Inhibition of SERPINe1 Reduces Rhabdoviral Infections in Zebrafish.” *Fish & Shellfish Immunology* 47 (1): 264–70. doi:10.1016/j.fsi.2015.09.017.

Fabregat, Antonio, Konstantinos Sidiropoulos, Guilherme Viteri, Oscar Forner, Pablo Marin-Garcia, Vicente Arnau, Peter D’Eustachio, Lincoln Stein, and Henning Hermjakob. 2017. “Reactome Pathway Analysis: A High-Performance in-Memory Approach.” *BMC Bioinformatics* 18 (1): 142. doi:10.1186/s12859-017-1559-2.

Fenwick, Craig, Soon-Young Na, Reinhard E. Voll, Haihong Zhong, Suhn-Young Im, Jae Woon Lee, and Sankar Ghosh. 2000. “A Subclass of Ras Proteins That Regulate the Degradation of IκB.” *Science* 287 (5454). http://icproxy.sabanciuniv.edu:2112/content/287/5454/869.full.

Fernandes-Alnemri, Teresa, Je-Wook Yu, Pinaki Datta, Jianghong Wu, and Emad S Alnemri. 2009. “AIM2 Activates the Inflammasome and Cell Death in Response to Cytoplasmic DNA.” *Nature* 458 (7237). NIH Public Access: 509–13. doi:10.1038/nature07710.

Fire, Andrew, SiQun Xu, Mary K. Montgomery, Steven A. Kostas, Samuel E. Driver, and Craig C. Mello. 1998. “Potent and Specific Genetic Interference by Double-Stranded RNA in Caenorhabditis Elegans.” *Nature* 391 (6669). Nature Publishing Group: 806–11. doi:10.1038/35888.

Franchi, Luigi, Neil Warner, Kyle Viani, and Gabriel Nuñez. 2009. “Function of Nod-like Receptors in Microbial Recognition and Host Defense.” *Immunological Reviews* 227 (1). NIH Public Access: 106–28. doi:10.1111/j.1600-065X.2008.00734.x.

Friedmann, T, and R Roblin. 1972. “Gene Therapy for Human Genetic Disease?” *Science (New York, N.Y.)* 175 (4025): 949–55. http://www.ncbi.nlm.nih.gov/pubmed/5061866.

Gack, Michaela U. 2014. “Mechanisms of RIG-I-like Receptor Activation and Manipulation by Viral Pathogens.” *Journal of Virology* 88 (10). American Society for Microbiology (ASM): 5213–16. doi:10.1128/JVI.03370-13.

Gale, M, S L Tan, M G Katze, and Michael G. Katze. 2000. “Translational Control of Viral Gene Expression in Eukaryotes.” *Microbiology and Molecular Biology Reviews : MMBR* 64 (2). American Society for Microbiology (ASM): 239–80. http://www.ncbi.nlm.nih.gov/pubmed/10839817.

Gale, Michael, and Michael G. Katze. 1997. “What Happens Inside Lentivirus or Influenza Virus Infected Cells: Insights into Regulation of Cellular and Viral Protein Synthesis.” *Methods* 11 (4): 383–401. doi:10.1006/meth.1996.0436.

Gehl, J. 2003. “Electroporation: Theory and Methods, Perspectives for Drug Delivery, Gene Therapy and Research.” *Acta Physiologica Scandinavica* 177 (4). Blackwell Science Ltd: 437–47. doi:10.1046/j.1365-201X.2003.01093.x.

Gloggnitzer, Jiradet, Svetlana Akimcheva, Arunkumar Srinivasan, Branislav Kusenda, Nina Riehs, Hansjörg Stampfl, Jaqueline Bautor, et al. 2014. “Nonsense-Mediated mRNA Decay Modulates Immune Receptor Levels to Regulate Plant Antibacterial Defense.” *Cell Host & Microbe* 16 (3): 376–90. doi:10.1016/j.chom.2014.08.010.

Groschel, Bettina, and Frederic Bushman. 2005. “Cell Cycle Arrest in G2/M Promotes Early Steps of Infection by Human Immunodeficiency Virus.” *Journal of Virology* 79 (9). American Society for Microbiology (ASM): 5695–5704. doi:10.1128/JVI.79.9.5695-5704.2005.

Guan, Y., D. R. E. Ranoa, S. Jiang, S. K. Mutha, X. Li, J. Baudry, and R. I. Tapping. 2010. “Human TLRs 10 and 1 Share Common Mechanisms of Innate Immune Sensing but Not Signaling.” *The Journal of Immunology* 184 (9): 5094–5103. doi:10.4049/jimmunol.0901888.

Guerrero, Santiago, Julien Batisse, Camille Libre, Serena Bernacchi, Roland Marquet, and Jean-Christophe Paillart. 2015. “HIV-1 Replication and the Cellular Eukaryotic Translation Apparatus.” *Viruses* 7 (1). Multidisciplinary Digital Publishing Institute (MDPI): 199–218. doi:10.3390/v7010199.

Haley, Kathryn P., Alberto G. Delgado, M. Blanca Piazuelo, Brittany L. Mortensen, Pelayo Correa, Steven M. Damo, Walter J. Chazin, Eric P. Skaar, and Jennifer A. Gaddy. 2015. “The Human Antimicrobial Protein Calgranulin C Participates in Control of Helicobacter Pylori Growth and Regulation of Virulence.” Edited by S. R. Blanke. *Infection and Immunity* 83 (7): 2944–56. doi:10.1128/IAI.00544-15.

Hamerman, Jessica A., Jessica R. Jarjoura, Mary Beth Humphrey, Mary C. Nakamura, William E. Seaman, and Lewis L. Lanier. 2006. “Cutting Edge:

Inhibition of TLR and FcR Responses in Macrophages by Triggering Receptor Expressed on Myeloid Cells (TREM)-2 and DAP12." *The Journal of Immunology* 177 (4). http://www.jimmunol.org/content/177/4/2051.short.

Heaton, Steven M., Natalie A. Borg, and Vishva M. Dixit. 2016. "Ubiquitin in the Activation and Attenuation of Innate Antiviral Immunity." *Journal of Experimental Medicine* 213 (1). http://jem.rupress.org/content/213/1/1.

Hewson, Christopher a, Alice Jardine, Michael R Edwards, Vasile Laza-stanca, and Sebastian L Johnston. 2005. "Toll-Like Receptor 3 Is Induced by and Mediates Antiviral Activity against Rhinovirus Infection of Human Bronchial Epithelial Cells." *J. Virol.* 79 (19): 12273–79. doi:10.1128/JVI.79.19.12273.

Hsu, Patrick?D., Eric?S. Lander, and Feng Zhang. 2014. "Development and Applications of CRISPR-Cas9 for Genome Engineering." *Cell* 157 (6): 1262–78. doi:10.1016/j.cell.2014.05.010.

Ichinohe, Takeshi, Iris K Pang, and Akiko Iwasaki. 2010. "Influenza Virus Activates Inflammasomes via Its Intracellular M2 Ion Channel." *Nature Immunology* 11 (5): 404–10. doi:10.1038/ni.1861.

Ikegame, Satoshi, Makoto Takeda, Shinji Ohno, Yuichiro Nakatsu, Yoichi Nakanishi, and Yusuke Yanagi. 2010. "Both RIG-I and MDA5 RNA Helicases Contribute to the Induction of Alpha/beta Interferon in Measles Virus-Infected Human Cells." *Journal of Virology* 84 (1). American Society for Microbiology (ASM): 372–79. doi:10.1128/JVI.01690-09.

Ishikawa, Hiroki, and Glen N Barber. 2008. "STING Is an Endoplasmic Reticulum Adaptor That Facilitates Innate Immune Signalling." *Nature* 455 (7213). NIH Public Access: 674–78. doi:10.1038/nature07317.

Ishikawa, Hiroki, Zhe Ma, and Glen N. Barber. 2009. "STING Regulates Intracellular DNA-Mediated, Type I Interferon-Dependent Innate Immunity." *Nature* 461 (7265): 788–92. doi:10.1038/nature08476.

Jackson, Emmanuel, Saffron Little, Dana S. Franklin, Jennifer A. Gaddy, and Steven M. Damo. 2017. "Expression, Purification, and Antimicrobial Activity of S100A12." *Journal of Visualized Experiments*, no. 123(May). doi:10.3791/55557.

Janeway, Charles A, and Ruslan Medzhitov. 2002. "Innate Immune Recognition." *Annu. Rev. Immunol.*, no. 2: 197–216. doi:10.1146/annurev.immunol.20.083001.084359.

Jiang, Xiaomo, Lisa N Kinch, Chad A Brautigam, Xiang Chen, Fenghe Du, Nick V Grishin, Zhijian J Chen, et al. 2012. "Ubiquitin-Induced Oligomerization of the RNA Sensors RIG-I and MDA5 Activates Antiviral Innate Immune Response." *Immunity* 36 (6). Elsevier: 959–73. doi:10.1016/j.immuni.2012.03.022.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier. 2012. "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive

Bacterial Immunity." *Science* 337 (6096): 816–21. doi:10.1126/science.1225829.

Joung, Julia, Silvana Konermann, Jonathan S Gootenberg, Omar O Abudayyeh, Randall J Platt, Mark D Brigham, Neville E Sanjana, and Feng Zhang. 2016. "Protocol: Genome-Scale CRISPR-Cas9 Knockout and Transcriptional Activation Screening." *bioRxiv* 12 (4). Nature Publishing Group: 59626. doi:10.1101/059626.

Kato, Hiroki, Osamu Takeuchi, Eriko Mikamo-Satoh, Reiko Hirai, Tomoji Kawai, Kazufumi Matsushita, Akane Hiiragi, Terence S Dermody, Takashi Fujita, and Shizuo Akira. 2008. "Length-Dependent Recognition of Double-Stranded Ribonucleic Acids by Retinoic Acid-Inducible Gene-I and Melanoma Differentiation-Associated Gene 5." *The Journal of Experimental Medicine* 205 (7). The Rockefeller University Press: 1601–10. doi:10.1084/jem.20080091.

Kato, Hiroki, Osamu Takeuchi, Shintaro Sato, Mitsutoshi Yoneyama, Masahiro Yamamoto, Kosuke Matsui, Satoshi Uematsu, et al. 2006. "Differential Roles of MDA5 and RIG-I Helicases in the Recognition of RNA Viruses" 441 (May): 101–6. doi:10.1038/nature04734.

Kawai, Taro, and Shizuo Akira. 2010. "The Role of Pattern-Recognition Receptors in Innate Immunity: Update on Toll-like Receptors." *Nature Immunology* 11 (5): 373–84. doi:10.1038/ni.1863.

Kim, Hyongbum, and Jin-Soo Kim. 2014. "A Guide to Genome Engineering with Programmable Nucleases." *Nature Reviews Genetics* 15 (5): 321–34. doi:10.1038/nrg3686.

Kim, Y G, J Cha, and S Chandrasegaran. 1996. "Hybrid Restriction Enzymes: Zinc Finger Fusions to Fok I Cleavage Domain." *Proceedings of the National Academy of Sciences of the United States of America* 93 (3). National Academy of Sciences: 1156–60. http://www.ncbi.nlm.nih.gov/pubmed/8577732.

