

“Interplay of Mast cells and Natural killer cells”

A Major Project dissertation submitted in partial fulfilment of the requirement for the degree of

**Master of Technology
In
Bioinformatics**

Submitted by

ANIRUDDHA BANIK

(2K15/BIO/02)

Delhi Technological University, Delhi, India

Under the supervision of

Dr. Asmita Das

(Assistant Professor)



Department of Biotechnology

Delhi Technological University

Shahbad Daultpur, Main Bawana Road,

Delhi – 110042, INDIA



CERTIFICATE

This is to certify that the M.Tech. Dissertation entitled “*Interplay of Mast cells and Natural killer cells*”, submitted by **ANIRUDDHA BANIK (2K15/BIO/02)** in partial fulfilment of the requirement for the reward of the degree of Master of Technology in Bioinformatics, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi), is an authentic record of the candidate’s own work carried out by him under my guidance.

The information and data enclosed in this dissertation is original and has not been submitted elsewhere for honouring of any other degree.

Dr. Asmita Das

(Supervisor)

Assistant Professor

Department of Biotechnology

Delhi Technological University

Shahbad Bawana Road, Delhi-110042

Prof. D. Kumar

(Co-Supervisor)

Head of the Department

Department of Biotechnology

Delhi Technological University

Shahbad Bawana Road, Delhi- 110042

DECLARATION

I, **Aniruddha Banik**, hereby declare that the M. Tech. report entitled “*Interplay of Mast cells and Natural killer cells*” submitted in partial fulfilment of the requirement for the award of the degree of Master of Technology, Delhi Technological University, is a record of original and independent research work done by me under the supervision and guidance of **Dr. Asmita Das, (Assistant Professor)** Department of Biotechnology at Delhi Technological University, Delhi and the thesis has not formed the basis of the award of any Degree/Diploma/Associate ship/Fellowship or other similar title to any candidate of any University/institution.

Date:

Signature of the Candidate

ACKNOWLEDGEMENT

I would first like to thank my thesis advisor Dr. Asmita Das, (Assistant Professor) Department of Biotechnology at Delhi Technological University. The door to Prof. Asmita Das office was always open whenever I ran into a trouble spot or had a question about my research or writing. She consistently allowed this thesis to be my own work, but steered me in the right direction whenever she thought I needed it.

I would also like to express my sincere thanks to Vice-chancellor Yogesh Singh and Prof. D. Kumar, H.O.D., Department of Biotechnology, D.T.U., Delhi for giving me this opportunity to undertake this project.

I would also like to express my sincere thanks to Harikrishan, Razi Khan, Mottadi Shiva, Kapil Jangra for helping in project work and their much needed cooperation.

I also wish to express my gratitude to my teachers and other non-teaching staff members of the Department of Biotechnology, DTU; especially Mr. Chail Bihari, Mr. Jitendra Singh for providing us their valuable teachings with constant support and guidance.

Finally, I must express my very profound gratitude to my parents and my friend Dr. Debarati Karmakar for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Place : Delhi

ANIRUDDHA BANIK

Date :

M.Tech. (Bioinformatics)

(2K15/BIO/02)

CONTENTS

TOPIC

Page No.

List of Figures

List of Abbreviations

1. ABSTRACT	9
2. INTRODUCTION	10
3. REVIEW OF LITERATURE	12
4. TOOLS AND DATABASES USED	17
5. METHODOLOGY AND APPROACH	24
6. RESULTS AND DISCUSSION	28
7. CONCLUSION	57
8. FUTURE PERSPECTIVES	58
9. REFERENCES	59

LIST OF FIGURES

S.No.	Title	Page No.
1.	Methodology	16
2.	Microarray gene expression data of Mast cells	21
3.	Un-stimulated gene expression profile graph of Mast cells	21
4.	Anti-IgE stimulated gene expression profile graph of Mast cells	22
5.	Mast cells gene expression data in experimental conditions	22
6.	R-Script of Box plot	23
7.	Box plot of Non-Normalized gene expression data of Mast cells	24
8.	Q-Q plot of Box plot	25
9.	Q-Q plot of Non-Normalized gene expression data of Mast cells	26
10.	General formula of Mean and Simple Normalization	27
11.	Box plot of Normalized gene expression data of Mast cells	28
12.	Q-Q plot of Normalized gene expression data of Mast cells	29
13.	Fold change value of Normalized gene expression data of Mast cells	31
15.	Graphical view of Fold change of Mast cells genes	32
16.	Threshold calculated using Quantile function	32
17.	Graphical view of Up-regulated genes in Mast cells	33

18.	Graphical view of Up-regulated genes in Mast cells	36
19.	Identification of common Up-regulated genes between Mast cells and NK cells	39
20.	List of common Up-regulated genes between Mast cells and NK cells	40
21.	Find out the interaction between common genes or proteins using String database	41
22.	Design a Mast cells TLR-mediated pathway	43
23.	Mast cells degranulation pathway and NK cells receptor-mediated pathway	45
24.	NK cells receptor-mediated and TLR-mediated pathway	47

LIST OF ABBREVIATIONS

MC	- Mast Cell
NK	- Natural killer Cells
TLR	- Toll like receptor
IL2	- Interleukin 2
TNF- α	- Tumour necrosis factor α
STRING	- Search Tool for the Retrieval of Interacting genes
NCBI	- National Centre for Biotechnology Information
GEO	- Gene Expression Omnibus
MYD88	- Myeloid differentiation primary response gene 88
Log FC	- Logarithm of Fold Change value
NN	- Non- Normalized data
TIRAP	- TIR Domain Containing Adaptor Protein
IRAK1	- Interleukin 1 Receptor Associated Kinase 1
TRAF6	- TNF Receptor Associated Factor 6
IRF5	- Interferon Regulatory Factor 5

ABSTRACT

Mast cells (MCs) that are well known for their important effector function in IgE-associated immune responses play a key role in innate immune defences. Natural killer (NK) cells are important in innate immunity, first described as guardians for the detection and clearance of transformed or virus-infected cells. In this study, we investigate the interaction between MCs and NK cells. After analysing Gene Expression Microarray data of Anti-IgE stimulated Mast cells and K562 erythroleukemic tumour stimulated human NK cells, a pathway was designed through which natural killer cells may be activated by mast cells secreting different kinds of inflammatory cytokines such as IL-8, IL-2, TNF-alpha etc. In case of in-vitro condition we can activate both natural killer cells and mast cells by using synthetic ligands such as Pam3CSK4 which is recognized by TLR-2 present in both of the cells.

INTRODUCTION:

To develop an effective immune response, the cells of the immune system are required to communicate between each other through secretion of soluble mediators and direct cell–cell interaction. Among the cells of the immune system, mast cell (MC) is one of the most powerful in terms of ability to respond to multiple stimuli and to selectively release different types and amounts of mediators. MCs originate from multipotent hematopoietic progenitors in bone marrow, and then migrate through blood to tissues where they mature [4]. MCs are positioned throughout the vascularized tissues where they constitute one of the first cell types of the immune system able to interact with allergens and antigens [4].

Mast cells heterogeneity depending on the tissue distribution and it is reflected by their ability to react to multiple stimuli and by immunoglobulin E (IgE)-dependent and -independent activation pathways [6] [30]. FcεRI and Fcγ receptors, Toll like receptors (TLRs), complement receptors, cytokine and chemokine receptors can regulate MC activation [2] [3]. Depending on the type of the stimuli they receive, MCs secrete a diverse and wide range of biologically active products that can trigger or suppress the immune response [6] [30]. Preformed mediators, such as histamine and certain cytokines are rapidly and instantaneously released upon MC activation and newly synthesized mediators, such as cytokines, chemokines start to be synthesized after MC activation [2] [3]. Although these products are all important in both innate and acquired immunity and the rapid release of MC mediators is very crucial for the initiation of the immune response at the site of infection and they are able to modulate the immune-cell trafficking to provide co-stimulatory signals for different cells activation [2] [3].

The cellular interactions that play a role in innate immune defense, emerging evidences show MC-dependent natural killer (NK) cell recruitment and activation [36]. NK cells are granular cytotoxic and circulating lymphocytes involved in the clearance of transformed and pathogen-infected cells. In innate immune system, NK cells recruitment to the site of infection is mediated by a large spectrum of chemokines which bind to the chemokine receptors, CCR2, CCR5 and CXCR3 on NK cells [36].

In our study we designed a signalling pathway which is TLR-mediated MYD88 pathway, which may be present in Mast cells and different inflammatory cytokines (IL-2, IL-4, IL-8 and TNF-alpha) are released through this pathway from Mast cells which helps in the activation of natural killer cells when these cytokines (IL-2, IL-4, IL-8 and TNF-alpha) bind to the receptors on Natural killer cells. During literature survey from different research papers we got the information about the interaction between Mast cells and Natural killer cells but on this topic very certain amount of research papers we have got. Because still the topic (Mast cells and Natural killer cells interaction) is unclear to the researchers how mast cells can activate natural killer cells during tumour killing. In research papers, they told about only the cytokines or chemokines mediated natural killer cells activation but they did not told about the pathway through which Mast cells interact with Natural killer cells.

For this approach we gone through the microarray data of Mast cells (**Human Mast cells activation via Fc (Epsilon) RI** (GEO accession id: **GSE3271**) (Set of arrays organised by shared biological context, such as tumours types) and Natural killer cells (**Gene expression analysis of tumour stimulated human Natural killer cells**) (GEO accession id is **GSE55977**) Collected both of these microarray data files from NCBI's GEO (Gene Expression Omnibus Repository) and analysed our result with GEO2R (a GEO Tool). Then, we downloaded gene expression data file for that particular data set. After importing expression data in excel sheet then arranging and fixing data (The Data Pre-processing Step). After the Data Pre-processing step we have to do normalization of the genes expression data, because the expression data of genes was collected from 'RAW' file which contains various noise that can hamper our accurate result that's why data Normalization is the very important step during handling with microarray data. Then we adopt quantile function to calculate Threshold value which is based upon Log FC value of the particular gene. Two different Threshold values have calculated for Mast cells and Natural killer cells microarray data. This Threshold value was applied on the Log FC value which represents one gene Fold change value then the genes are sorted out whose expression data pass the Threshold value. Both of the mast cells and Natural killer cells up-regulated (The expression value of the particular gene passes the Threshold value) genes are compared using the Java program. After successfully running the java program we got the matched genes which are common between Mast cells and Natural killer cells. After that Studied about the signalling pathway present in both of the MCs and NK cells in KEGG pathway. Then validating the Genes and Proteins present in signalling pathway, also present in microarray data of MCs and NK cells or not and also check their Log FC value and after getting the proteins also check the PPI (Protein-Protein interaction network) in string then design the signalling pathway which may be present in Mast cells and also help in the activation of Natural killer cells.

REVIEW OF LITERATURE:

Mast cells: Mast cells are important cells of the immune system and are of the hematopoietic lineage [4]. Mast cells are originated from pluripotent progenitor cells of the bone marrow, and mature under the influence of the c-kit ligand and stem cell factor in the presence of other distinct growth factors provided by the microenvironment of the tissue where they are destined to reside [4]. Under normal conditions, mature mast cells do not circulate in the bloodstream [4]. However, mast cell progenitors migrate into tissues and differentiate into mast cells under the influence of stem cell factor and various cytokines [4]. Mast cells are present throughout the body and they play important roles in the maintenance of many physiological functions as well as in the pathophysiology of diseases [4]. Accordingly, this review is focused on the role of mast cells in a wide range of physiological functions and pathogenesis of a variety of disease states [4].

Location of Mast cells: Mast cells are found in mucosal and epithelial tissues throughout the body [4] [14]. In rodents, mast cells also reside in peritoneal and thoracic cavities [14]. Mast cells are found in all vascularized tissues except for the central nervous system and the retina [14]. Mast cells are located at the junction point of the host and external environment at places of entry of antigen (gastrointestinal tract, skin, respiratory epithelium) [14]. Mast cells are located in areas below the epithelium in connective tissue surrounding blood cells, smooth muscle, mucous, and hair follicles [14].

The cytoplasm of the mast cell contains 50–200 large granules that store inflammatory mediators, including histamine, heparin, and a variety of cytokines, chondroitin sulfate, and neutral proteases [4]. In order for mast cells to migrate to their target locations, the co-ordinated effects of integrins, adhesion molecules, chemokines, cytokines, and growth factors are necessary [4].

There are two phenotypes of human mast cells: mucosal mast cells that produce only tryptase and connective tissue mast cells that produce chymase, tryptase, and carboxypeptidases [14]. Mast cell activation and mediator release have different effects in various tissues and organs [14]. Most common sites in the body exposed to antigens are the mucosa of the respiratory tract (airborne), gastrointestinal tract (food borne), blood (wounds, absorption from respiratory tract/gastrointestinal tract), and connective tissues [14].

When the gastrointestinal tract is exposed to an antigen, its response is to increase fluid secretion, increase smooth muscle contraction, and increase peristalsis [14]. Proteins derived from different plants and animals can act as antigens and activate the immune system in vulnerable subjects [14]. The antigen (peptide) permeates through the epithelial layer of the mucosa of the gut and binds to IgE on mucosal mast cells [14]. These peptides are presented to Th2 cells, and if there is an IgE antibody against the peptide present, it will cause activation of the mast cell resulting in an immune response [14]. This causes mast cells to degranulate and release a variety of inflammatory mediators [14]. These mediators increase vascular permeability, causing edema in the gut epithelium and smooth muscle contraction, which lead to vomiting and diarrhea [14]. This type of reaction can occur in response to peptides found in certain medications [14].

In the respiratory tract, the immune response to mast cell activation results in airway constriction, increased mucous production, and cough. The most common introduction of antigens to the respiratory tract is via inhalation [14]. Mucosal mast cells in the nasal epithelium are activated by antigens that diffuse across the mucosa after being inhaled [14]. In the respiratory tract, mast cell degranulation increases vascular permeability and local edema, which can obstruct nasal airways and lead to congestion [14]. There is increased production of mucus and its accumulation can block off the sinuses and result in a bacterial infection [14].

Mechanism of activation of Mast cells: Mast cells are known for their main mechanism of action: IgE-mediated allergic reactions through the Fc ϵ RI receptor. IgE antibodies are produced by mature B cells in response to CD4⁺ Th2 cells [2] [3]. Naïve mature B cells produce IgM and IgD antibodies [3]. Once they become activated by an antigen, B cells will proliferate [3]. If these B cells interact with cytokines, such as IL-4 (which is modulated by CD4⁺ Th2 cells), the antibody class switches from IgM to IgE [3]. IgE is mostly found bound to Fc ϵ RI receptors on the mast cell and very little IgE is found as a soluble antibody in circulation [2] [3]. When an antigen comes in contact with the mast cell, it crosslinks two or more Fc ϵ RI molecules and activates the release of granules from the mast cell. IgE is found in the connective tissue under epithelial layers of the skin, in the respiratory tract, and also in the gastrointestinal tract [3]. In addition to Fc ϵ RI, mast cells also express Fc receptors for IgA and IgG, receptors for adenosine, C3a, chemokines, cytokines, and pathogen-associated molecular patterns (PAMPs), as well as toll-like receptors (TLRs), all of which are involved in mast cell activation and immune response [3].

The most common physiological pathway for mast cell activation is via antigen/IgE/Fc ϵ RI cross linking [2] [3]. Fc ϵ RI consists of an α -chain that binds to IgE, a β -chain, which spans the membrane, and γ chains, which are a disulfide-linked homodimer. Fc ϵ RI interacts with LYN tyrosine kinase, which phosphorylates the tyrosine in its immunoreceptor tyrosine bases activation motifs (ITAMs) on the B and γ chains of the Fc ϵ RI [2] [3]. Lyn activates Syk tyrosine kinases, which phosphorylates signaling proteins, such as LAT1 and LAT2 (linkers for activation of T cells) [3]. Phosphorylated PLC γ hydrolyzes phosphatidylinositol-4, 5-bisphosphate to make inositol-1, 4, 5-triphosphate (IP3) and diacylglycerol (DAG) [3]. IP3 and DAG are second messengers and IP3 causes calcium mobilization from the endoplasmic reticulum [3]. Calcium release activates and causes NF κ B to translocate to the nucleus of the cell, which results in transcription of cytokines, such as IL-6, TNF α , and IL-13. Zeb2 is involved in regulation of degranulation upon stimulation via Fc ϵ RI [3]. Activation of Fc ϵ RI activates Fyn (Src kinase) [3]. Fyn regulates mast cell degranulation, which is complementary to the Lyn signaling pathway [6]. Fyn activates PI3K, which activates Akt and produces PIP3 [6]. This activates mTOR, which is involved in mast cell chemotaxis and cytokine production. There are also receptors for IgG called Fc γ R [6]. The γ -chain homodimer is the same in Fc γ RI as in Fc ϵ RI so the signal sent from Fc γ R can crosstalk with Fc ϵ RI [6].

Physiological roles of Mast cells: Mast cells are involved in the regulation of variety of physiological functions, including vasodilation, angiogenesis, bacterial, and parasite elimination [30]. In addition, mast cells regulate functions of many cell types, such as dendritic cells, macrophages, T cells, B cells, fibroblasts, eosinophils, endothelial cells, and epithelial cells [30]. Since, mast cells generate and release multi-potent molecules, such as histamine, proteases,

prostanoids, leukotrienes, heparin, and many cytokines, chemokines, and growth factors, they have the capacity to be involved in regulating the functions of many organs and tissues [30]. One of the mostly studied functions of the mast cell is its role in vascular and bronchial homeostasis [30]. Mast cells also play a significant role in the regulation of bone growth, remodeling, and mineral homeostasis [30].

Activation and Mediators Release: Activation of mast cells occurs when an antigen crosslinks IgE molecules that are bound to FcεRI on the surface of the mast cell [2]. FcεRI receptor for IgE has an affinity 100 times greater for the Fc of IgE than of IgG. Because of this, IgE is found bound to the FcεRI receptors on the mast cell even when there are no antigens present [3]. As a result, this makes the response of the mast cell to an antigen very fast. FcεRI signaling uses the Lyn-dependent phosphorylation of ITAMs on the B and γ subunits of the FcεRI [30]. Protein kinase Syk is activated and auto-phosphorylated after being recruited to the ITAMs [6]. Syk phosphorylates linker activation of T cells (LAT) and non-T cell activation linker (NTAL) [6]. LAT phosphorylates PLC, which produces IP3 and DAG, which activates intracellular calcium influx and PKC activation. NTAL activates PI3K, which also helps with calcium release [6]. This results in degranulation of the mast cells, lipid mediator production, and cytokine production [3].

Degranulation occurs a few seconds after crosslinking and results in release of the inflammatory mediators that are stored in the granules [2] [3]. Many of the mediators that are stored or newly synthesized by the mast cells attract leukocytes (eosinophils, basophils, Th2 lymphocytes, neutrophils) to the inflammatory site and amplify the inflammatory response [6]. The inflammatory mediators increase the permeability of the blood vessels so that the immune cells can move from the blood stream to the affected tissue [3]. After degranulation, mast cells resynthesize the mediators and repopulate granules [2] [3]. Histamine and heparin are also stored in pre-packaged granules and are involved with vascular permeability and smooth muscle contraction [3] [6]. Histamine is the most important mediator released from the mast cell involved with an allergic response [6]. Histamine is derived from the amino acid histidine and works through three different receptors (H1, H2, H3) [6]. Stimulation of H1 receptors by the binding of histamine induces the classic allergic reaction [3]. H1 receptors are found on smooth muscle cells and endothelial cells [3]. Activation of H1 receptors on endothelial cells results in increased vascular permeability and activation of smooth muscle cells resulting in contraction, constriction of airways, and mucous secretion [6]. TNFα, also stored in the MCG, activates macrophages, endothelium, and cytokines [6]. TNF-α binds to endothelial cells and resulting in increased adhesion molecule expression. Leukocytes can bind to these adhesion molecules and then are brought to the site of inflammation [3] [6].

Natural Killer cells: Years ago, the histologic and functional definition of an NK cell was that of a large granular lymphocyte that could kill a target cell “naturally,” that is, in a spontaneous fashion that did not require any priming and was not restricted by the target cell's expression of major histocompatibility complex (MHC) molecules. Experiments in mouse models of bone marrow graft rejection led to the proposal that NK cells would kill any target that lacked self-major histocompatibility complex (MHC) class I molecules (the “missing self” hypothesis) [18]. This extraordinary idea was developed before anyone knew what the NK cell was using to “see” its targets [18]. It is now clear that NK cells have a multitude of inhibitory and activating receptors that

engage MHC class I molecules, MHC class I-like molecules, and molecules unrelated to MHC [18]. Thus, NK cells are indeed restricted in what target cells they can engage by the expression of the target's MHC ligands, but in a very complex fashion that remains incompletely understood [18]. Notably, orthologs of more recently discovered NK-cell receptor families cannot be found beyond mammals, suggesting that the composite modern day NK cell emerged well after T and B cells appeared to define the vertebrate adaptive immune system. Furthermore, the complementary roles that NK and cytolytic T cells have in target recognition and host defense, and their similar mechanisms of cytolysis, suggest that these 2 cell types may have each evolved from a common ancestral cytolytic effector cell [20]. Finally, a subset of human NK cells produce abundant cytokines with modest or no ability to lyse target cells [18]. NK cells are crucial components of the innate immune system and, as the name suggests, they do not require pre-stimulation to perform their effector functions [18]. Morphologically, they are characterized as large, granular, bone marrow-derived lymphocytes and phenotypically, they are defined as CD56⁺ CD3⁻ in humans [18]. They represent 10% of the cells in the total peripheral blood mononuclear cell (MNC) population of circulating human lymphocytes and they comprise the third largest population of lymphocytes following B and T cells [18]. They are also found in the peritoneal cavity, spleen, liver, lung, lymph nodes, thymus, and in uterus during gestation [18].

NK-cell Development: It is generally accepted that NK cells develop primarily in the bone marrow, similar to B cells and myeloid origin cells [20]. However, recent studies have shown that NK cells can also develop in lymph nodes and liver [20]. The generation of NK cells from hematopoietic stem cells (HSC) is a continuous process. In the first step, the HSC shows commitment towards NK-cell lineage [20]. NK-cell precursors (NKP) have been identified in the hematopoietic population, which differentiates into NK cells but not to other lineages [20]. This process is followed by phenotypic and functional NK-cell maturation [20]. In the final step, NK cells undergo homeostasis [20]. Several transcription factors as well as soluble and membrane factors have been identified that regulate NK-cell development and maturation [20]. Transcription factors involved in the generation of NKP include Ets-1, Id2, Ikaros and PU.1. Maturation of immature NK cells is regulated by Gata-3 and IRF-2 and functional differentiation of matured NK cells involves CEBP- γ , MEF and MITF [20]. The cytokine IL-15 has been shown to be essential for NK-cell development, homeostasis and survival [20]. Studies by Freud and Ferlazzo have implicated the role of T cell derived IL-2 in the cytolytic functional maturation of NK cells [20].

Natural Killer cells Function: NK cells can secrete cytokines and chemokines that influence the host's immune response, and/or kill certain infected or transformed cells via perforin/granzyme or death receptor (eg. Fas, TRAIL)-related pathways. Interferon gamma (IFN- γ) is considered the prototypic NK-cell cytokine, and its production by NK cells is known to shape the Th1 immune response, activate APCs to further up-regulate MHC class I expression, activate macrophage killing of obligate intracellular pathogens and have anti-proliferative effects on viral- and malignant-transformed cells [25]. For many of these functions, it would make sense for NK cells to be in close proximity to APCs and T cells [25]. NK-cell functioning is controlled by a wide range of receptors that are expressed on the cell surface [25]. These receptors are either inhibitory or activating in nature. The family of inhibitory receptors consists of the killer immunoglobulin-like receptors (KIR) or Ig-like receptors (CD158), the C type lectin receptors (CD94-NKG2A) and leukocyte

inhibitory receptors (LIR1, LAIR-1) [12]. Activating receptors are the natural cytotoxicity receptors (NKp46, NKp44), C type lectin receptors (NKG2D, CD94-NKG2C), and Ig-like receptors (2B4) [12]. A particular NK cell typically expresses two to four inhibitory receptors in addition to an array of activation receptors [12] [13]. As different NK cells express different combinations of inhibitory or activating receptors, there is sizeable heterogeneity within the NK-cell population [25]. It is for this reason that NK cells are considered to have the ability to respond to a variety of stimuli and to participate in immune responses under different pathological conditions [25]. NK-cell cytotoxicity is tightly regulated by a balance between activating and inhibitory signals [12] [25]. The inhibitory NK-cell receptors recognize self-MHC class I molecule, and this prevents NK-cell activation [12]. This explains self-tolerance and prevention of host cell killing. It was earlier discovered that NK cells are activated when they encounter cells which lack self-MHC class I molecule [13]. This is known as the 'missing-self' hypothesis. Moreover, NK cells can discriminate between normal host cells and infected or abnormal cells by recognition of MHC class I molecules [13]. Virally infected cells and tumour cells often down-regulate MHC class I expression to escape recognition by cytotoxic T lymphocytes (CTL), but this results in their vulnerability towards NK-cell attack [25]. In this condition, activation receptors are no longer suppressed and they induce potent stimulatory signals, therefore tipping the balance in favour of NK-cell activation. This condition is often referred to as induced-self recognition [12] [13] [25].