Kolde, Raivo, Sven Laur, Priit Adler, and Jaak Vilo. 2012. "Robust Rank Aggregation for Gene List Integration and Meta-Analysis." *Bioinformatics* 28 (4): 573–80. doi:10.1093/bioinformatics/btr709.

Koyama, Shohei, Ken J. Ishii, Cevayir Coban, and Shizuo Akira. 2008. "Innate Immune Response to Viral Infection." *Cytokine* 43 (3): 336–41. doi:10.1016/j.cyto.2008.07.009.

Krug, Anne, Gary D. Luker, Winfried Barchet, David A. Leib, Shizuo Akira, and Marco Colonna. 2004. "Herpes Simplex Virus Type 1 Activates Murine Natural Interferon-Producing Cells through Toll-like Receptor 9." *Blood* 103 (4). http://www.bloodjournal.org/content/103/4/1433/tab-e-letters.

Ksienzyk, A., B. Neumann, R. Nandakumar, K. Finsterbusch, M. Grashoff, R. Zawatzky, G. Bernhardt, H. Hauser, and A. Kroger. 2011. "IRF-1 Expression Is Essential for Natural Killer Cells to Suppress Metastasis." *Cancer Research* 71 (20): 6410–18. doi:10.1158/0008-5472.CAN-11-1565.

Kurt-jones, Evelyn A, Lana Popova, Laura Kwinn, Lia M Haynes, Les P Jones, Ralph A Tripp, Edward E Walsh, et al. 2000. "Pattern Recognition Receptors TLR4 and CD14 Mediate Response to Respiratory Syncytial Virus" 1 (5).

Lanier, L L, A M Le, C I Civin, M R Loken, and J H Phillips. 1986. "The Relationship of CD16 (Leu-11) and Leu-19 (NKH-1) Antigen Expression on Human Peripheral Blood NK Cells and Cytotoxic T Lymphocytes." *The Journal of Immunology* 136 (12). http://www.jimmunol.org/content/136/12/4480.long.

Leth-Larsen, Rikke, Peter Garred, Henriette Jensenius, Joseph Meschi, Kevan Hartshorn, Jens Madsen, Ida Tornoe, et al. 2005. "A Common Polymorphism in the SFTPD Gene Influences Assembly, Function, and Concentration of Surfactant Protein D." *The Journal of Immunology* 174 (3). http://www.jimmunol.org/content/174/3/1532.short.

Li, Ge, Hyeon U Park, Dong Liang, and Richard Y Zhao. 2010. "Cell Cycle G2/M Arrest through an S Phase-Dependent Mechanism by HIV-1 Viral Protein R." *Retrovirology* 7 (1): 59. doi:10.1186/1742-4690-7-59.

Li, L, L P Wu, and S Chandrasegaran. 1992. "Functional Domains in Fok I Restriction Endonuclease." *Proceedings of the National Academy of Sciences of the United States of America* 89 (10): 4275–79. http://www.ncbi.nlm.nih.gov/pubmed/1584761.

Li, Wei, Han Xu, Tengfei Xiao, Le Cong, Michael I Love, Feng Zhang, Rafael A Irizarry, Jun S Liu, Myles Brown, and X Shirley Liu. 2014. "MAGeCK Enables Robust Identification of Essential Genes from Genome-Scale CRISPR/Cas9 Knockout Screens." *Genome Biology* 15 (12): 554. doi:10.1186/s13059-014-0554-4.

Lobry, Camille, Tatiana Lopez, Alain Israël, and Robert Weil. 2007. "Negative Feedback Loop in T Cell Activation through IkappaB Kinase-Induced Phosphorylation and Degradation of Bcl10." *Proceedings of the National Academy of Sciences of the United States of America* 104 (3): 908–13. doi:10.1073/pnas.0606982104.

Lozach, Pierre-Yves, Laura Burleigh, Isabelle Staropoli, and Ali Amara. 2007. "The C Type Lectins DC-SIGN and L-SIGN." In *Methods in Molecular Biology (Clifton, N.J.)*, 379:51–68. doi:10.1007/978-1-59745-393-6_4.

Lucas, Peter C., Masakatsu Yonezumi, Naohiro Inohara, Linda M. McAllister-Lucas, Mohamed E. Abazeed, Felicia F. Chen, Shoji Yamaoka, Masao Seto, and Gabriel Núñez. 2001. "Bcl10 and MALT1, Independent Targets of Chromosomal Translocation in MALT Lymphoma, Cooperate in a Novel NF-κB Signaling Pathway." *Journal of Biological Chemistry* 276 (22): 19012–19. doi:10.1074/jbc.M009984200.

Lund, Jennifer, Ayuko Sato, Shizuo Akira, Ruslan Medzhitov, and Akiko Iwasaki. 2003. "Toll-like Receptor 9-Mediated Recognition of Herpes Simplex Virus-2 by Plasmacytoid Dendritic Cells." *The Journal of Experimental Medicine* 198

(3). The Rockefeller University Press: 513–20. doi:10.1084/jem.20030162.

Makarova, Kira S, Daniel H Haft, Rodolphe Barrangou, Stan J J Brouns, Emmanuelle Charpentier, Philippe Horvath, Sylvain Moineau, et al. 2011. "Evolution and Classification of the CRISPR-Cas Systems." *Nature Reviews. Microbiology* 9 (6). NIH Public Access: 467–77. doi:10.1038/nrmicro2577.

Martinez, Jennifer, Xiaopei Huang, and Yiping Yang. 2010. "Direct TLR2 Signaling Is Critical for NK Cell Activation and Function in Response to Vaccinia Viral Infection." *PLoS Pathogens* 6 (3). Public Library of Science: e1000811. doi:10.1371/journal.ppat.1000811.

Marzi, A., T. Gramberg, G. Simmons, P. Moller, A. J. Rennekamp, M. Krumbiegel, M. Geier, et al. 2004. "DC-SIGN and DC-SIGNR Interact with the Glycoprotein of Marburg Virus and the S Protein of Severe Acute Respiratory Syndrome Coronavirus." *Journal of Virology* 78 (21): 12090–95. doi:10.1128/JVI.78.21.12090-12095.2004.

Masumi, Atsuko, Masahiko Ito, Keiko Mochida, Isao Hamaguchi, Takuo Mizukami, Haruka Momose, Madoka Kuramitsu, et al. 2010. "Enhanced RIG-I Expression Is Mediated by Interferon Regulatory Factor-2 in Peripheral Blood B Cells from Hepatitis C Virus-Infected Patients." *Biochemical and Biophysical Research Communications* 391 (4): 1623–28. doi:10.1016/j.bbrc.2009.12.092.

Merz, Andreas, Gang Long, Marie-Sophie Hiet, Britta Brügger, Petr Chlanda, Patrice Andre, Felix Wieland, Jacomine Krijnse-Locker, and Ralf Bartenschlager. 2011. "Biochemical and Morphological Properties of Hepatitis C Virus Particles and Determination of Their Lipidome." *Journal of Biological Chemistry* 286 (4): 3018–32. doi:10.1074/jbc.M110.175018.

Monteiro, João T, and Bernd Lepenies. 2017. "Myeloid C-Type Lectin Receptors in Viral Recognition and Antiviral Immunity." *Viruses* 9 (3). Multidisciplinary Digital Publishing Institute (MDPI). doi:10.3390/v9030059.

Mulligan, Richard C. 1993. "The Basic Science of Gene Therapy." *Science*. American Association for the Advancement of Science. doi:10.1126/science.8493530.

Naldini, Luigi, Ulrike Blömer, Philippe Gallay, Daniel Ory, Richard Mulligan, Fred H. Gage, Inder M. Verma, and Didier Trono. 1998. "In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector." *Science* 272. American Association for the Advancement of Science: 263–67. doi:10.2307/2889637.

Naldini, Luigi, Didier Trono, and Inder M. Verma. 2016. "Lentiviral Vectors, Two Decades Later." *Science* 353 (6304). http://science.sciencemag.org/content/353/6304/1101.long.

Ozinsky, A., D. M. Underhill, J. D. Fontenot, A. M. Hajjar, K. D. Smith, C. B. Wilson, L. Schroeder, and A. Aderem. 2000. "The Repertoire for Pattern Recognition of Pathogens by the Innate Immune System Is Defined by

Cooperation between Toll-like Receptors." *Proceedings of the National Academy of Sciences* 97 (25): 13766–71. doi:10.1073/pnas.250476497.

Pavletich, N P, and C O Pabo. 1991. "Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at 2.1 A." *Science (New York, N.Y.)* 252 (5007): 809–17. http://www.ncbi.nlm.nih.gov/pubmed/2028256.

Pichlmair, Andreas, Oliver Schulz, Choon-Ping Tan, Jan Rehwinkel, Hiroki Kato, Osamu Takeuchi, Shizuo Akira, Michael Way, Giampietro Schiavo, and Caetano Reis e Sousa. 2009. "Activation of MDA5 Requires Higher-Order RNA Structures Generated during Virus Infection." *Journal of Virology* 83 (20). American Society for Microbiology (ASM): 10761–69. doi:10.1128/JVI.00770-09.

Pichlmair, Andreas, Oliver Schulz, Choon Ping Tan, Tanja I. Näslund, Peter Liljeström, Friedemann Weber, and Caetano Reis e Sousa. 2006. "RIG-I-Mediated Antiviral Responses to Single-Stranded RNA Bearing 5'-Phosphates." *Science* 314 (5801). http://science.sciencemag.org/content/314/5801/997.full.

Pichlmair, Andreas, and Caetano Reis Sousa. 2007. "Review Innate Recognition of Viruses." *Immunity* 27 (3): 370–83. doi:10.1016/j.immuni.2007.08.012.

Pisegna, Simona, Gianluca Pirozzi, Mario Piccoli, Luigi Frati, Angela Santoni, and Gabriella Palmieri. 2017. "p38 MAPK Activation Controls the TLR3-Mediated up-Regulation of Cytotoxicity and Cytokine Production in Human NK Cells." Accessed July 17. doi:10.1182/blood-2004-05-1860.

Prencipe, G., G. Minnone, R. Strippoli, L. De Pasquale, S. Petrini, I. Caiello, L. Manni, F. De Benedetti, and L. Bracci-Laudiero. 2014. "Nerve Growth Factor Downregulates Inflammatory Response in Human Monocytes through TrkA." *The Journal of Immunology* 192 (7): 3345–54. doi:10.4049/jimmunol.1300825.

Punt, Jenni, Judith Owen, and M. A. Caligiuri. 2001. "The Biology of Human Natural Killer-Cell Subsets." *Trends in Immunology* 22 (11): 633–40. doi:10.1016/S1471-4906(01)02060-9.

Rajakylä, Eeva Kaisa, Tiina Viita, Salla Kyheröinen, Guillaume Huet, Richard Treisman, and Maria K. Vartiainen. 2015. "RNA Export Factor Ddx19 Is Required for Nuclear Import of the SRF Coactivator MKL1." *Nature Communications* 6 (January): 5978. doi:10.1038/ncomms6978.

Ralainirina, Natacha, Nicolaas H. C. Brons, Wim Ammerlaan, Céline Hoffmann, François Hentges, and Jacques Zimmer. 2010. "Mouse Natural Killer (NK) Cells Express the Nerve Growth Factor Receptor TrkA, Which Is Dynamically Regulated." Edited by Ben C. B. Ko. *PLoS ONE* 5 (12): e15053. doi:10.1371/journal.pone.0015053.