Interaction between Mast cells and NK cells: The cellular interactions that play a role in innate immune defense, emerging evidences show MC-dependent natural killer (NK) cell recruitment and activation [36]. NK cells are granular cytotoxic and circulating lymphocytes involved in the clearance of transformed and pathogen-infected cells [36]. As a part of the innate immune system, their recruitment to the site of infection is mediated by a large spectrum of chemokines which bind to the chemokine receptors, CCR2, CCR5 and CXCR3 on NK cells [36]. Activated MCs can induce NK cell accumulation in different disease models. For instance, immune surveillance by MCs is important for NK cell recruitment and viral clearance during Dengue infection [36]. Human cord blood-derived MCs stimulated with virus-associated TLR3 agonist can recruit human NK via the CXCL8 and CXCR1 axis, underlining MC role as sentinel cell during early viral infections [36]. Lipopolysaccharide (LPS)-activated BMDCs induce cell contact-dependent IFN- γ secretion by NK cells, without affecting cell-mediated cytotoxicity [36]. Cellular interaction is partly mediated by OX40L expression on MCs [36]. In the cited work, authors underline that different MC signals of activation confer different results in terms of NK activation [36]. In fact, in addition to LPS, stimulation of MCs via TLR3 and TLR9, but not with IgE/antigen, amplifies IFN- γ secretion by NK cells [36]. Similarly, in a model of hepato-carcinoma, MC pro-tumoral role is associated with reduction of NK cell number and activation [36]. This effect was due to the fact that, in the tumour micro environment, SCF-activated MCs release adenosine that inhibits production of IFN- γ by NK cells [36]. Enhanced CCL3-mediated recruitment of NK cells is instead observed in an orthotopic melanoma model in which TLR2-activated MCs exert anticancer properties by secreting large amounts of this chemokine [36].

TOOLS AND DATABASE USED:

Here, we have a list of various tools and databases, which were being used during this project work. In this project different tools, d=bioinformatics databases and also different programming languages were used for the development of signalling pathway.

1) **GEO (Gene Expression Omnibus)**: The Gene Expression Omnibus (GEO) database is an international public repository that archives and freely distributes high-throughput gene expression and other functional genomics data sets. Created in 2000 as a worldwide resource for gene expression studies, GEO has evolved with rapidly changing technologies and now accepts high-throughput data for many other data applications, including those that examine genome methylation, chromatin structure, and genome–protein interactions. GEO supports community-derived reporting standards that specify provision of several critical study elements including raw data, processed data, and descriptive metadata. The database not only provides access to data for tens of thousands of studies, but also offers various Web-based tools and strategies that enable users to locate data relevant to their specific interests, as well as to visualize and analyze the data. This chapter includes detailed descriptions of methods to query and download GEO data and use the analysis and visualization tools. Gene Expression Omnibus (GEO) is a database supported by the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM) that accepts raw and processed data with written descriptions of experimental design, sample attributes, and methodology for studies of high-throughput gene expression and genomics. The introduction of DNA microarrays and the Serial Analysis of Gene Expression (SAGE) protocol as methods of simultaneously assaying gene expression of multiple genes in 1995 enabled scientists to study gene expression of hundreds to thousands of genes, thereby vastly increasing the experimental scale and providing a far more complete understanding of biological processes compared to earlier single-gene studies.

In 2000, NCBI launched the GEO database as a repository for high-throughput gene expression data. In 2002, major journals started to require deposit of microarray data into public repositories, and consequently, the content of GEO grew quickly. Furthermore, the nature of high-throughput genomic experiments expanded rapidly since the first microarrays used to analyze gene expression, and thus the GEO database similarly evolved to keep pace with the changing technologies and applications. Today, GEO accepts data from a wide variety of technologies, including DNA microarrays, protein or tissue arrays, high-throughput nucleic acid sequencing, SAGE, and RT-PCR. And while the majority, approximately 90%, of the data in GEO are indeed gene expression data, the applications have also expanded to include studies on genome methylation, genome binding/occupancy, protein profiling, chromosome conformation studies, and genome variation/copy number.

The GEO homepage is at <http://www.ncbi.nlm.nih.gov/geo/>.

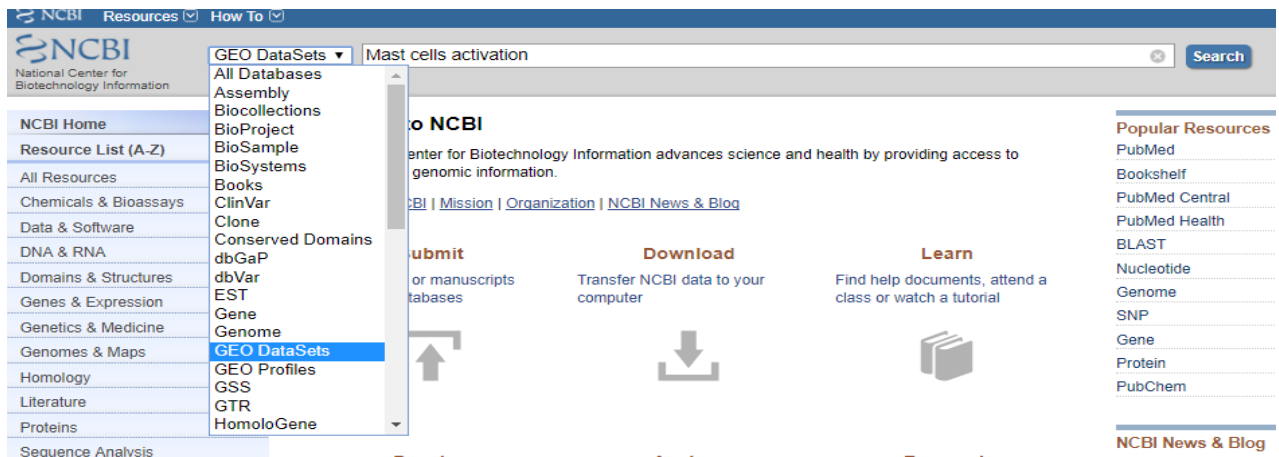


Fig 1: Home Page of NCBI with Search tab

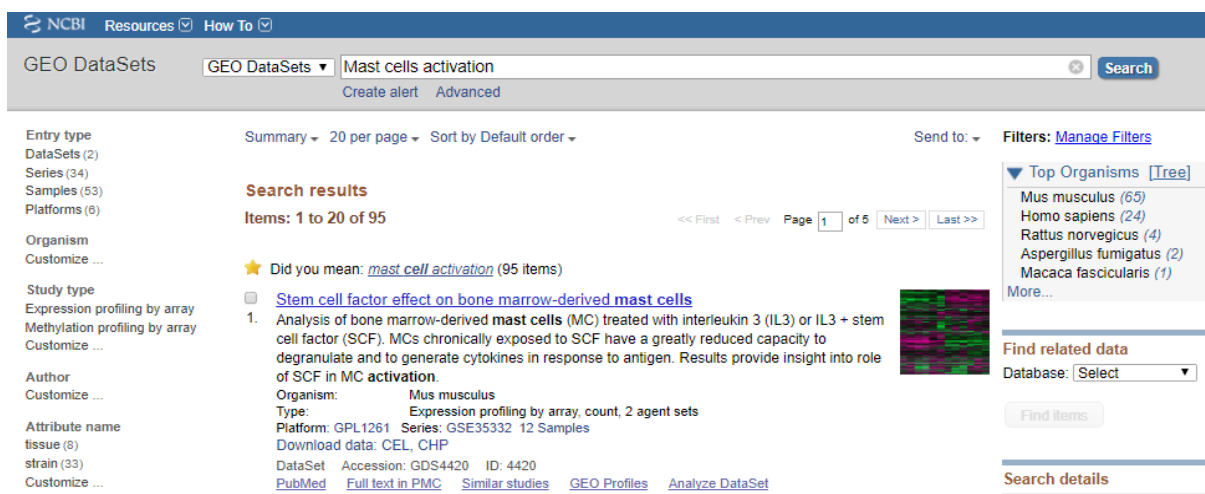


Fig2: After selecting Geo Datasets and entering query in search box the result page contains Microarray data

2) **GEO2R (GEO-Tool)** : GEO2R is an interactive web tool that allows users to compare two or more groups of Samples in a GEO Series in order to identify genes that are differentially expressed across experimental conditions. Results are presented as a table of genes ordered by significance.

GEO2R performs comparisons on original submitter-supplied processed data tables using the GEO-query and limma R packages from the Bioconductor project. Bioconductor is an open source software project based on the R programming language that provides tools for the analysis of high-throughput genomic data. The *GEO-query* R package parses GEO data into R data structures that can be used by other R packages. The *limma* (Linear Models for Microarray Analysis) R package has emerged as one of the most widely used statistical tests for identifying differentially expressed genes. It handles a wide range of experimental designs and data types and applies multiple-testing corrections on P-values to help correct for the occurrence of false positives. Thus, GEO2R provides a simple interface that allows users to perform R statistical analysis without command line expertise.

[Gene expression analysis of tumor stimulated human NK cells](#)

39. (Submitter supplied) Genome wide expression profiling of human NK cells stimulated with K562 erythroleukemic tumor cells after four hours of NK-tumor co-culture. Responding NK cells were compared to non-responding NK cells, delineated by display of CD107 on the NK cell surface following cytotoxic granule release. We hypothesized that tumor responses would initiate rapid changes in gene expression in the NK cell that would identify new features of the anti-tumor response of NK cells. [more...](#)

Organism: Homo sapiens
 Type: Expression profiling by array
 Platform: GPL6883 12 Samples
 Download data: TXT

Series Accession: GSE55977 ID: 200055977

[PubMed](#) [Full text in PMC](#) [Similar studies](#) [Analyze with GEO2R](#)



Fig3: For GEO2R Analysis click on the link ‘Analyze with GEO2R’

GEO accession Gene expression analysis of tumor stimulated human NK cells

▼ Samples ▼ Define groups

Enter a group name:

Group	Accession		Source name	Developmental stage	Status	Cell type	Sort pop
CD107Pos	GSM1349722		NK cell	adult	healthy	NK	CD107+
CD107Pos	GSM1349723		NK cell	adult	healthy	NK	CD107+
CD107Pos	GSM1349724		NK cell	adult	healthy	NK	CD107+
CD107Neg	GSM1349725	Donor1 CD107neg Rep1	NK cell	adult	healthy	NK	CD107-
CD107Neg	GSM1349726	Donor1 CD107neg Rep2	NK cell	adult	healthy	NK	CD107-
CD107Neg	GSM1349727	Donor1 CD107neg Rep3	NK cell	adult	healthy	NK	CD107-
CD107Pos	GSM1349728	Pooled CD107pos Rep1	NK cell	adult	healthy	NK	CD107+
CD107Pos	GSM1349729	Pooled CD107pos Rep2	NK cell	adult	healthy	NK	CD107+
CD107Pos	GSM1349730	Pooled CD107pos Rep3	NK cell	adult	healthy	NK	CD107+
CD107Neg	GSM1349731	Pooled CD107neg Rep1	NK cell	adult	healthy	NK	CD107-
CD107Neg	GSM1349732	Pooled CD107neg Rep2	NK cell	adult	healthy	NK	CD107-
CD107Neg	GSM1349733	Pooled CD107neg Rep3	NK cell	adult	healthy	NK	CD107-

Fig4: For GEO2R Analysis Define groups of the available cell lines 2 groups must be defined

CD107Pos	GSM1349729	Pooled CD107pos Rep2	NK cell	adult	healthy	NK	CD107+
CD107Pos	GSM1349730	Pooled CD107pos Rep3	NK cell	adult	healthy	NK	CD107+
CD107Neg	GSM1349731	Pooled CD107neg Rep1	NK cell	adult	healthy	NK	CD107-
CD107Neg	GSM1349732	Pooled CD107neg Rep2	NK cell	adult	healthy	NK	CD107-
CD107Neg	GSM1349733	Pooled CD107neg Rep3	NK cell	adult	healthy	NK	CD107-

GEO2R

▼ Quick start

- Specify a GEO Series accession and a Platform if prompted.
- Click 'Define groups' and enter names for the groups of Samples you plan to compare, e.g., test and control.
- Assign Samples to each group. Highlight Sample rows then click the group name to assign those Samples to the group. Use the Sample metadata (title, source and characteristics) columns to which group.
- Click 'Top 250' to perform the calculation with default settings.
- Results are presented as a table of genes ordered by significance. The top 250 genes are presented and may be viewed as profile graphs. Alternatively, the complete results table may be saved.
- You may change settings in Options tab.

How to use



Fig5: Click on the ‘Save all results’ and copy the results into Spread sheet.