Rathinam, Vijay A K, and Katherine A Fitzgerald. 2011a. "Innate Immune Sensing of DNA Viruses." *Virology* 411 (2). NIH Public Access: 153–62. doi:10.1016/j.virol.2011.02.003.

———. 2011b. “Innate Immune Sensing of DNA Viruses.” *Virology* 411 (2). NIH Public Access: 153–62. doi:10.1016/j.virol.2011.02.003.

Rathinam, Vijay A K, Zhaozhao Jiang, Stephen N Waggoner, Shruti Sharma, Leah E Cole, Lisa Waggoner, Sivapriya Kailasan Vanaja, et al. 2010. “The AIM2 Inflammasome Is Essential for Host Defense against Cytosolic Bacteria and DNA Viruses.” *Nature Immunology* 11 (5). NIH Public Access: 395–402. doi:10.1038/ni.1864.

Realegeno, Susan, Kindra M. Kelly-Scumpia, Angeline Tilly Dang, Jing Lu, Rosane Teles, Philip T. Liu, Mirjam Schenk, et al. 2016. “S100A12 Is Part of the Antimicrobial Network against Mycobacterium Leprae in Human Macrophages.” Edited by Padmini Salgame. *PLOS Pathogens* 12 (6). Public Library of Science: e1005705. doi:10.1371/journal.ppat.1005705.

Rehwinkel, Jan, Choon Ping Tan, Delphine Goubau, Oliver Schulz, Andreas Pichlmair, Katja Bier, Nicole Robb, et al. 2010. “RIG-I Detects Viral Genomic RNA during Negative-Strand RNA Virus Infection.” *Cell* 140 (3): 397–408. doi:10.1016/j.cell.2010.01.020.

Reimand, Jüri, Tambet Arak, Priit Adler, Liis Kolberg, Sulev Reisberg, Hedi Peterson, and Jaak Vilo. 2016. “g:Profiler—a Web Server for Functional Interpretation of Gene Lists (2016 Update).” *Nucleic Acids Research* 44 (W1): W83–89. doi:10.1093/nar/gkw199.

Ren, Gang, Kairong Cui, Zhiying Zhang, and Keji Zhao. 2015. “Division of Labor between IRF1 and IRF2 in Regulating Different Stages of Transcriptional Activation in Cellular Antiviral Activities.” *Cell & Bioscience* 5 (1): 17. doi:10.1186/s13578-015-0007-0.

Rigby, Rachel E., and Jan Rehwinkel. 2015. “RNA Degradation in Antiviral Immunity and Autoimmunity.” *Trends in Immunology* 36 (3): 179–88. doi:10.1016/j.it.2015.02.001.

Robertson, MJ, and J Ritz. 1990. “Biology and Clinical Relevance of Human Natural Killer Cells.” *Blood* 76 (12). http://www.bloodjournal.org/content/76/12/2421.long?sso-checked=true.

Rossetti, Maura, Silvia Gregori, Ehud Hauben, Brian D. Brown, Lucia Sergi Sergi, Luigi Naldini, and Maria-Grazia Roncarolo. 2011. “HIV-1-Derived Lentiviral Vectors Directly Activate Plasmacytoid Dendritic Cells, Which in Turn Induce the Maturation of Myeloid Dendritic Cells.” *Human Gene Therapy* 22 (2): 177–88. doi:10.1089/hum.2010.085.

Ruland, Jürgen, Gordon S Duncan, Andrew Elia, Ivan del Barco Barrantes, Linh Nguyen, Sue Plyte, Douglas G Millar, et al. 2001. “Bcl10 Is a Positive Regulator of Antigen Receptor–Induced Activation of NF-κ B and Neural Tube Closure.” *Cell* 104 (1): 33–42. doi:10.1016/S0092-8674(01)00189-1.

Sabbah, Ahmed, Te Hung Chang, Rosalinda Harnack, Victoria Frohlich, Kaoru Tominaga, Peter H Dube, Yan Xiang, and Santanu Bose. 2009. “Activation of

Innate Immune Antiviral Responses by Nod2." *Nature Immunology* 10 (10). Nature Publishing Group: 1073–80. doi:10.1038/ni.1782.

Sanborn, Keri B., Gregory D. Rak, Saumya Y. Maru, Korey Demers, Analisa Difeo, John A. Martignetti, Michael R. Betts, Rémi Favier, Pinaki P. Banerjee, and Jordan S. Orange. 2009. "Myosin IIA Associates with NK Cell Lytic Granules to Enable Their Interaction with F-Actin and Function at the Immunological Synapse." *The Journal of Immunology* 182 (11). http://www.jimmunol.org/content/182/11/6969.short.

Sancar, Aziz. 1996. "DNA Excision Repair." *Annual Review of Biochemistry* 65 (1). Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA : 43–81. doi:10.1146/annurev.bi.65.070196.000355.

Sanjana, Neville E. 2016. "Genome-Scale CRISPR Pooled Screens." *Analytical Biochemistry*, June. doi:10.1016/j.ab.2016.05.014.

Sanjana, Neville E, Ophir Shalem, and Feng Zhang. 2014. "Improved Vectors and Genome-Wide Libraries for CRISPR Screening." *Nature Methods* 11 (8). NIH Public Access: 783–84. doi:10.1038/nmeth.3047.

Sasai, Miwa, and Masahiro Yamamoto. 2013. "Pathogen Recognition Receptors: Ligands and Signaling Pathways by Toll-like Receptors." *International Reviews of Immunology* 32 (2): 116–33. doi:10.3109/08830185.2013.774391.

Sayitoglu, Canan. 2017. "Intracellular Immunodynamics Of Lentiviral Gene Delivery In Human Natural Killer Cells." Sabanci University.

Schaefer, Esperance A K, James Meixiong, Christina Mark, Amy Deik, Daniel L Motola, Dahlene Fusco, Andrew Yang, et al. 2016. "Apolipoprotein B100 Is Required for Hepatitis C Infectivity and Mipomersen Inhibits Hepatitis C." *World Journal of Gastroenterology* 22 (45). Baishideng Publishing Group Inc: 9954–65. doi:10.3748/wjg.v22.i45.9954.

Segovia, Jesus, Ahmed Sabbah, Victoria Mgbemena, Su-Yu Tsai, Te-Hung Chang, Michael T Berton, Ian R Morris, Irving C Allen, Jenny P-Y Ting, and Santanu Bose. 2012. "TLR2/MyD88/NF-κB Pathway, Reactive Oxygen Species, Potassium Efflux Activates NLRP3/ASC Inflammasome during Respiratory Syncytial Virus Infection." *PloS One* 7 (1). Public Library of Science: e29695. doi:10.1371/journal.pone.0029695.

Shalem, Ophir, Neville E. Sanjana, Ella Hartenian, Xi Shi, David A. Scott, Tarjei S. Mikkelsen, Dirk Heckl, et al. 2014a. "Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells." *Science* 343 (6166). NIH Public Access: 84–87. doi:10.1126/science.1247005.

Shalem, Ophir, Neville E Sanjana, Ella Hartenian, Xi Shi, David A Scott, Tarjei S Mikkelsen, Dirk Heckl, et al. 2014b. "Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells." *Science (New York, N.Y.)* 343 (6166). NIH Public Access: 84–87. doi:10.1126/science.1247005.

Soto-Rifo, R., P. S. Rubilar, and T. Ohlmann. 2013. “The DEAD-Box Helicase DDX3 Substitutes for the Cap-Binding Protein eIF4E to Promote Compartmentalized Translation Initiation of the HIV-1 Genomic RNA.” *Nucleic Acids Research* 41 (12): 6286–99. doi:10.1093/nar/gkt306.

Sundaram, Meera V. 2005. “The Love-Hate Relationship between Ras and Notch.” *Genes & Development* 19 (16). Cold Spring Harbor Laboratory Press: 1825–39. doi:10.1101/gad.1330605.

Sutlu, Tolga, Sanna Nyström, Mari Gilljam, Birgitta Stellan, Steven E Applequist, and Evren Alici. 2012. “Inhibition of Intracellular Antiviral Defense Mechanisms Augments Lentiviral Transduction of Human Natural Killer Cells: Implications for Gene Therapy.” *Human Gene Therapy* 23 (10): 1090–1100. doi:10.1089/hum.2012.080.

Sutton, R E, M J Reitsma, N Uchida, and P O Brown. 1999. “Transduction of Human Progenitor Hematopoietic Stem Cells by Human Immunodeficiency Virus Type 1-Based Vectors Is Cell Cycle Dependent.” *Journal of Virology* 73 (5). American Society for Microbiology: 3649–60. http://www.ncbi.nlm.nih.gov/pubmed/10196257.

Tabeta, Koichi, Kasper Hoebe, Edith M Janssen, Xin Du, Philippe Georgel, Karine Crozat, Suzanne Mudd, et al. 2006. “The Unc93b1 Mutation 3d Disrupts Exogenous Antigen Presentation and Signaling via Toll-like Receptors 3, 7 and 9.” *Nature Immunology* 7 (2). Nature Publishing Group: 156–64. doi:10.1038/ni1297.

Takaoka, Akinori, ZhiChao Wang, Myoung Kwon Choi, Hideyuki Yanai, Hideo Negishi, Tatsuma Ban, Yan Lu, et al. 2007. “DAI (DLM-1/ZBP1) Is a Cytosolic DNA Sensor and an Activator of Innate Immune Response.” *Nature* 448 (7152). Nature Publishing Group: 501–5. doi:10.1038/nature06013.

Takeuchi, Osamu, and Shizuo Akira. 2008. “MDA5/RIG-I and Virus Recognition.” *Current Opinion in Immunology* 20 (1): 17–22. doi:10.1016/j.coi.2008.01.002.

———. 2010. “Pattern Recognition Receptors and Inflammation.” *Cell* 140 (6): 805–20. doi:10.1016/j.cell.2010.01.022.

Tassaneetrithep, Boonrat, Timothy H. Burgess, Angela Granelli-Piperno, Christine Trumpfheller, Jennifer Finke, Wellington Sun, Michael A. Eller, et al. 2003. “DC-SIGN (CD209) Mediates Dengue Virus Infection of Human Dendritic Cells.” *The Journal of Experimental Medicine* 197 (7): 823–29. doi:10.1084/jem.20021840.

Thomas, Clare E., Anja Ehrhardt, and Mark A. Kay. 2003. “Progress and Problems with the Use of Viral Vectors for Gene Therapy.” *Nature Reviews Genetics* 4 (5): 346–58. doi:10.1038/nrg1066.

Unterholzner, Leonie, Sinead E Keating, Marcin Baran, Kristy A Horan, Søren B Jensen, Shruti Sharma, Cherilyn M Sirois, et al. 2010. “IFI16 Is an Innate Immune Sensor for Intracellular DNA.” *Nature Immunology* 11 (11). NIH

Public Access: 997–1004. doi:10.1038/ni.1932.

Verma, I. M., L. Naldini, T. Kafri, H. Miyoshi, M. Takahashi, U. Blömer, N. Somia, L. Wang, and F. H. Gage. 2000. "Gene Therapy: Promises, Problems and Prospects." In *Genes and Resistance to Disease*, 147–57. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-56947-0_13.