Quick start

Log-transformation has been applied to the data. You can change this in the Options tab.

Recalculate if you changed any options. Save all results Select columns

ID	adj.P.Val	P.Value	t	B	logFC	Gene.symbol
ILMN_1798204	1.14e-07	5.34e-12	-29.47	16.29	-2.149	IL21R
ILMN_1723004	4.58e-07	4.30e-11	-24.41	14.92	-2.563	CD72
ILMN_1697554	4.39e-06	6.19e-10	19.15	12.9	1.341	SASH3
ILMN_1661861	4.51e-06	1.25e-09	-17.95	12.31	-4.788	CSF2

Fig6: The View of Result page after analysis.

3) **R Programming Language:** R is a programming language and software environment for statistical analysis, graphics representation and reporting. R was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand, and is currently developed by the R Development Core Team. R is freely available under the GNU General Public License, and pre-compiled binary versions are provided for various operating systems like Linux, Windows and Mac. This programming language was named **R**, based on the first letter of first name of the two R authors (Robert Gentleman and Ross Ihaka), and partly a play on the name of the Bell Labs Language **S**.

For generating Box-plot ‘Limma’ and ‘Readxl’ both of these packages are required. The ‘Limma’ Package which is downloaded from Bio-conductor site. Mainly from the Excel file R-Script reads only the different Genes and their respective Expression data. In case of analysis the data from Gene expression experiments ‘Limma’ package provides an integrated solution. This Package is very helpful for reading, normalizing and exploring differential expression analysis of microarray and high- throughput PCR data.

For generating Q-Q plot ‘Limma’, ‘Stats’ and ‘gplots’ both of these packages are required. The ‘Limma’ Package which is downloaded from Bio-conductor site. Mainly from the Excel file R-Script reads only the different Genes and their respective Expression data. In case of analysis the data from Gene expression experiments ‘Limma’ package provides an integrated solution. This Package is very helpful for reading, normalizing and exploring differential expression analysis of microarray and high- throughput PCR data. The ‘Stats’ package contains functions for statistical calculations. The ‘Gplots’ package is used in plotting data, including: calculating and plotting locally smoothed summary function as ‘bandplot’. With the help of this above R-Script Histogram is also generated.

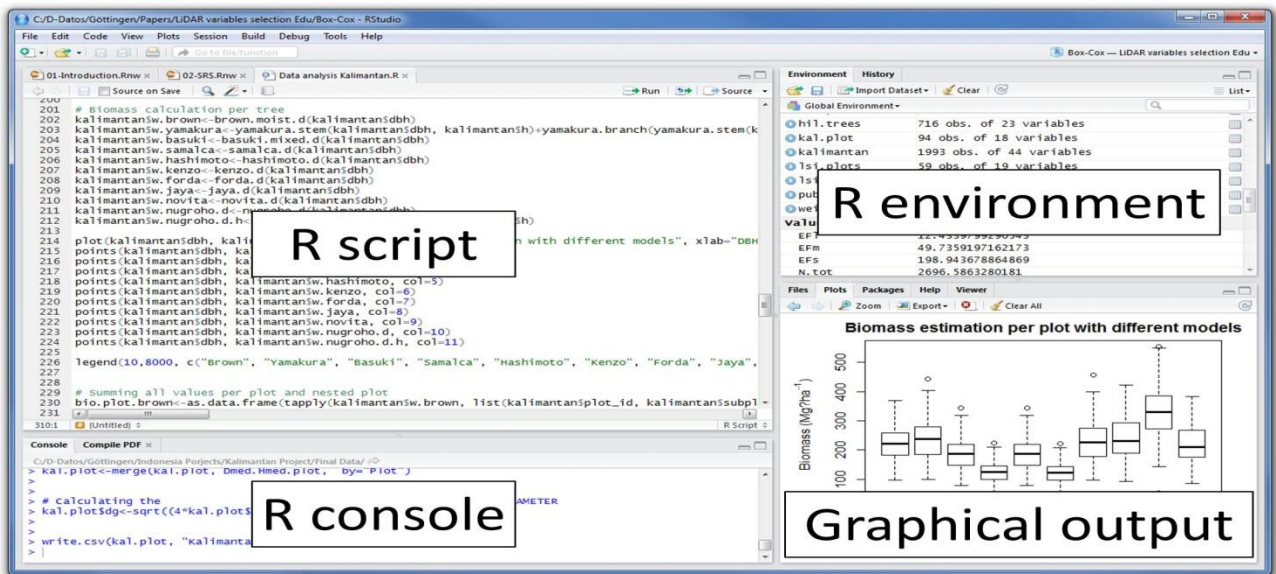


Fig7: Above Figure showing R studio IDE environment for R Language Interface

4) **Netbeans(Software for Java Programming):** NetBeans is a software development platform written in Java. The NetBeans Platform allows applications to be developed from a set of modular software components called *modules*. Applications based on the NetBeans Platform, including the NetBeans integrated development environment (IDE), can be extended by third party developers.

The NetBeans IDE is primarily intended for development in Java, but also supports other languages, in particular PHP, C/C++ and HTML5.

NetBeans is cross-platform and runs on Microsoft Windows, Mac OS X, Linux, Solaris and other platforms supporting a compatible JVM.

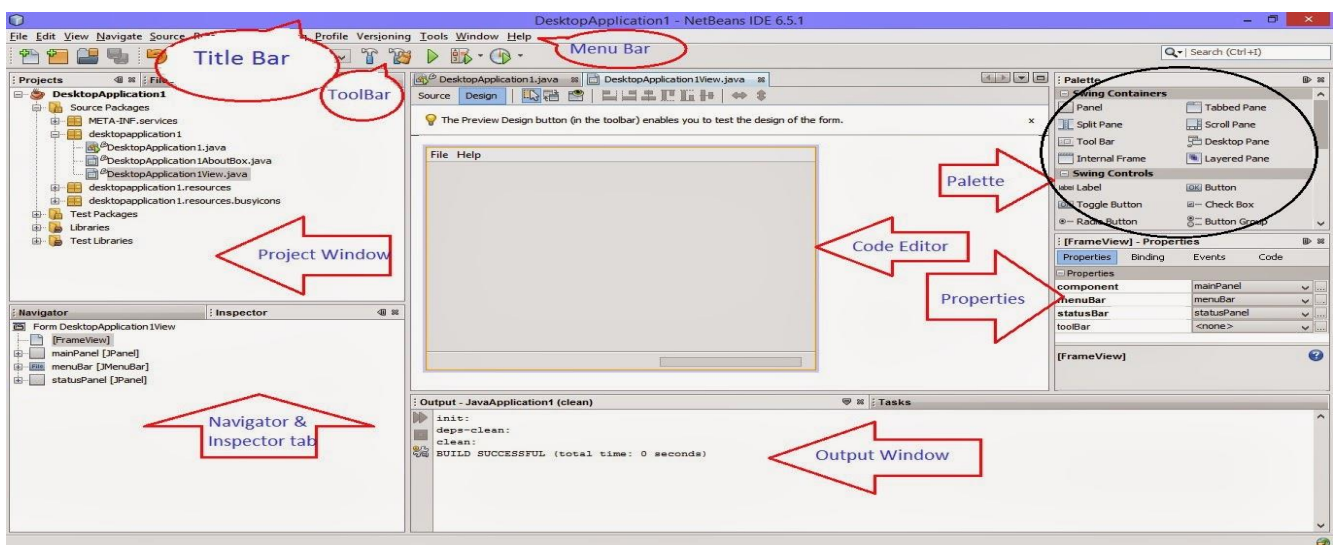


Fig8: User Interface of Netbeans for the Development of Java program.

5) **Cytoscape:** Cytoscape is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating with gene expression profiles and other state data. Additional features are available as plugins. Plugins are available for network and molecular profiling analyses, new layouts, additional file format support and connection with databases and searching in large networks. Plugins may be developed using the Cytoscape open Java software architecture by anyone and plugin community development is encouraged.^{[1][2]} Cytoscape also has a JavaScript-centric sister project named Cytoscape.js that can be used to analyse and visualise graphs in JavaScript environments, like a browser.

Cytoscape was originally created at the Institute of Systems Biology in Seattle in 2002. Now, it is developed by an international consortium of open source developers. Cytoscape was initially made public in July, 2002 (v0.8); the second release (v0.9) was in November, 2002, and v1.0 was released in March 2003.

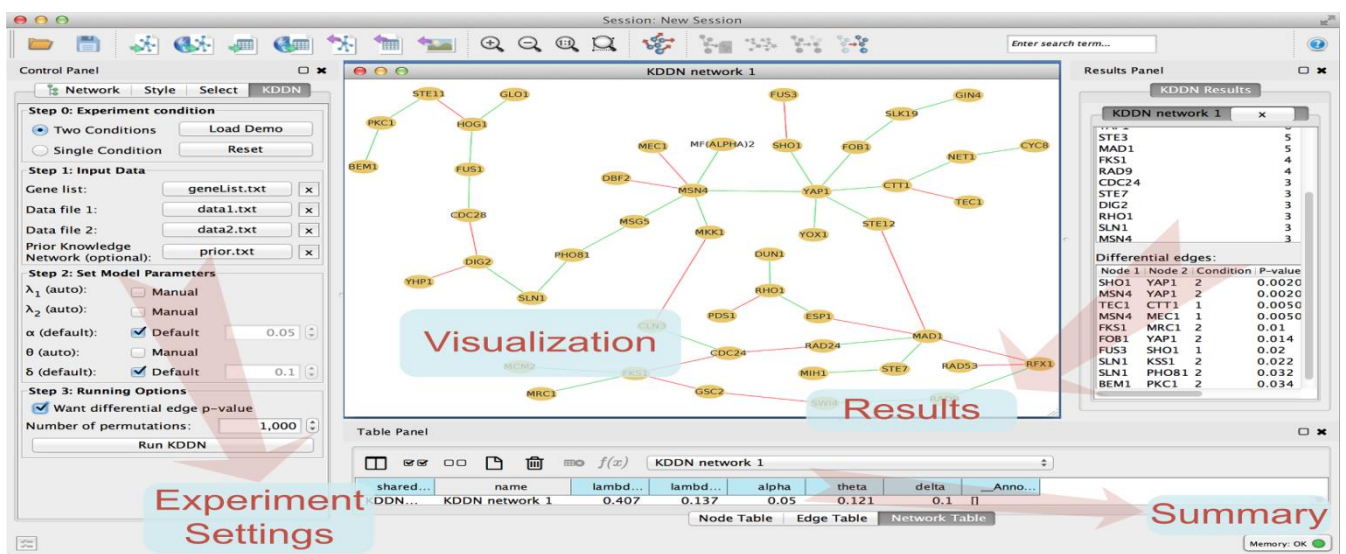


Fig9: User Interface of Cytoscape Software for the development of Models (Signalling pathways)

6) **KEGG Pathway:** KEGG PATHWAY is a collection of manually drawn pathway maps representing our knowledge on the molecular interaction, reaction and relation networks for:

1. Metabolism
2. Genetic Information Processing
3. Environmental Information Processing
4. Cellular Processes
5. Organismal Systems
6. Human Diseases
7. Drug Development

KEGG PATHWAY is a reference database for **Pathway Mapping**.



Fig10: Home Page of KEGG pathway database

Pathway Text Search

Number of entries in a page: 20 | Hide thumbnail

Items: 1 - 5 of 5

Entry	Thumbnail Image	Name	Description	Object
hsa05310		Asthma - Homo sapiens (human)	... Inhaled allergens encounter antigen presenting cells (APC) that line the airway. Upon recognition o...	...64: Fc epsilon RI signaling pathway hsa04660: T cell receptor signaling pathway hsa04612: Antigen pr...
hsa04664		Fc epsilon RI signaling pathway - Homo sapiens (human)	Fc epsilon RI-mediated signaling pathways in mast cells are initiated by the interaction of antigen ...	C00696 (Prostaglandin D2) C05952 (Leukotriene E4) C05951 (Leukotriene D4) C02166 (Leukotriene C4) CO...
hsa05140		Leishmaniasis - Homo sapiens (human)	Leishmania is an intracellular protozoan parasite of macrophages that causes visceral, mucosal, and ...	C00076 (Calcium cation) C00533 (Nitric oxide) C02737 (Phosphatidylserine) C00195 (N-Acylsphingosine)...

Fig11: The result page contains Pathway lists after selecting the organism and entering the query in search box.

METHODOLOGY:

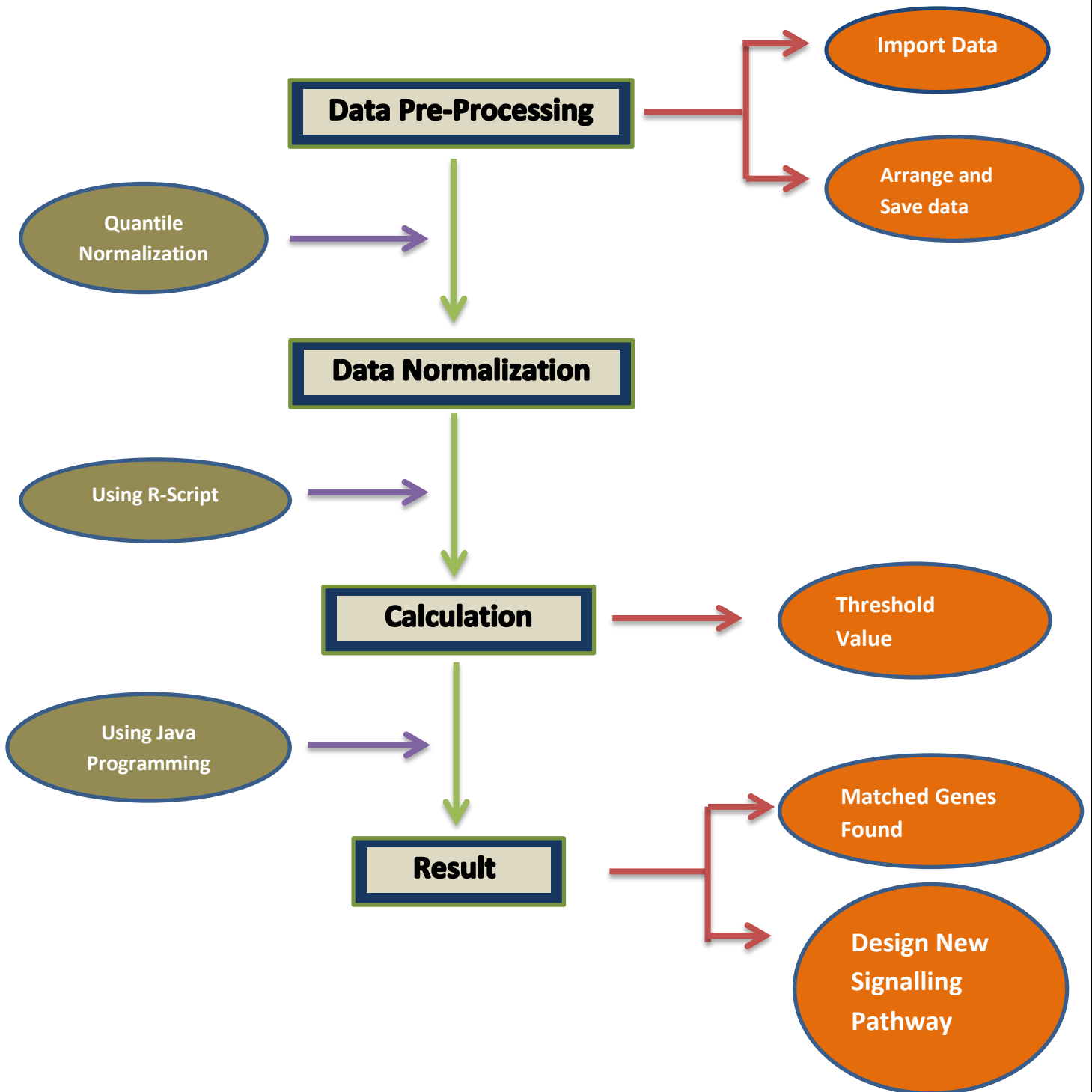


Fig1: Different Stages were used to Design a New Signalling Pathway

APPROACH:

Step 1: Gone through several research papers to get information about the mast cells and natural killer cells such as –

- (i) Different TLRs present on mast cells and natural killer cells.
- (ii) Different signalling pathways present in mast cells and natural killer cells.
- (iii) Different types of cytokines, chemokines are released in both of the mast cells and natural killer cells signalling pathway.
- (iv) Different kinds of synthetic ligands available for different TLRs.



Step 2: Microarray data was collected from Gene Expression Omnibus Repository Datasets of NCBI.

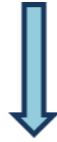
- (i) **Human Mast cells activation via Fc (Epsilon) RI** (GEO accession id: **GSE3271**) (Set of arrays organised by shared biological context, such as tumours types).
- (ii) **Gene expression analysis of tumour stimulated human Natural killer cells** (GEO accession id is **GSE55977**).



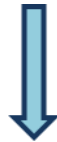
Step 3: Next step was pre-processing of the data by arranging them on the basis of genes with their respective ref IDs and expression values according to the given cell lines.



Step 4: Microarray gene expression data was normalised by using **any global Median Normalisation** or **Simple Normalisation** to convert raw data of expression files to appropriate linearly arranged data.



Step 5: From the Normalised genes expression data of each cell median was calculated for control sets and also calculated for disease sets. Data was normalised for mast cells and natural killer cells.



Step 6: Fold change of genes expression was calculated with the help of below formula:

$$\text{FC (Control set/Disease set)} \sim \text{Signal Intensity (Control set)/Signal Intensity (Disease set)}.$$



Step 7: R script was used to set the threshold value for mast cells and natural killer cells by using Quantile function. This threshold value was mainly applied on gene expression data.



Step 8: On the basis of threshold value of mast cells and natural killer cells only the above fold change values were selected (up-regulated genes) from mast cells and natural killer cells microarray data.



Step 9: In excel file only the up-regulated mast cells genes and up-regulated natural killer cells genes are imported and using Java programming the common genes were found out.



Step 10: After finding common genes thoroughly studied the pathway of Natural killer cells activation (NCR mediated and TLR-mediated) and pathway of Mast cells activation (Degranulation) in KEGG pathway to find out how many genes are involved from common genes list in those above mentioned pathways (NCR mediated and TLR mediated NK cells activation and Mast cells degranulation).



Step 11: Using String the interaction between the genes or proteins which are present in a particular signalling pathway were found out and designed a new signalling pathway of mast cells which is not present in KEGG pathway.

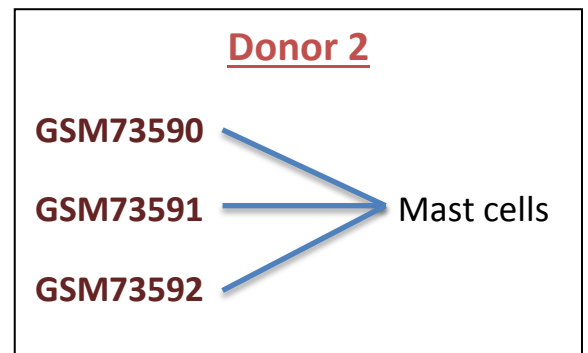
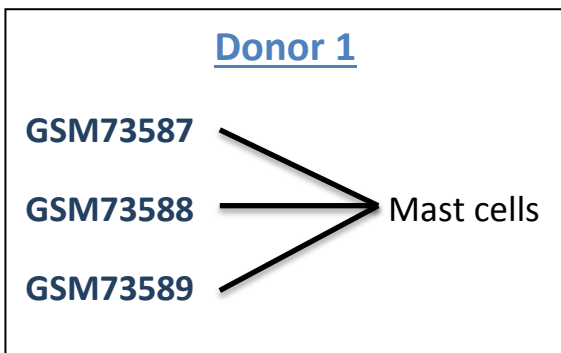


Step 12: Found out the ligands of a particular TLR which is commonly present on Both Mast cells and Natural killer cells.

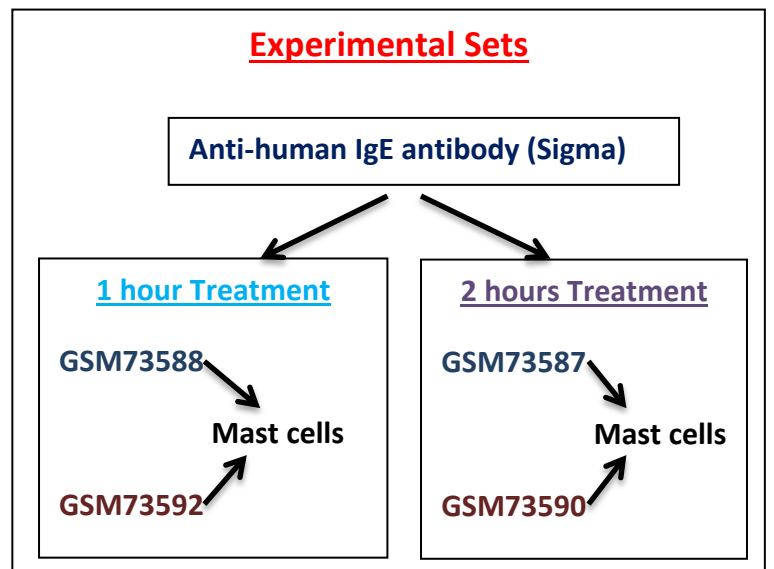
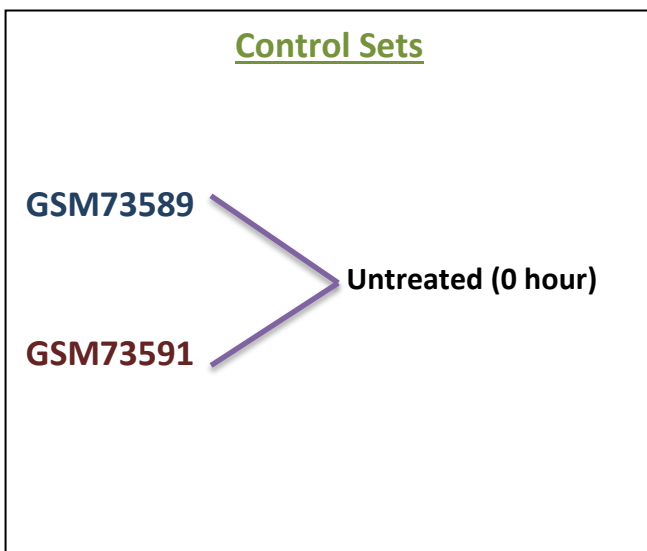
RESULTS AND DISCUSSION:

4.1 Stimulation condition of Mast cells Microarray data

Mast cells microarray data was collected from Gene expression omnibus datasets (GEO) with the help of GEO accession id (GSE3271). The substantial levels of histamine release, reflecting a high level of FcεRI-dependent mast cell activation, could be achieved by stimulating the mast cells with 10 or 1 μg/ml of anti-human IgE antibody (Sigma). Umbilical cord blood samples were taken from two different **Donors**.



Mast cells were stimulated with Anti-human IgE antibody treatment in the following manner.



Umbilical cord blood samples were taken from two different donors (Donor 1 and Donor 2). GSM73587, GSM73588, GSM73589 represent donor 1 Mast cells where GSM73587 cell was treated with anti-human IgE antibody (Sigma) for 2 hours and GSM73588 cell was treated with anti-human IgE antibody (Sigma) for 1 hours and GSM73589 was taken as a control. GSM73590, GSM73591, GSM73592 represent donor 2 Mast cells where GSM73590 cell was treated with anti-IgE for 2 hours and GSM73592 cell was treated with anti-IgE for 1 hours and GSM73591 was taken as a control.

The Gene expression data of non-stimulated Mast cells (for 0 hour) and stimulated Mast cells (with anti-human IgE antibody (Sigma) for 1 hour and 2 hours).

1	ID_REF	GSM73587	GSM73588	GSM73589	GSM73590	GSM73591	GSM73592	Gene.symbol
2	1	-0.213	-0.541	-0.242	-0.422	-0.358	0.049	MT1A
3	2	0.325	-0.044	0.17	0.196	-0.013	-0.133	MT1F
4	3	-0.246	-0.826	-0.489	-0.283	-0.267	0.056	MT1X
5	4	-1.555	-1.243	-1.234	-1.446	-1.245	-1.42	CSF2
6	5	-2.514	1.777	1.616	1.592	-3.571	-3.077	XCL1
7	6	-0.091	0.696	0.79	0.331	0.419	0.271	MT2A
8	7	1.655	1.841	1.873	1.565	1.375	1.715	BIRC3
9	8	-0.128	0.445	0.691	0.245	0.026	-0.014	TNF
10	9	1.202	2.091	1.844	1.684	1.294	1.029	IL21R
11	10	-0.206	0.084	-0.04	-0.317	0.171	0.091	C10orf54
12	11	0.686	0.052	0.317	0.04	0.348	0.298	NR4A3
13	12	0.904	0.814	0.766	1.082	0.3	0.617	PHLDA1
14	13	-0.94	-0.373	0.13	-0.975	-0.374	-0.726	CD72
15	14	-1.965	-2.22	-2.251	-2.398	-2.884	-2.66	IFNG

Fig1: Microarray Gene expression data (GSE3271) of Mast cells.