Walsh, Derek, Michael B Mathews, and Ian Mohr. 2013. "Tinkering with Translation: Protein Synthesis in Virus-Infected Cells." *Cold Spring Harbor Perspectives in Biology* 5 (1). Cold Spring Harbor Laboratory Press: a012351. doi:10.1101/cshperspect.a012351.

Wang, Tian, Terrence Town, Lena Alexopoulou, John F Anderson, Erol Fikrig, and Richard A Flavell. 2004. "Toll-like Receptor 3 Mediates West Nile Virus Entry into the Brain Causing Lethal Encephalitis." *Nature Medicine* 10 (12): 1366–73. doi:10.1038/nm1140.

Wang, Tim, Jenny J. Wei, David M. Sabatini, Eric S. Lander, and Eric S. Lander Tim Wang, Jenny J. Wei, David M. Sabatini. 2014. "Genetic Screens in Human Cells Using the CRISPR-Cas9 System." *Science* 80 (January): 80–85. doi:10.1126/science.1246981.

Weber, Friedemann, Valentina Wagner, Simon B Rasmussen, Rune Hartmann, and Søren R Paludan. 2006. "Double-Stranded RNA Is Produced by Positive-Strand RNA Viruses and DNA Viruses but Not in Detectable Amounts by Negative-Strand RNA Viruses." *Journal of Virology* 80 (10). American Society for Microbiology: 5059–64. doi:10.1128/JVI.80.10.5059-5064.2006.

Williams, D E, A E Namen, D Y Mochizuki, and R W Overell. 1990. "Clonal Growth of Murine Pre-B Colony-Forming Cells and Their Targeted Infection by a Retroviral Vector: Dependence on Interleukin-7." *Blood* 75 (5): 1132–38. http://www.ncbi.nlm.nih.gov/pubmed/2407299.

Wilson, Ross C., and Jennifer A. Doudna. 2013. "Molecular Mechanisms of RNA Interference." *Annual Review of Biophysics* 42 (1). Annual Reviews : 217–39. doi:10.1146/annurev-biophys-083012-130404.

Yang, Yanwu, Xiaoxia Wang, Cheryl A Hawkins, Kan Chen, Julia Vaynberg, Xian Mao, Yizeng Tu, et al. 2009. "Structural Basis of Focal Adhesion Localization of LIM-Only Adaptor PINCH by Integrin-Linked Kinase." *The Journal of Biological Chemistry* 284 (9): 5836–44. doi:10.1074/jbc.M805319200.

Yang, Yong, Mingwei Tong, Li Yi, Yuening Cheng, Miao Zhang, Zhigang Cao, Jianke Wang, Peng Lin, and Shipeng Cheng. 2016. "Identification and Characterization of the Toll-like Receptor 8 Gene in the Chinese Raccoon Dog (Nyctereutes Procyonoides)." *Immunology Letters* 178 (October): 50–60. doi:10.1016/j.imlet.2016.07.012.

Yoneyama, Mitsutoshi, Mika Kikuchi, Takashi Natsukawa, Noriaki Shinobu, Tadaatsu Imaizumi, Makoto Miyagishi, Kazunari Taira, Shizuo Akira, and Takashi Fujita. 2004. "The RNA Helicase RIG-I Has an Essential Function in

Double-Stranded RNA-Induced Innate Antiviral Responses." *Nature Immunology* 5 (7): 730–37. doi:10.1038/ni1087.

Zhang, F., Y. Wen, and X. Guo. 2014. "CRISPR/Cas9 for Genome Editing: Progress, Implications and Challenges." *Human Molecular Genetics* 23 (R1). Oxford University Press: R40–46. doi:10.1093/hmg/ddu125.

Zhang, Shangming, Guiandre Joseph, Karen Pollok, Lionel Berthoux, Lakshmi Sastry, Jeremy Luban, and Kenneth Cornetta. 2006. "G2 Cell Cycle Arrest and Cyclophilin A in Lentiviral Gene Transfer." *Molecular Therapy : The Journal of the American Society of Gene Therapy* 14 (4). Elsevier: 546–54. doi:10.1016/j.ymthe.2006.05.022.

Zhou, Honglin, Ingrid Wertz, Mark Ultsch, Somasekar Seshagiri, Michael Eby, Wei Xiao, and Vishva M Dixit. 2003. "Bcl10 Activates the NF-kB Pthway through Ubiquitination of NEMO." *Nature*, 167–71. doi:10.1038/nature02269.1.

Zhou, Kai, Jing Wang, An Li, Wenming Zhao, Dongfang Wang, Wei Zhang, Jinghua Yan, George Fu Gao, Wenjun Liu, and Min Fang. 2016. "Swift and Strong NK Cell Responses Protect 129 Mice against High-Dose Influenza Virus Infection." *The Journal of Immunology* 196 (4): 1842–54. doi:10.4049/jimmunol.1501486.

Zufferey, R, T Dull, R J Mandel, A Bukovsky, D Quiroz, L Naldini, and D Trono. 1998. "Self-Inactivating Lentivirus Vector for Safe and Efficient in Vivo Gene Delivery." *Journal of Virology* 72 (12): 9873–80. http://www.ncbi.nlm.nih.gov/pubmed/9811723.

APPENDIX A : g:profiler Antiviral Gene Enrichment Map Output:

Rank	Reactome ID	p-value	Pathway Name	Gene List
1	REAC:168928	0,000000178	RIG-I/MDA5 mediated induction of IFN-alpha/beta pathways	UBC,S100A12,NKIRAS2,IRF2,IFNB1
2	REAC:5668914	0,0000524	Diseases of metabolism	PPP1R3C,SFTPD,UBC
3	REAC:168898	0,0000934	Toll-Like Receptors Cascades	APOB,UBC,S100A12,NKIRAS2
4	REAC:68827	0,000125	CDT1 association with the CDC6:ORC:origin complex	ORC2,UBC,PSMD1
5	REAC:5663084	0,000128	Diseases of carbohydrate metabolism	PPP1R3C,UBC
6	REAC:3229121	0,000128	Glycogen storage diseases	PPP1R3C,UBC
7	REAC:3785653	0,000128	Myoclonic epilepsy of Lafora	PPP1R3C,UBC
8	REAC:1169091	0,000181	Activation of NF-kappaB in B cells	BCL10,UBC,PSMD1
9	REAC:68867	0,000189	Assembly of the pre-replicative complex	ORC2,UBC,PSMD1
10	REAC:5619084	0,000189	ABC transporter disorders	SFTPD,UBC,PSMD1
11	REAC:1834949	0,000197	Cytosolic sensors of pathogen-associated DNA	UBC,S100A12,NKIRAS2
12	REAC:69275	0,000199	G2/M Transition	CEP250,UBC,PCNT,PSMD1
13	REAC:453274	0,000208	Mitotic G2-G2/M phases	CEP250,UBC,PCNT,PSMD1
14	REAC:69052	0,000215	Switching of origins to a post-replicative state	ORC2,UBC,PSMD1
15	REAC:68949	0,000215	Orc1 removal from chromatin	ORC2,UBC,PSMD1
16	REAC:69300	0,000233	Removal of licensing factors from origins	ORC2,UBC,PSMD1
17	REAC:5619115	0,000252	Disorders of transmembrane transporters	SFTPD,UBC,PSMD1

18	REAC:69304	0,000262	Regulation of DNA replication	ORC2,UBC,PSMD1
19	REAC:1168372	0,000279	Downstream signaling events of B Cell Receptor (BCR)	FGF10,BCL10,UBC,PSMD1
20	REAC:69298	0,000298	Association of licensing factors with the pre-replicative complex	ORC2,UBC
21	REAC:983712	0,000312	Ion channel transport	CLCA4,BEST3,UBC,HTR3A
22	REAC:3322077	0,000341	Glycogen synthesis	PPP1R3C,UBC
23	REAC:168142	0,000351	Toll Like Receptor 10 (TLR10) Cascade	UBC,S100A12,NKIRAS2
24	REAC:168176	0,000351	Toll Like Receptor 5 (TLR5) Cascade	UBC,S100A12,NKIRAS2
25	REAC:975871	0,000351	MyD88 cascade initiated on plasma membrane	UBC,S100A12,NKIRAS2
26	REAC:68874	0,000363	M/G1 Transition	ORC2,UBC,PSMD1
27	REAC:69002	0,000363	DNA Replication Pre-Initiation	ORC2,UBC,PSMD1
28	REAC:2565942	0,000389	Regulation of PLK1 Activity at G2/M Transition	CEP250,UBC,PCNT
29	REAC:975138	0,000443	TRAF6 mediated induction of NFkB and MAP kinases upon TLR7/8 or 9 activation	UBC,S100A12,NKIRAS2
30	REAC:168181	0,000472	Toll Like Receptor 7/8 (TLR7/8) Cascade	UBC,S100A12,NKIRAS2
31	REAC:975155	0,000472	MyD88 dependent cascade initiated on endosome	UBC,S100A12,NKIRAS2
32	REAC:181438	0,000487	Toll Like Receptor 2 (TLR2) Cascade	UBC,S100A12,NKIRAS2
33	REAC:168188	0,000487	Toll Like Receptor TLR6:TLR2 Cascade	UBC,S100A12,NKIRAS2
34	REAC:168179	0,000487	Toll Like Receptor TLR1:TLR2 Cascade	UBC,S100A12,NKIRAS2
35	REAC:166058	0,000487	MyD88:Mal cascade initiated on plasma membrane	UBC,S100A12,NKIRAS2
36	REAC:202424	0,000518	Downstream TCR signaling	BCL10,UBC,PSMD1

37	REAC:168138	0,000533	Toll Like Receptor 9 (TLR9) Cascade	UBC,S100A12, NKIRAS2
38	REAC:168164	0,000566	Toll Like Receptor 3 (TLR3) Cascade	UBC,S100A12, NKIRAS2
39	REAC:166166	0,000566	MyD88-independent TLR3/TLR4 cascade	UBC,S100A12, NKIRAS2
40	REAC:937061	0,000566	TRIF-mediated TLR3/TLR4 signaling	UBC,S100A12, NKIRAS2
41	REAC:69239	0,000583	Synthesis of DNA	ORC2,UBC,PSMD1
42	REAC:5607764	0,000617	CLEC7A (Dectin-1) signaling	BCL10,UBC,PSMD1
43	REAC:2672351	0,000653	Stimuli-sensing channels	CLCA4,BEST3, UBC
44	REAC:1810476	0,000714	RIP-mediated NFkB activation via ZBP1	S100A12,NKIRAS2
45	REAC:69306	0,000728	DNA Replication	ORC2,UBC,PSMD1
46	REAC:166054	0,000851	Activated TLR4 signalling	UBC,S100A12, NKIRAS2
47	REAC:933542	0,000914	TRAF6 mediated NF-kB activation	S100A12,NKIRAS2
48	REAC:202403	0,000918	TCR signaling	BCL10,UBC,PSMD1
49	REAC:1606322	0,000986	ZBP1(DAI) mediated induction of type I IFNs	S100A12,NKIRAS2
50	REAC:166016	0,00111	Toll Like Receptor 4 (TLR4) Cascade	UBC,S100A12, NKIRAS2
51	REAC:983705	0,00115	Signaling by the B Cell Receptor (BCR)	FGF10,BCL10, UBC,PSMD1
52	REAC:69206	0,00119	G1/S Transition	ORC2,UBC,PSMD1
53	REAC:445989	0,00122	TAK1 activates NFkB by phosphorylation and activation of IKKs complex	S100A12,NKIRAS2
54	REAC:2173796	0,00139	SMAD2/SMAD3:SMAD4 heterotrimer regulates transcription	SERPINE1,UBC
55	REAC:933541	0,00148	TRAF6 mediated IRF7 activation	IRF2,IFNB1
56	REAC:5654726	0,00148	Negative regulation of FGFR1 signaling	FGF10,UBC
57	REAC:69242	0,00151	S Phase	ORC2,UBC,PSMD1