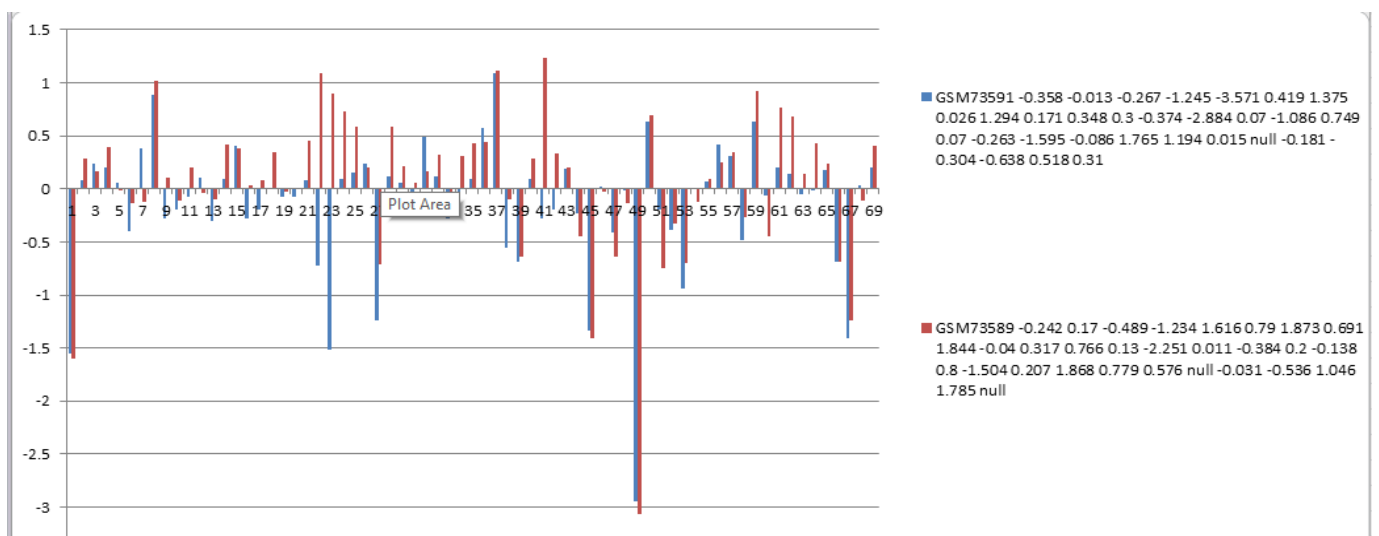


Fig2: Un-stimulated gene expression profile graph of mast cells

Gene expression profile graph was generated using un-stimulated mast cells of donor 1 and donor 2. Here two un-stimulated cells, such as GSM73589, GSM73591 were used and only 100 genes expression data were taken to make this graph because the graph would not be understandable in the case of large amounts of genes expression data. This graph represents the un-stimulated mast cells gene expression profile. The GSM73589 of donor 1 cell was untreated for 0 hour whereas GSM73591 of donor 2 cell was untreated for 0 hour.

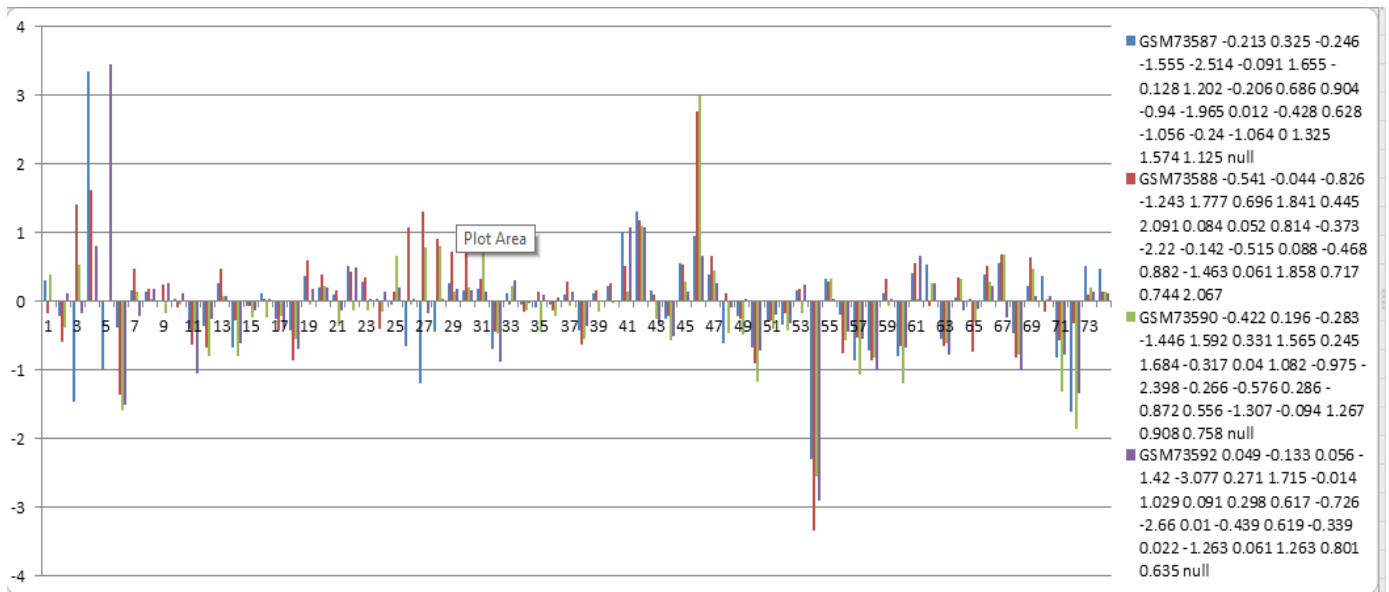


Fig3: Anti-IgE stimulated gene expression profile graph of mast cells

Gene expression profile graph was generated using anti-IgE stimulated mast cells of donor 1 and donor 2. Here four anti-IgE stimulated cells, such as GSM73587, GSM73588, GSM73590, GSM73592 were used and only 100 genes expression data were taken to make this graph. This graph represents the anti-IgE stimulated mast cells gene expression profile. The GSM73587 of donor 1 and GSM73590 of donor 2 cells were treated with anti-IgE for 2 hours whereas GSM73588 of donor 1 and GSM73592 of donor 2 cells were treated with anti-IgE for 1 hour.

Gene.symbol	Untreated (0 hour)		Anti-IgE treated for 1 hour		Anti-IgE treated for 2 hours	
	GSM73589	GSM73591	GSM73588	GSM73592	GSM73587	GSM73590
MYD88	-0.089	-0.019	0.348	0.144	0.395	0.561
IRAK4	0.851	0.159	-0.901	-1.775	-0.781	-0.457
TLR6	0.56	0.07	0.473	0.535	0.829	0.47
IRAK1	-1.197	-1.382	-0.829	-1.622	-2.536	-1.267
TRAF6	0.353	-0.009	0.547	0.627	0.036	0.584

Fig4: The Gene Expression data of Mast cells in different Experimental Conditions.

The above table represents each gene expression data in different Experimental Conditions. In case of MYD88 Gene –

Mast cells MYD88 Gene Expression Data					
Untreated (0 hour)		Anti-IgE treated for 1 hour		Anti-IgE treated for 2 hours	
GSM73589	GSM73591	GSM73588	GSM73592	GSM73587	GSM73590
-0.089	-0.019	0.348	0.144	0.395	0.561

4.2 Microarray data Normalization

R-Studio is C++ Programming based software which helps in executing R-Script. Using ‘Boxplot R-Script’ the Box-Plot of Genes expression data was created. From the Box-plot we can easily identify the Microarray genes expression data of Mast cells which was downloaded as a series matrix file (GEO accession id ([GSE3271](#))) is ‘Normalized’ or ‘Non-Normalized’.

```

library(limma)
library(readxl)
n <- read_excel("C:/Users/my/Desktop/t2.xlsx")
View(n)
names(n)
#k4$ID <- NULL
#k4$`Gene Title` <- NULL
#View(k4)
n2 <- n$`log FC`
n3 <- n$`Gene(20761)`
boxplot(n2)
quantile(n2, probs = c(0,0.25,0.5,0.75,1))
boxplot(n2, main = "Boxplot", ylab = "Nor.MEAN", las = 1)
boxplot(n2 ~ n3, main = 'Colon ipsc Boxplot by Genes')

```

Fig5: R-script was used to create Box-plot

Using the above 'R-script' we can easily create box-plot of gene expression data. Only expression data of the genes is required as an input. The above program read the 'Genes Expression' value and according to the value it plotted the graph. To see the graph which was plotted by using the 'R-script' and 'Genes Expression value', we can easily understand the genes expression data was biased or not. If the data is biased it means the data is non-normalized. So in that case first the expression data of genes was normalized by using different techniques (Global Normalization, Simple Normalization, Quantile Normalization) then use the gene expression data for further analysis.

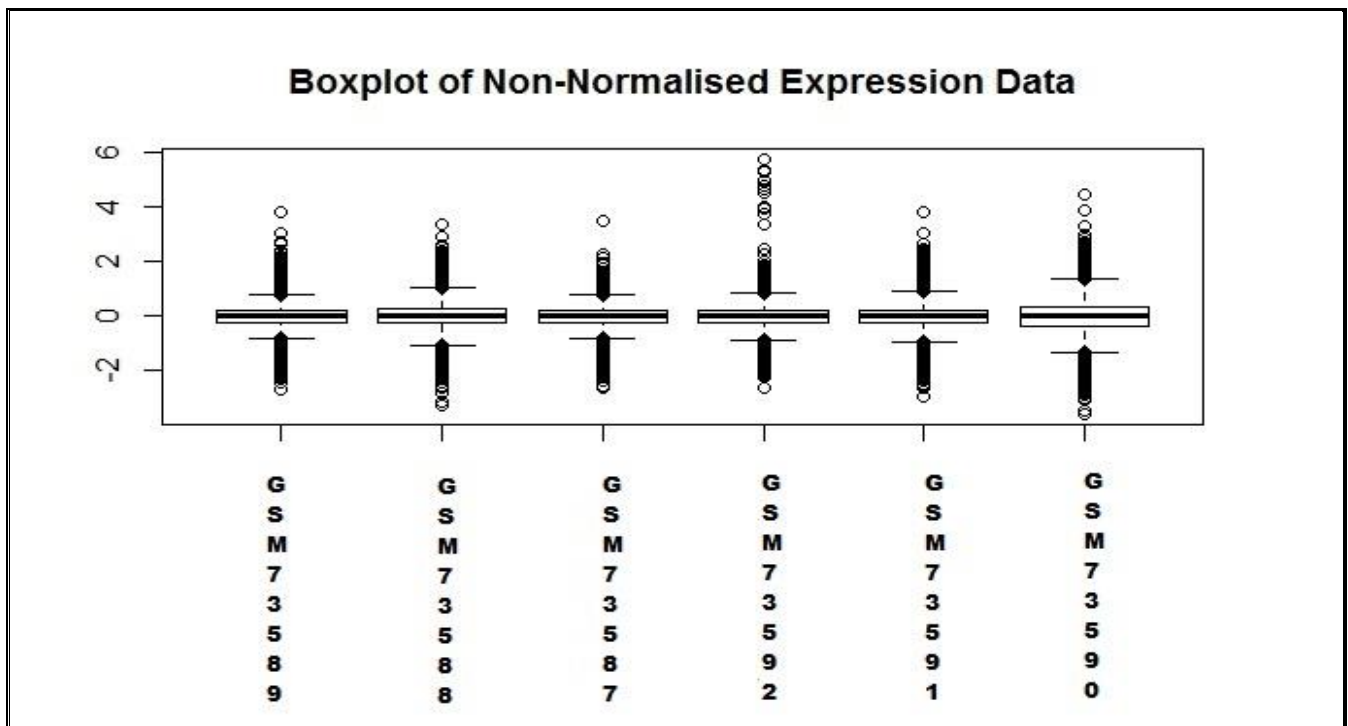


Fig6: In This Box-plot the 'X'-axis represents the Number of Cells and 'Y'-axis represents their respective Gene's Expression data of Mast cells.

The 6 different Cells containing Excel file was imported in R-Script and 'X' axis in this Box-plot represents the number of cells and in each cell among 6 cell files contained a different number of Genes and their respective Expression data. Here in this Box-plot the Expression data of same genes in different cells lied in different range according to the 'Y' axis range value. Due to the Non-Normalised data this kind of Box-Plot was generated.