58	REAC:5654727	0,00157	Negative regulation of FGFR2 signaling	FGF10,UBC
59	REAC:5621481	0,00167	C-type lectin receptors (CLRs)	BCL10,UBC,PSMD1
60	REAC:453279	0,00173	Mitotic G1-G1/S phases	ORC2,UBC,PSMD1
61	REAC:2871837	0,00198	FCERI mediated NF-kB activation	BCL10,UBC,PSMD1
62	REAC:2424491	0,00238	DAP12 signaling	FGF10,TREM2,UBC,PSMD1
63	REAC:2467813	0,00246	Separation of Sister Chromatids	UBC,MIS12,PSMD1
64	REAC:2173793	0,00262	Transcriptional activity of SMAD2/SMAD3:SMAD4 heterotrimer	SERPINE1,UBC
65	REAC:69481	0,00263	G2/M Checkpoints	ORC2,UBC,PSMD1
66	REAC:2172127	0,00268	DAP12 interactions	FGF10,TREM2,UBC,PSMD1
67	REAC:68882	0,00295	Mitotic Anaphase	UBC,MIS12,PSMD1
68	REAC:2555396	0,003	Mitotic Metaphase and Anaphase	UBC,MIS12,PSMD1
69	REAC:211733	0,00351	Regulation of activated PAK-2p34 by proteasome mediated degradation	UBC,PSMD1
70	REAC:75815	0,00364	Ubiquitin-dependent degradation of Cyclin D	UBC,PSMD1
71	REAC:69229	0,00364	Ubiquitin-dependent degradation of Cyclin D1	UBC,PSMD1
72	REAC:69017	0,00364	CDK-mediated phosphorylation and removal of Cdc6	UBC,PSMD1
73	REAC:349425	0,00378	Autodegradation of the E3 ubiquitin ligase COP1	UBC,PSMD1
74	REAC:180534	0,00378	Vpu mediated degradation of CD4	UBC,PSMD1
75	REAC:69613	0,00393	p53-Independent G1/S DNA damage checkpoint	UBC,PSMD1
76	REAC:69610	0,00393	p53-Independent DNA Damage Response	UBC,PSMD1

77	REAC:69601	0,00393	Ubiquitin Mediated Degradation of Phosphorylated Cdc25A	UBC,PSMD1
78	REAC:913531	0,00399	Interferon Signaling	UBC,IRF2,IFNB1
79	REAC:3000178	0,00407	ECM proteoglycans	SERPINE1,DAG1
80	REAC:169911	0,00407	Regulation of Apoptosis	UBC,PSMD1
81	REAC:180585	0,00407	Vif-mediated degradation of APOBEC3G	UBC,PSMD1
82	REAC:174113	0,00422	SCF-beta-TrCP mediated degradation of Emi1	UBC,PSMD1
83	REAC:4641257	0,00422	Degradation of AXIN	UBC,PSMD1
84	REAC:450408	0,00422	AUF1 (hnRNP D0) binds and destabilizes mRNA	UBC,PSMD1
85	REAC:5362768	0,00437	Hh mutants that don't undergo autocatalytic processing are degraded by ERAD	UBC,PSMD1
86	REAC:69620	0,0044	Cell Cycle Checkpoints	ORC2,UBC,PSMD1
87	REAC:4641258	0,00452	Degradation of DVL	UBC,PSMD1
88	REAC:69541	0,00452	Stabilization of p53	UBC,PSMD1
89	REAC:5387390	0,00467	Hh mutants abrogate ligand secretion	UBC,PSMD1
90	REAC:5676590	0,00483	NIK-->noncanonical NF-kB signaling	UBC,PSMD1
91	REAC:187577	0,00499	SCF(Skp2)-mediated degradation of p27/p21	UBC,PSMD1
92	REAC:8852276	0,00499	The role of GTSE1 in G2/M progression after G2 checkpoint	UBC,PSMD1
93	REAC:5610785	0,00499	GLI3 is processed to GLI3R by the proteasome	UBC,PSMD1
94	REAC:5610783	0,00499	Degradation of GLI2 by the proteasome	UBC,PSMD1
95	REAC:5610780	0,00499	Degradation of GLI1 by the proteasome	UBC,PSMD1

APPENDIX B : g:profiler Proviral Pathways Results

Rank	Reactome ID	p-value	Pathway Name	Gene List
1	REAC:69275	9,08E-09	G2/M Transition	DYNLL1,EP300,PSMA6,PPP2R1A,TUBG1,CCNA2,MZT2B
2	REAC:453274	9,78E-09	Mitotic G2-G2/M phases	DYNLL1,EP300,PSMA6,PPP2R1A,TUBG1,CCNA2,MZT2B
3	REAC:72312	4,72E-08	rRNA processing	RPL18,NIP7,RPL10,UTP14A,RPL13,TSR1,RPLP2
4	REAC:72766	4,73E-07	Translation	RPL18,IARS2,SRP54,RPL10,RPL13,RPLP2
5	REAC:1799339	7,32E-07	SRP-dependent cotranslational protein targeting to membrane	RPL18,SRP54,RPL10,RPL13,RPLP2
6	REAC:927802	8,34E-07	Nonsense-Mediated Decay (NMD)	RPL18,PPP2R1A,RPL10,RPL13,RPLP2
7	REAC:975957	8,34E-07	Nonsense Mediated Decay (NMD) enhanced by the Exon Junction Complex (EJC)	RPL18,PPP2R1A,RPL10,RPL13,RPLP2
8	REAC:71291	9,3E-07	Metabolism of amino acids and derivatives	RPL18,PSMA6,OGDH,RPL10,RPL13,SLC3A2,RPLP2
9	REAC:69242	2,12E-06	S Phase	POLD1,PDS5B,PSMA6,CCNA2,E2F4
10	REAC:2467813	4,89E-06	Separation of Sister Chromatids	PDS5B,PSMA6,PPP2R1A,CDC20,SPC24
11	REAC:68886	0,0000056	M Phase	PDS5B,PSMA6,PPP2R1A,CDC20,SPC24,H2AFZ
12	REAC:68882	6,69E-06	Mitotic Anaphase	PDS5B,PSMA6,PPP2R1A,CDC20,SPC24

13	REAC:2555396	6,88E-06	Mitotic Metaphase and Anaphase	PDS5B,PSMA6,PPP2R1A,CDC20,SPC24
14	REAC:380287	8,11E-06	Centrosome maturation	DYNLL1,PPP2R1A,TUBG1,MZT2B
15	REAC:380270	8,11E-06	Recruitment of mitotic centrosome proteins and complexes	DYNLL1,PPP2R1A,TUBG1,MZT2B
16	REAC:192823	0,0000112	Viral mRNA Translation	RPL18,RPL10,RPL13,RPLP2
17	REAC:156902	0,0000112	Peptide chain elongation	RPL18,RPL10,RPL13,RPLP2
18	REAC:72764	0,0000133	Eukaryotic Translation Termination	RPL18,RPL10,RPL13,RPLP2
19	REAC:2408557	0,0000133	Selenocysteine synthesis	RPL18,RPL10,RPL13,RPLP2
20	REAC:156842	0,0000139	Eukaryotic Translation Elongation	RPL18,RPL10,RPL13,RPLP2
21	REAC:975956	0,0000145	Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	RPL18,RPL10,RPL13,RPLP2
22	REAC:2500257	0,0000177	Resolution of Sister Chromatid Cohesion	PDS5B,PPP2R1A,CDC20,SPC24
23	REAC:2132295	0,0000177	MHC class II antigen presentation	DNM2,DYNLL1,ACTR1B,AP2A1
24	REAC:72689	0,0000184	Formation of a pool of free 40S subunits	RPL18,RPL10,RPL13,RPLP2
25	REAC:5663205	0,0000216	Infectious disease	RPL18,PSMA6,RPL10,RPL13,RPLP2,AP2A1
26	REAC:187037	0,000022	NGF signalling via TRKA from the plasma membrane	DNM2,PSMA6,PPP2R1A,CAMK2G,GDNF,AP2A1
27	REAC:68877	0,0000239	Mitotic Prometaphase	PDS5B,PPP2R1A,CDC20,SPC24
28	REAC:156827	0,0000266	L13a-mediated translational silencing of	RPL18,RPL10,RPL13,RPLP2

			Ceruloplasmin expression	
29	REAC:72706	0,0000276	GTP hydrolysis and joining of the 60S ribosomal subunit	RPL18,RPL10,RPL13,RPLP2
30	REAC:2408522	0,0000327	Selenoamino acid metabolism	RPL18,RPL10,RPL13,RPLP2
31	REAC:72613	0,000035	Eukaryotic Translation Initiation	RPL18,RPL10,RPL13,RPLP2
32	REAC:72737	0,000035	Cap-dependent Translation Initiation	RPL18,RPL10,RPL13,RPLP2
33	REAC:2586552	0,0000356	Signaling by Leptin	PSMA6,PPP2R1A,STAT5A,CAMK2G,GDNF
34	REAC:451927	0,0000437	Interleukin-2 signaling	PSMA6,PPP2R1A,STAT5A,CAMK2G,GDNF
35	REAC:69206	0,0000464	G1/S Transition	PSMA6,PPP2R1A,CCNA2,E2F4
36	REAC:512988	0,0000514	Interleukin-3, 5 and GM-CSF signaling	PSMA6,PPP2R1A,STAT5A,CAMK2G,GDNF
37	REAC:168273	0,0000555	Influenza Viral RNA Transcription and Replication	RPL18,RPL10,RPL13,RPLP2
38	REAC:6791312	0,000068	TP53 Regulates Transcription of Cell Cycle Genes	EP300,CCNA2,E2F4
39	REAC:166520	0,0000704	Signalling by NGF	DNM2,PSMA6,PPP2R1A,CAMK2G,GDNF,AP2A1
40	REAC:168255	0,0000714	Influenza Life Cycle	RPL18,RPL10,RPL13,RPLP2
41	REAC:2262752	0,0000728	Cellular responses to stress	DYNLL1,EP300,PSMA6,CCNA2,CAMK2G,H2AFZ
42	REAC:453279	0,0000774	Mitotic G1-G1/S phases	PSMA6,PPP2R1A,CCNA2,E2F4
43	REAC:168254	0,0000951	Influenza Infection	RPL18,RPL10,RPL13,RPLP2
44	REAC:195721	0,000143	Signaling by Wnt	EP300,PSMA6,PPP2R1A,H2AFZ,AP2A1

45	REAC:1433557	0,000151	Signaling by SCF-KIT	PSMA6,PPP2R1A,STAT5A,CAMK2G,GDNF
46	REAC:1236394	0,000159	Signaling by ERBB4	PSMA6,PPP2R1A,STAT5A,CAMK2G,GDNF
47	REAC:186763	0,000185	Downstream signal transduction	PSMA6,PPP2R1A,STAT5A,CAMK2G,GDNF
48	REAC:380284	0,000198	Loss of proteins required for interphase microtubule organization__from the centrosome	DYNLL1,PPP2R1A,TUBG1
49	REAC:380259	0,000198	Loss of Nlp from mitotic centrosomes	DYNLL1,PPP2R1A,TUBG1
50	REAC:174184	0,000224	Cdc20:Phospho-APC/C mediated degradation of Cyclin A	PSMA6,CDC20,CCNA2
51	REAC:8854518	0,000224	AURKA Activation by TPX2	DYNLL1,PPP2R1A,TUBG1
52	REAC:179419	0,000234	APC:Cdc20 mediated degradation of cell cycle proteins prior to satisfation of the cell cycle checkpoint	PSMA6,CDC20,CCNA2
53	REAC:186797	0,000251	Signaling by PDGF	PSMA6,PPP2R1A,STAT5A,CAMK2G,GDNF
54	REAC:176409	0,000253	APC/C:Cdc20 mediated degradation of mitotic proteins	PSMA6,CDC20,CCNA2
55	REAC:176814	0,000263	Activation of APC/C and APC/C:Cdc20 mediated degradation of mitotic proteins	PSMA6,CDC20,CCNA2

56	REAC:380320	0,000276	Recruitment of NuMA to mitotic centrosomes	TUBG1,MZT2B
57	REAC:177504	0,000276	Retrograde neurotrophin signalling	DNM2,AP2A1
58	REAC:176408	0,000306	Regulation of APC/C activators between G1/S and early anaphase	PSMA6,CDC20,CCNA2
59	REAC:69656	0,000317	Cyclin A:Cdk2-associated events at S phase entry	PSMA6,CCNA2,E2F4
60	REAC:69202	0,000328	Cyclin E associated events during G1/S transition	PSMA6,CCNA2,E2F4
61	REAC:453276	0,000377	Regulation of mitotic cell cycle	PSMA6,CDC20,CCNA2
62	REAC:174143	0,000377	APC/C-mediated degradation of cell cycle proteins	PSMA6,CDC20,CCNA2
63	REAC:2565942	0,00039	Regulation of PLK1 Activity at G2/M Transition	DYNLL1,PPP2R1A,TUBG1
64	REAC:194315	0,000417	Signaling by Rho GTPases	PPP2R1A,ARHGAP21,CDC20,SPC24,H2AFZ
65	REAC:6804114	0,000463	TP53 Regulates Transcription of Genes Involved in G2 Cell Cycle Arrest	EP300,E2F4
66	REAC:201681	0,000479	TCF dependent signaling in response to WNT	EP300,PSMA6,PPP2R1A,H2AFZ
67	REAC:179812	0,000527	GRB2 events in EGFR signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
68	REAC:180336	0,000527	SHC1 events in EGFR signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
69	REAC:1250347	0,000527	SHC1 events in ERBB4 signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
70	REAC:112412	0,000527	SOS-mediated signalling	PSMA6,PPP2R1A,CAMK2G,GDNF
71	REAC:5673001	0,000527	RAF/MAP kinase cascade	PSMA6,PPP2R1A,CAMK2G,GDNF

72	REAC:5654693	0,000535	FRS-mediated FGFR1 signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
73	REAC:5654700	0,000535	FRS-mediated FGFR2 signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
74	REAC:5654712	0,000535	FRS-mediated FGFR4 signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
75	REAC:5654706	0,000535	FRS-mediated FGFR3 signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
76	REAC:5620912	0,000536	Anchoring of the basal body to the plasma membrane	DYNLL1,PPP2R1A,TUBG1
77	REAC:170984	0,00056	ARMS-mediated activation	PSMA6,PPP2R1A,CAMK2G,GDNF
78	REAC:187706	0,00056	Signalling to p38 via RIT and RIN	PSMA6,PPP2R1A,CAMK2G,GDNF
79	REAC:170968	0,000569	Frs2-mediated activation	PSMA6,PPP2R1A,CAMK2G,GDNF
80	REAC:5684996	0,000578	MAPK1/MAPK3 signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
81	REAC:69239	0,000586	Synthesis of DNA	POLD1,PSMA6,CCNA2
82	REAC:169893	0,000587	Prolonged ERK activation events	PSMA6,PPP2R1A,CAMK2G,GDNF
83	REAC:167044	0,000623	Signalling to RAS	PSMA6,PPP2R1A,CAMK2G,GDNF
84	REAC:912526	0,000633	Interleukin receptor SHC signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
85	REAC:5218921	0,000642	VEGFR2 mediated cell proliferation	PSMA6,PPP2R1A,CAMK2G,GDNF
86	REAC:187687	0,000691	Signalling to ERKs	PSMA6,PPP2R1A,CAMK2G,GDNF
87	REAC:5205685	0,000697	Pink/Parkin Mediated Mitophagy	MFN2,TOMM20
88	REAC:69306	0,000733	DNA Replication	POLD1,PSMA6,CCNA2
89	REAC:375165	0,000842	NCAM signaling for neurite out-growth	PSMA6,PPP2R1A,CAMK2G,GDNF
90	REAC:5663220	0,000879	RHO GTPases Activate Formins	PPP2R1A,CDC20,SPC24
91	REAC:69273	0,000903	Cyclin A/B1 associated events during G2/M transition	PPP2R1A,CCNA2

92	REAC:1538133	0,00105	G0 and Early G1	CCNA2,E2F4
93	REAC:195258	0,00107	RHO GTPase Effectors	PPP2R1A,CDC20,SPC24,H2AFZ
94	REAC:112399	0,00107	IRS-mediated signalling	PSMA6,PPP2R1A,CAMK2G,GDNF
95	REAC:5683057	0,00107	MAPK family signaling cascades	PSMA6,PPP2R1A,CAMK2G,GDNF
96	REAC:74751	0,00112	Insulin receptor signalling cascade	PSMA6,PPP2R1A,CAMK2G,GDNF
97	REAC:2428924	0,00112	IGF1R signaling cascade	PSMA6,PPP2R1A,CAMK2G,GDNF
98	REAC:2428928	0,00112	IRS-related events triggered by IGF1R	PSMA6,PPP2R1A,CAMK2G,GDNF
99	REAC:437239	0,00113	Recycling pathway of L1	DNM2,AP2A1
100	REAC:2404192	0,00114	Signaling by Type 1 Insulin-like Growth Factor 1 Receptor (IGF1R)	PSMA6,PPP2R1A,CAMK2G,GDNF
101	REAC:5205647	0,00122	Mitophagy	MFN2,TOMM20
102	REAC:448424	0,00124	Interleukin-17 signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
103	REAC:399721	0,00139	Glutamate Binding, Activation of AMPA Receptors and Synaptic Plasticity	CAMK2G,AP2A1
104	REAC:399719	0,00139	Trafficking of AMPA receptors	CAMK2G,AP2A1
105	REAC:74752	0,0015	Signaling by Insulin receptor	PSMA6,PPP2R1A,CAMK2G,GDNF
106	REAC:4420097	0,00165	VEGFA-VEGFR2 Pathway	PSMA6,PPP2R1A,CAMK2G,GDNF
107	REAC:194138	0,0018	Signaling by VEGF	PSMA6,PPP2R1A,CAMK2G,GDNF
108	REAC:2871796	0,00184	FCERI mediated MAPK activation	PSMA6,PPP2R1A,CAMK2G,GDNF
109	REAC:6804757	0,00188	Regulation of TP53 Degradation	PPP2R1A,CCNA2
110	REAC:5654696	0,00192	Downstream signaling of activated FGFR2	PSMA6,PPP2R1A,CAMK2G,GDNF

111	REAC:5654716	0,00192	Downstream signaling of activated FGFR4	PSMA6,PPP2R1A,CAMK2G,GDNF
112	REAC:5654708	0,00192	Downstream signaling of activated FGFR3	PSMA6,PPP2R1A,CAMK2G,GDNF
113	REAC:6806003	0,00198	Regulation of TP53 Expression and Degradation	PPP2R1A,CCNA2
114	REAC:5654687	0,00198	Downstream signaling of activated FGFR1	PSMA6,PPP2R1A,CAMK2G,GDNF
115	REAC:5654743	0,00198	Signaling by FGFR4	PSMA6,PPP2R1A,CAMK2G,GDNF
116	REAC:3371571	0,00198	HSF1-dependent transactivation	EP300,CAMK2G
117	REAC:5654741	0,00201	Signaling by FGFR3	PSMA6,PPP2R1A,CAMK2G,GDNF
118	REAC:5654736	0,00209	Signaling by FGFR1	PSMA6,PPP2R1A,CAMK2G,GDNF
119	REAC:5633007	0,00223	Regulation of TP53 Activity	EP300,PPP2R1A,CCNA2
120	REAC:2424491	0,00227	DAP12 signaling	PSMA6,PPP2R1A,CAMK2G,GDNF
121	REAC:5663202	0,00242	Diseases of signal transduction	EP300,PSMA6,PPP2R1A,STAT5A
122	REAC:1912408	0,00243	Pre-NOTCH Transcription and Translation	NOTCH3,EP300
123	REAC:177929	0,00246	Signaling by EGFR	PSMA6,PPP2R1A,CAMK2G,GDNF
124	REAC:3700989	0,00256	Transcriptional Regulation by TP53	EP300,PPP2R1A,CCNA2,E2F4
125	REAC:2172127	0,00256	DAP12 interactions	PSMA6,PPP2R1A,CAMK2G,GDNF
126	REAC:5654738	0,00267	Signaling by FGFR2	PSMA6,PPP2R1A,CAMK2G,GDNF
127	REAC:1428517	0,00274	The citric acid (TCA) cycle and respiratory electron transport	OGDH,UQCRB,ATP5L
128	REAC:69236	0,0028	G1 Phase	PPP2R1A,E2F4
129	REAC:69231	0,0028	Cyclin D associated events in G1	PPP2R1A,E2F4

130	REAC:1852241	0,00283	Organelle biogenesis and maintenance	DYNLL1,PPP2R1A,TUBG1,ATP5L
131	REAC:190236	0,00283	Signaling by FGFR	PSMA6,PPP2R1A,CAMK2G,GDNF
132	REAC:6788656	0,00306	Histidine, lysine, phenylalanine, tyrosine, proline and tryptophan catabolism	OGDH,SLC3A2
133	REAC:5617833	0,00343	Assembly of the primary cilium	DYNLL1,PPP2R1A,TUBG1
134	REAC:69620	0,00446	Cell Cycle Checkpoints	PSMA6,CDC20,CCNA2
135	REAC:450408	0,00452	AUF1 (hnRNP D0) binds and destabilizes mRNA	PSMA6,HNRNPD
136	REAC:174113	0,00452	SCF-beta-TrCP mediated degradation of Emi1	PSMA6,CDC20
137	REAC:1912422	0,00468	Pre-NOTCH Expression and Processing	NOTCH3,EP300
138	REAC:3906995	0,005	Diseases associated with O-glycosylation of proteins	NOTCH3,ADAMTS7
139	REAC:881907	0,00501	Gastrin-CREB signalling pathway via PKC and MAPK	PSMA6,PPP2R1A,CAMK2G,GDNF
140	REAC:2454202	0,00521	Fc epsilon receptor (FCERI) signaling	PSMA6,PPP2R1A,CAMK2G,GDNF