R-Studio is a C++ Programming based software which helps in executing R-Script. After generating Box-plot, Q-Q plot was generated to recheck the genes expression data which was downloaded as a series matrix file from Gene Expression Omnibus, Normalised or Non-Normalised.

```
1 library(limma)
2 library(stats)
3 library(gplots)
4
5 setwd("C:/Users/my/Desktop")
6 library(limma)
7 #load("C:/Users/my/desktop/.xlsx")
8 a <- read_excel("C:/Users/my/Desktop/ani.xlsx")
9 View(a)
10 #ani$ID_REF__1 <- NULL
11 #View(a)
12 dim(a)
13 a[22906,20]
14 a[1:22906,1:20]
15 a[1,1]
16 #a[,1:2]
17 x = a$`Normalised(LogFC)`
18 hist(x,col = 'red')
19 hist(x,freq = FALSE, col = 'green')
20 xbar = mean(x)
21 s=sd(x)
22 curve(dnorm(x,xbar,s),col = 2,add = TRUE)
23 ones = c(xbar-s,xbar+s)
24 twos = c(xbar-2*s,xbar+2*s)
25 threes = c(xbar-3*s,xbar+3*s)
26 n=length(x)
27
28 ones
29 sum((x>ones[1])*(x<ones[2]))/n #percent within 1 sd
30
31 twos
32 sum((x>twos[1])*(x<twos[2]))/n #percent within 2 sd
33
34 threes
35 sum((x>threes[1])*(x<threes[2]))/n #percent within 3 sd
36
37 qqnorm(x)
38
```

Fig7: This R-script was used to create Q-Q Plot

Using the above ‘R-script’ we can easily create Q-Q plot of gene expression data. Here expression data of the genes was required as an input. The above program read the ‘Gene expression’ value and according to the value it plotted the graph. To see the graph which was plotted by using the ‘R-script’ and ‘Gene expression value’ we can understand the genes expression data was biased or not. If the data is biased it means the data is non-normalised. So in that case first the expression data of genes was normalised by using different techniques then used the gene expression data for further analysis.

The different Genes and their Respective Expression data containing Excel File were imported in R-Script. On the basis of Genes and their respective Expression data this Q-Q Plot was generated.

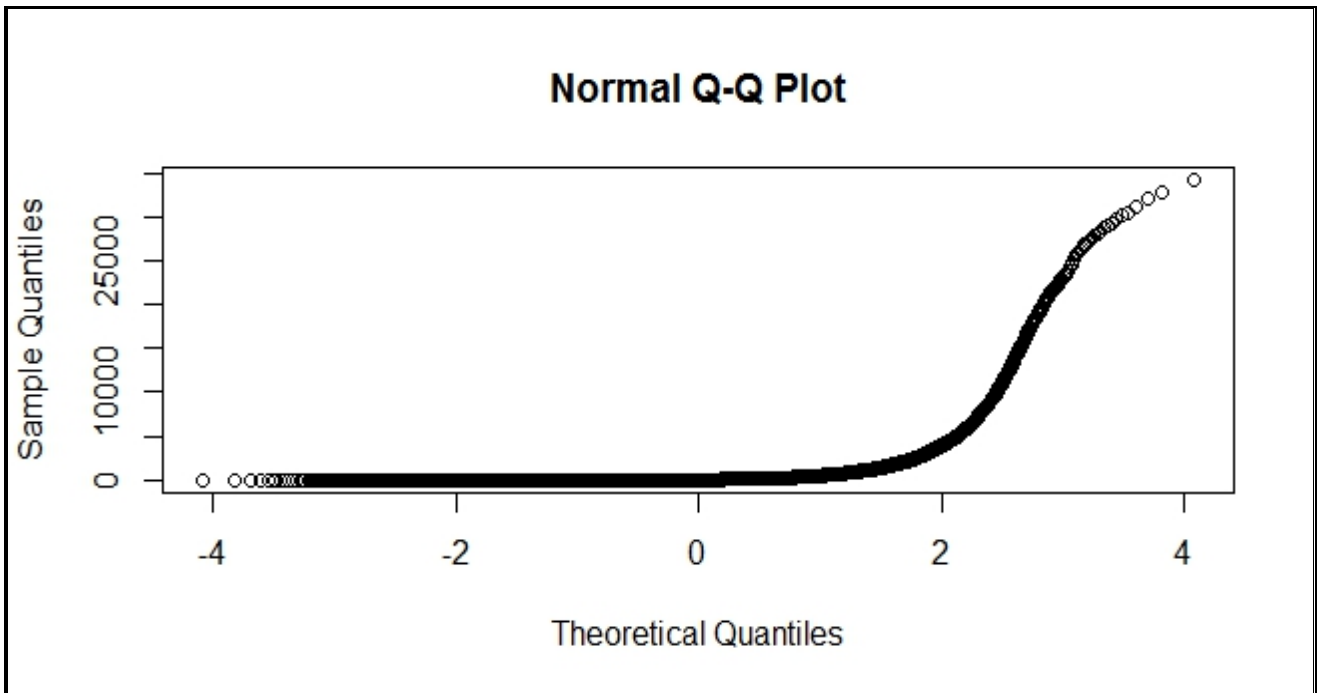


Fig8: In Non-Normalised Quantile Plot the ‘Y’-axis represents the number of Genes and ‘X’-axis represents their respective Fold Change Value of natural killer cells.

The Genes Expression data containing Excel file was imported in R-Script and ‘X’ axis in this Q-Q plot represents the range of gene expression data. Here in this Q-Q plot ‘Y’ axis represents the total number of genes Present in the excel file. Due to the Non-Normalised data this kind of Q-Q Plot was generated like as ‘S’.

So from the ‘**Q-Q plot**’ and ‘**Box plot**’ it was proved that the Series matrix file which was downloaded from Gene Expression Omnibus contained **Non-Normalised Expression data**. Many sources of systematic variation were occurred in microarray experiments. From this Q-Q plot and Box plot, it showed that in Genes Expression data there were lots of variations presented which can be affected the measured gene expression levels.

To Normalise the Gene Expression data first the formula was used to calculate the mean of the Gene Expression data –

Sample Mean
$\bar{x} = \frac{\sum x}{n}$

Fig9: General Formula of Mean

Where, $\sum X$ is the sum of all microarray Gene Expression data values of mast cells

n is the number of data items(Cell) in Sample

Normalisation

$= \frac{\text{(Expression value of Gene in particular cell)}}{\text{Mean of expression values of the same gene for all cells present in a single microarray data}}$
--

Fig10: General Formula of Simple Normalization

After importing the microarray gene expression data of mast cells in excel file which was downloaded as Mast cell activation via the Fc(epsilon)RI(**GSE3271**) from GEO datasets of National Centre for Biotechnology Information(NCBI) by analysing with GEO2R(A GEO Tool available for visualization of Microarray data along with other relevant calculated statistical parameters). Downloaded data was pasted into excel sheet with their respective gene expression values taken from series matrix file data. Now applying the above formulas the gene expression data was normalised on the basis of Simple Normalisation to convert our raw data to appropriate linearly arranged data.

After Normalisation with the help of above R-Script (**Fig5: ‘Box-plot R-script’**), the different Genes and their Respective Normalized Expression data containing Excel File was imported. On the basis of Genes and their respective Normalized Expression data this Box-Plot was generated.

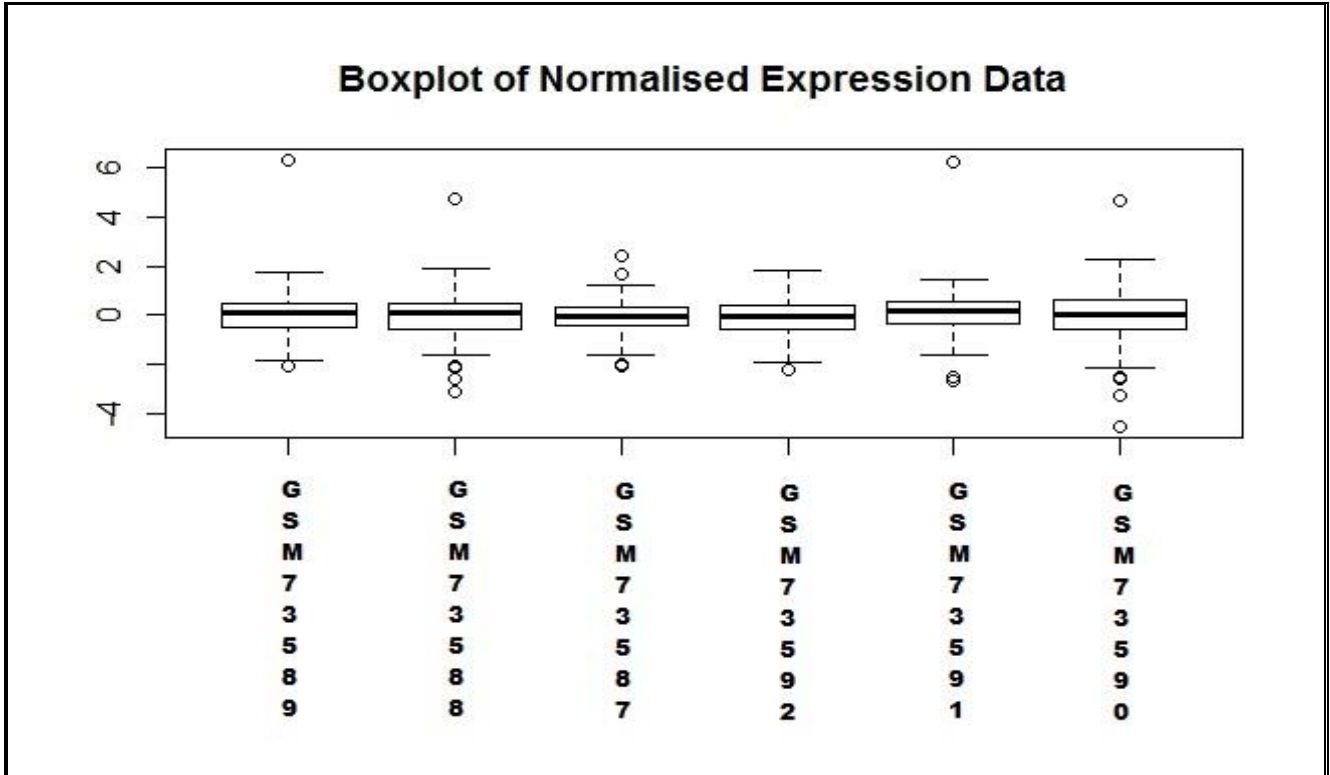


Fig11: In This Box-plot the ‘X’-axis represents the Number of Cells and ‘Y’-axis represents their respective Gene’s Expression value.

The Excel file which was imported in R-Script contained 6 different Cells which lie on ‘X’ axis in this Box-plot and in each cell among 6 cell files contained a different number of Genes and their respective Normalized Expression data. Here in this Box-plot the Expression data of same genes in different cells lied in almost proper range according to the ‘Y’ axis range value. Due to the Normalization variance of the gene expression data was removed data. So from the Box-plot it was proved that the Series matrix file which was downloaded from Gene Expression Omnibus contains Non-Normalised Expression data was successfully converted into normalised Gene Expression data.

After creating Box plot of Normalized genes expression data of Mast cells the different genes and their Respective Normalized Expression data containing Excel File was imported in 'Quantile R-Script'. On the basis of Genes and their respective Normalized Expression data this Q-Q Plot was generated by using the R-Script (Fig7: 'Quantile-plot R-script').

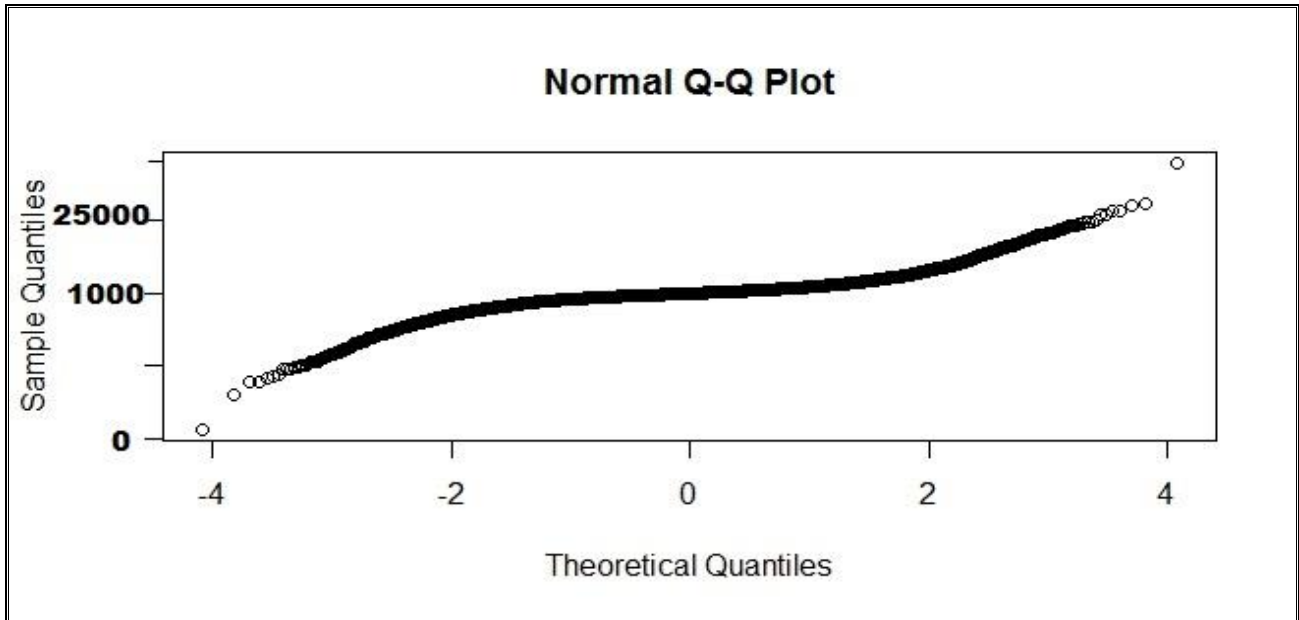


Fig12: In Normalised Quantile Plot the 'Y'-axis represents the Normalise range and 'X'-axis represents their respective expression data.