APPENDIX C FastQC per base quality reads of the libraries

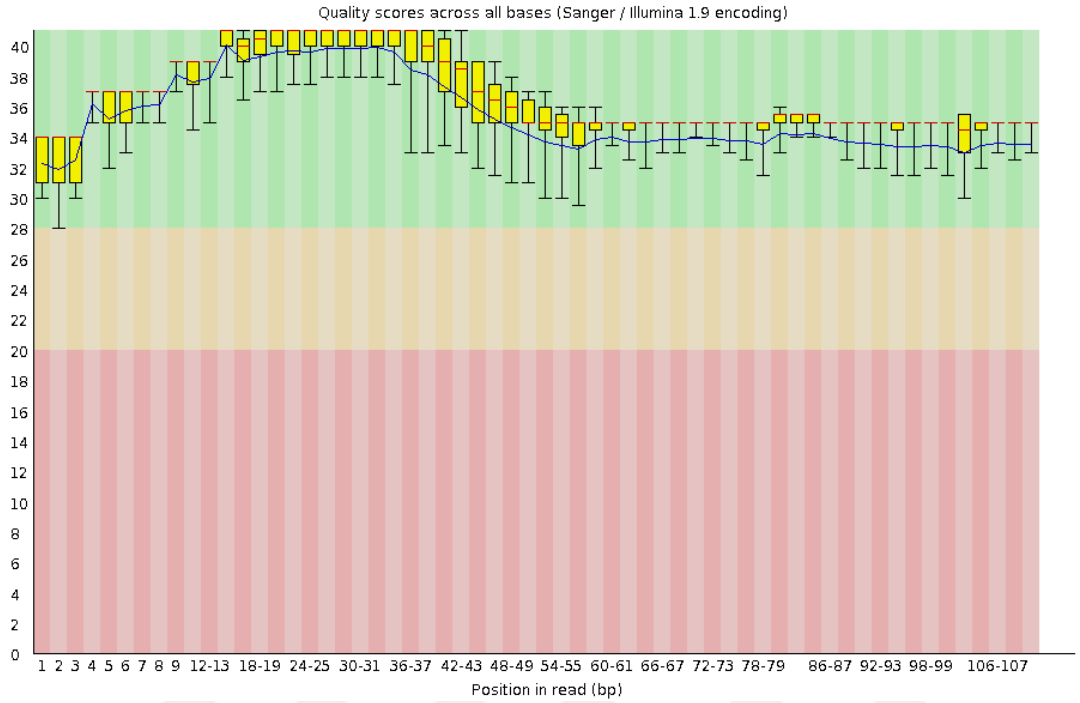

Figure C.1 FastQC base library A

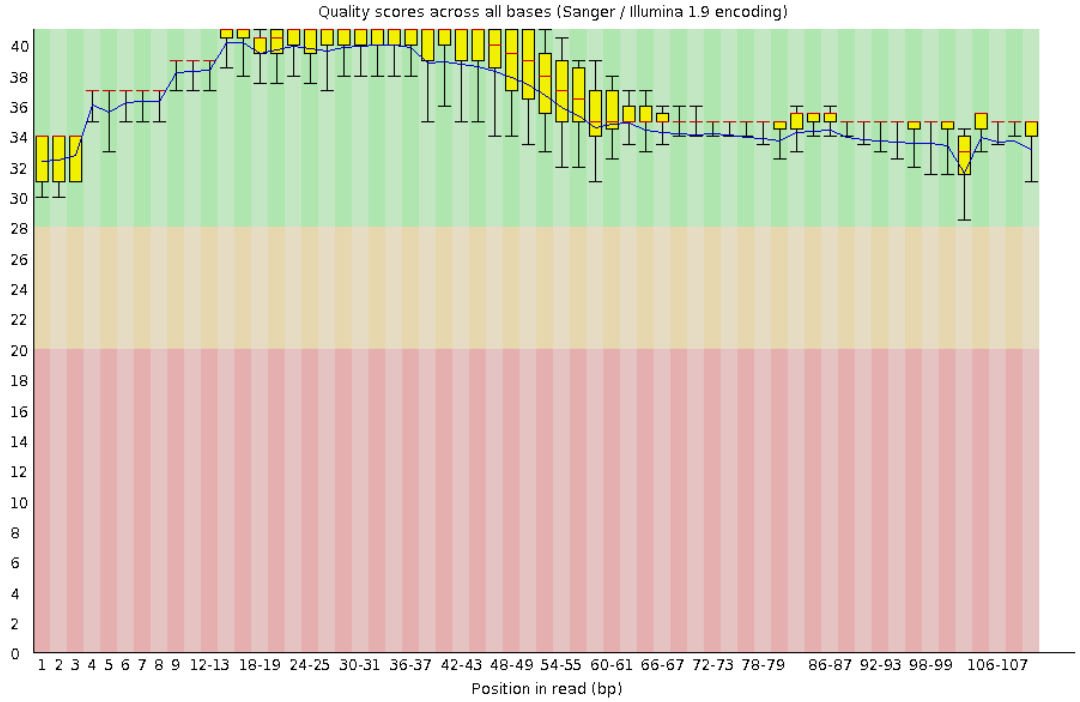

Figure C.2 FastQC of library A GFP+

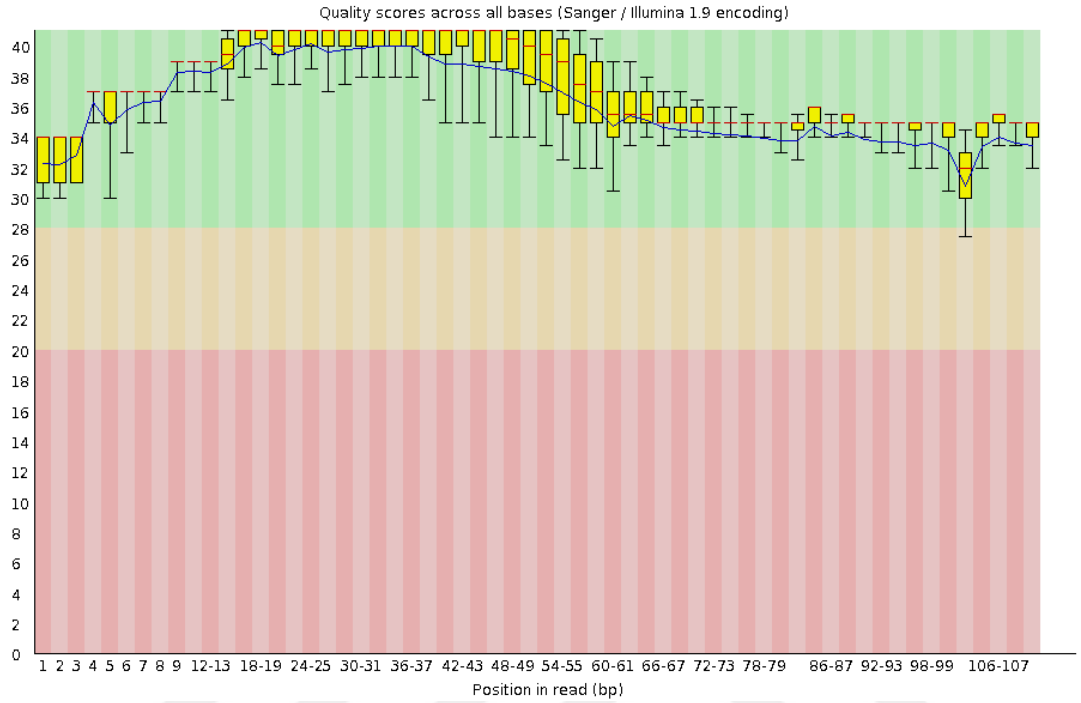

Figure C.3. FastQC of library A GFP-

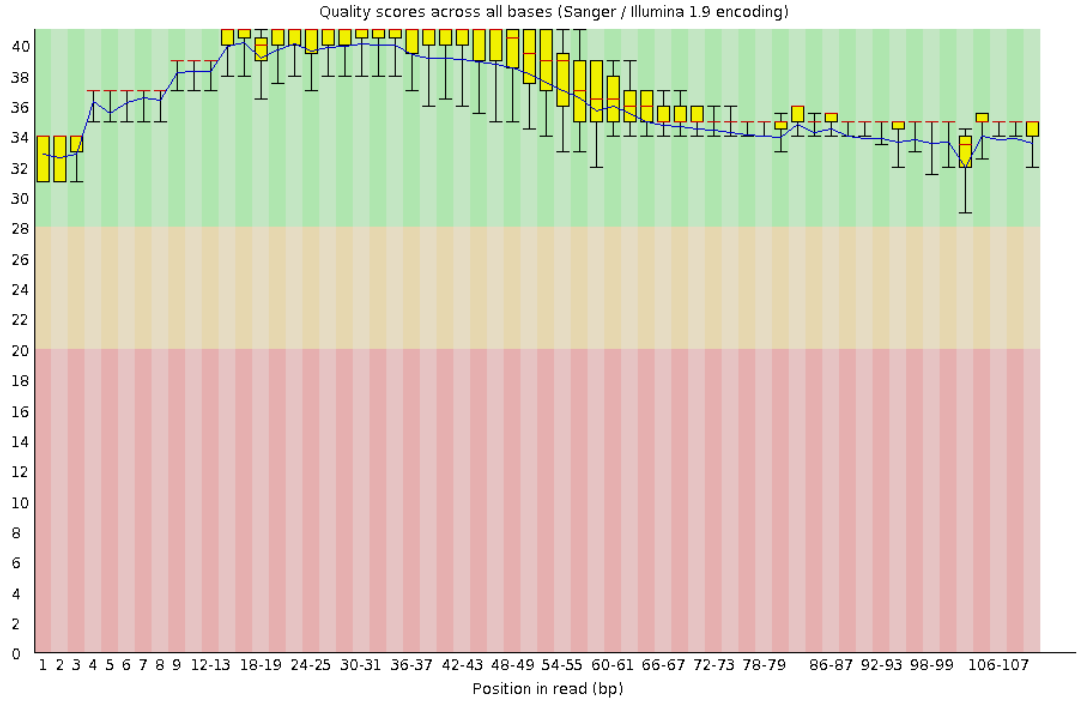

Figure C.4. FastQC base library B

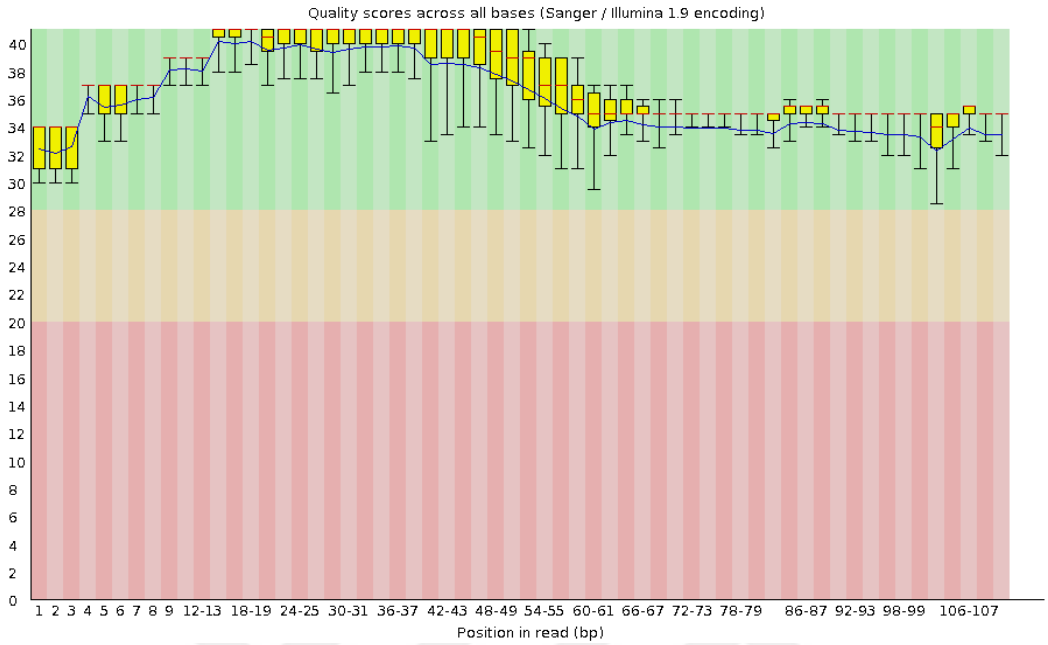

Figure C.5 FastQC base library B GFP+

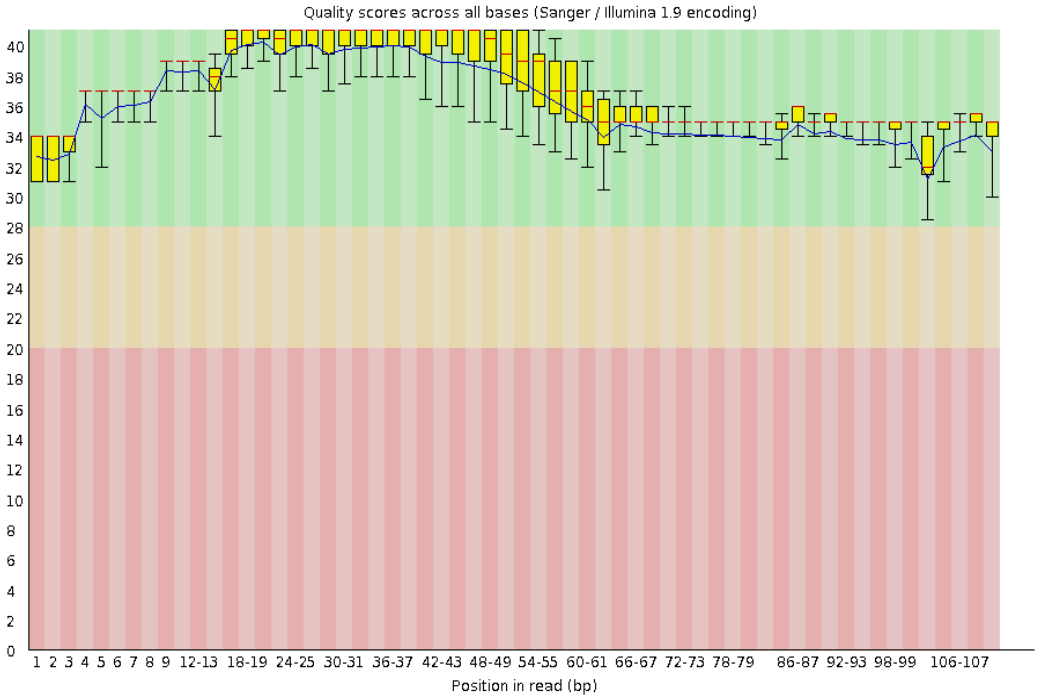

Figure C.6 FastQC base library B GFP-

APPENDIX D

GO Cellular Component Analysis via g:Profiler For Antiviral Pathway Candidate Genes

Rank	p-value	GO ID	Cellular Component Name	Gene list
1	6,39E-05	GO:0031012	extracellular matrix	FGF10,SERPINE1,TINAGL1,HNRNPU,MFAP4,DAG1
2	0,00378	GO:0005859	muscle myosin complex	MYH13,MYOM3
3	0,00532	GO:0016460	myosin II complex	MYH13,MYOM3
4	0,00631	GO:0098802	plasma membrane receptor complex	ACVR1C,BCL10,HTR3A
5	0,00668	GO:0060205	cytoplasmic vesicle lumen	APOB,SERPINE1,S100A12,PSMD1
6	0,00676	GO:0031983	vesicle lumen	APOB,SERPINE1,S100A12,PSMD1
7	0,00803	GO:0015629	actin cytoskeleton	MYH13,MTPN,CAP2,MYOM3
8	0,00915	GO:0000792	heterochromatin	ORC2,BEND3
9	0,01	GO:0016459	myosin complex	MYH13,MYOM3
10	0,0173	GO:0099512	supramolecular fiber	MYH13,MYOM3,BCL10,MFAP4
11	0,0179	GO:0045334	clathrin-coated endocytic vesicle	APOB,SFTPD
12	0,0183	GO:0099081	supramolecular polymer	MYH13,MYOM3,BCL10,MFAP4
13	0,0185	GO:0099080	supramolecular complex	MYH13,MYOM3,BCL10,MFAP4
14	0,0217	GO:0044420	extracellular matrix component	MFAP4,DAG1
15	0,0242	GO:0030139	endocytic vesicle	APOB,SFTPD,UBC
16	0,0317	GO:1902554	serine/threonine protein kinase complex	ACVR1C,CCNL1
17	0,0349	GO:0005814	centriole	CEP250,PCNT
18	0,0417	GO:1902911	protein kinase complex	ACVR1C,CCNL1
19	0,0443	GO:0043235	receptor complex	ACVR1C,BCL10,HTR3A

20	0,0453	GO:0071013	catalytic step 2 spliceosome	SKIV2L2,HNRNPU

GO Cellular Component Analysis via g:Profiler For Proviral Pathway Candidate Genes

Rank	**p-value**	**GO ID**	**Cellular component**	**Gene List**
1	2,51E-07	GO:0005813	centrosome	DYNLL1,ACTR1B,TUBG1,CKAP2,MZT2B,CCNF,CEP57L1,DCAF12
2	3,48E-06	GO:0022625	cytosolic large ribosomal subunit	RPL18,RPL10,RPL13,RPLP2
3	0,000042	GO:0015934	large ribosomal subunit	RPL18,RPL10,RPL13,RPLP2
4	4,35E-05	GO:0022626	cytosolic ribosome	RPL18,RPL10,RPL13,RPLP2
5	7,62E-05	GO:0000775	chromosome,centromeric region	PDS5B,DYNLL1,PPP2R1A,SPC24
6	8,69E-05	GO:0098687	chromosomal region	POLD1,PDS5B,DYNLL1,PPP2R1A,SPC24
7	0,000281	GO:0044391	ribosomal subunit	RPL18,RPL10,RPL13,RPLP2
8	0,000306	GO:0044455	mitochondrial membrane part	MFN2,UQCRB,ATP5L,TOMM20
9	0,000333	GO:0005819	spindle	DYNLL1,CDC20,TUBG1,MZT2B
10	0,000395	GO:0030684	preribosome	NIP7,UTP14A,TSR1
11	0,000482	GO:0005840	ribosome	RPL18,RPL10,RPL13,RPLP2
12	0,000631	GO:0044445	cytosolic part	RPL18,RPL10,RPL13,RPLP2
13	0,000875	GO:0000930	gamma-tubulincomplex	TUBG1,MZT2B
14	0,00118	GO:0031306	intrinsiccomponent of mitochondrial outer membrane	MFN2,TOMM20
15	0,00121	GO:0000228	nuclearchromosome	POLD1,SMARCB1,TUBG1,H2AFZ,E2F4

16	0,00198	GO:0044450	microtubule organizingcenter part	TUBG1,MZT2B,CCNF
17	0,00203	GO:0030666	endocytic vesicle membrane	DNM2,CAMK2G,AP2A1
18	0,00268	GO:0098798	mitochondrial proteincomplex	UQCRB,ATP5L,TOMM20
19	0,00283	GO:0005905	clathrin-coated pit	DNM2,AP2A1
20	0,00346	GO:0015629	actincytoskeleton	NOTCH3,ARHGAP21,SLC9A3R1,ACTR1B
21	0,00427	GO:0031461	cullin-RING ubiquitin ligasecomplex	CDC20,CCNF,DCAF12
22	0,00427	GO:0005874	microtubule	DNM2,TUBG1,CKAP2
23	0,00551	GO:0000785	chromatin	PDS5B,SMARCB1,H2AFZ,E2F4
24	0,00554	GO:0101002	ficolin-1-rich granule	DYNLL1,ACTR1B,IMPDH2
25	0,00749	GO:0099512	supramolecular fiber	DNM2,PSMA6,TUBG1,CKAP2
26	0,00764	GO:0098573	intrinsiccomponent of mitochondrial membrane	MFN2,TOMM20
27	0,00765	GO:0098589	membrane region	DNM2,EQTN,AP2A1
28	0,00793	GO:0099081	supramolecular polymer	DNM2,PSMA6,TUBG1,CKAP2
29	0,008	GO:0099080	supramolecularcomplex	DNM2,PSMA6,TUBG1,CKAP2
30	0,00861	GO:0044454	nuclearchromosome part	POLD1,SMARCB1,H2AFZ,E2F4
31	0,0102	GO:0097223	sperm part	SLC9A3R1,EQTN
32	0,011	GO:0030139	endocytic vesicle	DNM2,CAMK2G,AP2A1
33	0,0167	GO:0000151	ubiquitin ligasecomplex	CDC20,CCNF,DCAF12
34	0,017	GO:0000790	nuclearchromatin	SMARCB1,H2AFZ,E2F4
35	0,0192	GO:0099513	polymericcytoskeletal fiber	DNM2,TUBG1,CKAP2
36	0,0193	GO:1990204	oxidoreductasecomplex	OGDH,UQCRB

37	0,0219	GO:0000776	kinetochore	DYNLL1,SPC24
38	0,0315	GO:0005741	mitochondrial outer membrane	MFN2,TOMM20
39	0,0331	GO:0005875	microtubule associatedcomplex	DYNLL1,ACTR1B
40	0,0343	GO:0005925	focal adhesion	RPL18,DNM2,RPLP2
41	0,0347	GO:0098800	inner mitochondrial membrane proteincomplex	UQCRB,ATP5L
42	0,0348	GO:0005924	cell-substrate adherens junction	RPL18,DNM2,RPLP2
43	0,0352	GO:1904813	ficolin-1-rich granule lumen	ACTR1B,IMPDH2
44	0,0359	GO:0030055	cell-substrate junction	RPL18,DNM2,RPLP2
45	0,0363	GO:0031968	organelle outer membrane	MFN2,TOMM20
46	0,0385	GO:0019867	outer membrane	MFN2,TOMM20
47	0,0454	GO:0005759	mitochondrial matrix	IARS2,OGDH,HSPE1
48	0,0481	GO:0031300	intrinsiccomponent of organelle membrane	MFN2,TOMM20