This Quantile plot was generated with the help of R-Script (Fig7: 'Quantile-plot R-script') after generation of Box-plot of the normalized Gene Expression data to re-check the Expression data was properly normalised or not.

In this Q-Q plot 'X' axis represents Genes Expression and 'Y' axis represents the total number of genes Present in the excel file. Due to the Normalization of gene expression data, proper Q-Q Plot was generated. So from the Q-Q plot it was proved that the Series matrix file which was downloaded from Gene Expression Omnibus contained Non-Normalised Expression data was converted into normalised Gene Expression data. Many sources of systematic variation were occurred in microarray experiments with the help of Simple Normalization (Statistical Methods) these kind of variations were removed.

After all, the Gene Expression data which was downloaded from GEO (Gene Expression Omnibus) is normalised and on the basis of Box-plot and Q-Q plot the Mast cells gene Expression data was almost normalised.

Importance of Normalization:

Normalization of Microarray data is aimed to correct for the systematic measurement errors and bias in the observed data. The errors and bias may be introduced by several factors such as difference in probe labelling, concentration of target DNA/RNA sequence and instrumental noise due to the scanners and printers etc. the process allows data to be compared across a common reference.

So, Normalization is able to increase significance of the fold change of a gene with respect to control.

4.3 Analysis of Fold change in Gene expression of Mast cells upon Stimulation

After Normalisation of Gene Expression data which was downloaded as series matrix file from GEO datasets (Gene Expression Omnibus).Median was calculated for control sets as well as disease sets of Normalised Gene Expression data. The purpose of calculating Median of Gene Expression data was to find out the Fold change value. Mainly to find out the Fold change value for a particular gene the median value of control sets and median value of disease sets was required.

Gene Symbol	GSM73587	GSM73588	GSM73589	GSM73590	GSM73591	GSM73592	Normalised Median(Test Sets)	NormalisedMedian(Control Sets)
EEF1A1	100.66578	96.671081	100.66578	100.66578	100.66578	100.66578	100.6657838	100.3922365
SLC35E2	0.0465072	0.0317376	0.0421514	0.0471427	0.0589472	0.026123	0.044329307	0.040385994
RPS28	6.6233285	6.1520731	12.853973	5.4818781	12.373914	11.6023	9.112816553	8.225188221
IPO13	0.0789275	0.1225142	0.0746936	0.1088028	0.1090002	0.10677	0.107786576	0.147360177
AFAP1	0.1583323	0.1402283	0.0702627	0.1347947	0.0588486	0.096581	0.115687901	0.026594486
IL8	0.0478154	0.0328053	0.0281501	0.0228644	0.0142086	0.027559	0.02785459	0.019418792
CDT1	1.5287518	1.6445764	1.4869543	1.4035985	1.222116	1.296556	1.445276418	1.638797505
TRPV1	0.0035901	0.0021906	-0.010721	-0.000789	-0.038066	-0.035343	-0.005755302	-0.014707519
LPP	1.9421829	2.1566881	1.8431785	1.7767873	1.6821697	1.505118	1.809982889	1.775798505
CCNE2	-0.02718	-0.031789	-0.023405	-0.028129	-0.016369	-0.011321	-0.025292402	-0.02275241
HNRNPAB	12.872904	12.011832	10.804735	15.44131	13.85195	13.88796	13.36242668	14.90880561
BORCS5	0.1332698	0.1390648	0.2109656	0.0914688	0.1263393	0.119108	0.129804534	0.103605577
SNIP1	1.7029637	1.6259522	1.7827869	1.4620094	1.5949253	1.512515	1.610438789	1.247229901
BCL6B	0.0153346	0.0684265	0.1349386	0.0738726	0.1081686	0.066306	0.071149533	0.116126304
GRHL1	0.2098783	0.2497323	0.0426898	0.2343513	0.0672747	0.058139	0.138576473	0.114468936
TRAF7	0.0230217	0.0003742	0.0100691	0.0314278	0.0097955	0.012714	0.011391626	0.00214128
ATP13A4	-0.002402	-0.025202	-0.010394	-0.019643	-0.015808	-0.030048	-0.017725699	-0.008057017
TRAF6	0.023945	0.0380803	0.0039988	0.0305046	0.0030927	-0.005255	0.013971927	0.000950548
HMG20B	2.7736426	2.7414056	2.536227	2.3689854	2.436132	2.325096	2.486179504	2.87644715
IRAK5	0.0015622	0.0076899	-0.016916	-0.019786	-0.01931	-0.030506	-0.018112836	-0.019858493

Fig13: The Median value of Control sets as well as Experimental sets was calculated after Normalization.

To calculate the median of Control sets as well as Disease sets the formula was used-

$\{(n + 1) \div 2\}^{\text{th}}$ value, where **n** is the number of values in a set of data.

In Fig: 13 the 'Green' colour column represents the control sets where the Mast cells were untreated for 0 hour and 'Blue' colour column represents the experimental sets where the Mast cells were treated with Anti-IgE for 1 hour and 'Red' colour column represents the another experimental sets where the Mast cells were treated with Anti-IgE for 2 hour.

After calculating the median of control sets and disease sets of Gene Expression data of Mast cells the Fold change value was calculated with the help of this formula-

$$FC_1(A/B) \sim = \text{Control Median (A)}/\text{Disease Median (B)}$$

Gene Symbol	MEDIAN(Control)	MEDIAN(Disease)	FC VALUE	Normalised(LogFC)
EEF1A1	100.3922365	100.6657838	1.002724786	0.00118175
SLC35E2	0.040385994	0.044329307	1.097640626	0.040460173
RPS28	8.225188221	9.112816553	1.107915869	0.044506783
IPO13	0.147360177	0.107786576	0.731449824	-0.13581546
AFAP1	0.026594486	0.115687901	4.350070986	0.638496344
DMD	0.019418792	0.02785459	1.434414124	0.156674553
CDT1	1.638797505	1.445276418	0.881912752	-0.054574378
TRPV1	-0.014707519	-0.005755302	0.39131696	-0.407471329
LPP	1.775798505	1.809982889	1.019250148	0.008280783
CCNE2	-0.02275241	-0.025292402	1.111636148	0.045962661
HNRNPAB	14.90880561	13.36242668	0.896277477	-0.047557517
BORCS5	0.103605577	0.129804534	1.252872071	0.097906728
SNIP1	1.247229901	1.610438789	1.29121246	0.110997708
BCL6B	0.116126304	0.071149533	0.612690928	-0.21275855
GRHL1	0.114468936	0.138576473	1.210603315	0.083001859
BASP1	0.00214128	0.011391626	5.320008826	0.725912353
ATP13A4	-0.008057017	-0.017725699	2.200032521	0.342429101
RESP18	0.000950548	0.013971927	14.6988143	1.167282303
HMG20B	2.87644715	2.486179504	0.864323026	-0.063323917
CCT3	-0.019858493	-0.018112836	0.912095189	-0.039959835

Fig14: From Normalised data of Mast cells the Fc value is calculated with the help of above Formula.

The Fold change value was calculated because to know how much a particular gene is over-expressed from control condition to disease condition. Let's say for Gene A the expression value in control set was 50 and in disease set the expression value of the same gene was 100. Then the fold change of Gene A was $100/50=2$. This works well for over expressed genes as the number directly corresponded to how many times a gene was over expressed.

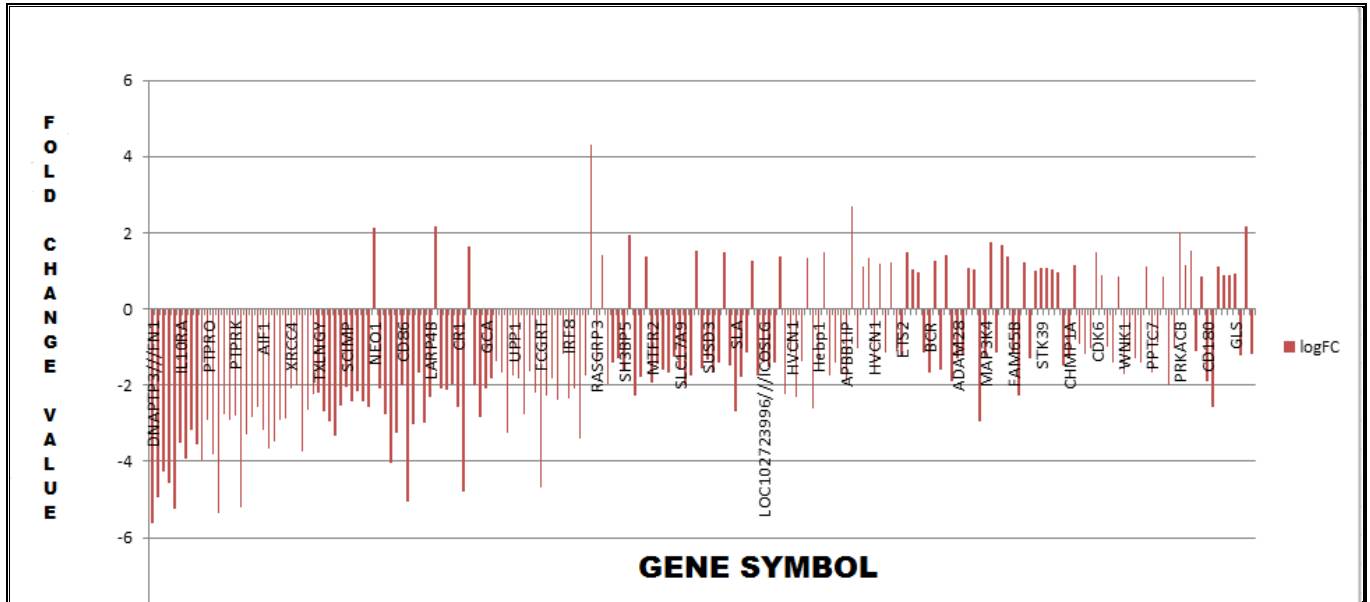


Fig15: Graphical view of Fold change of Mast cells Genes.

4.4 Setting of Threshold Value for Mast Cells Expression Data

After calculating the Fold change from the median value of control sets and experimental sets of Mast cells Gene Expression data, the Threshold value was calculated with the help of Quantile function. Here R-Script was used to perform the Quantile function.

```
library(limma)
library(readxl)
n <- read_excel("C:/Users/my/Desktop/T.xlsx")
n2 <- n$logFC
boxplot(n2)
quantile(n2, probs = c(0, 0.25, 0.5, 0.75, 1))

      0%      25%      50%      75%      100%
-0.07770  0.04715  0.19300  0.46425  5.62000

Threshold : 1.3343
```

Fig16: Threshold value was calculated from the Fc value using Quantile Function.

Quantile function arranges the whole Fold change values of in order of lowest to highest value (Such as 0%, 25%, 50%, 75%, 100%). Then on the basis of 0%, 25%, 50%, 75%, 100% it plots the graph of fold change values. Then from the graph it selects only that value which is present in the middle of the graph. This 'Quantile function' takes this value as Threshold. Here in case of mast cells the threshold value was - 0.0776. To find out the threshold value all of these steps were performed using R-Script. On the basis of this threshold value the Fold change value was sorted and with the fold change values of Mast cells their respective Genes was also selected. Beyond this threshold value the Genes were called Down-regulated genes and above this threshold value the genes were called Up-regulated genes. Here only up-regulated Mast cells genes were selected.

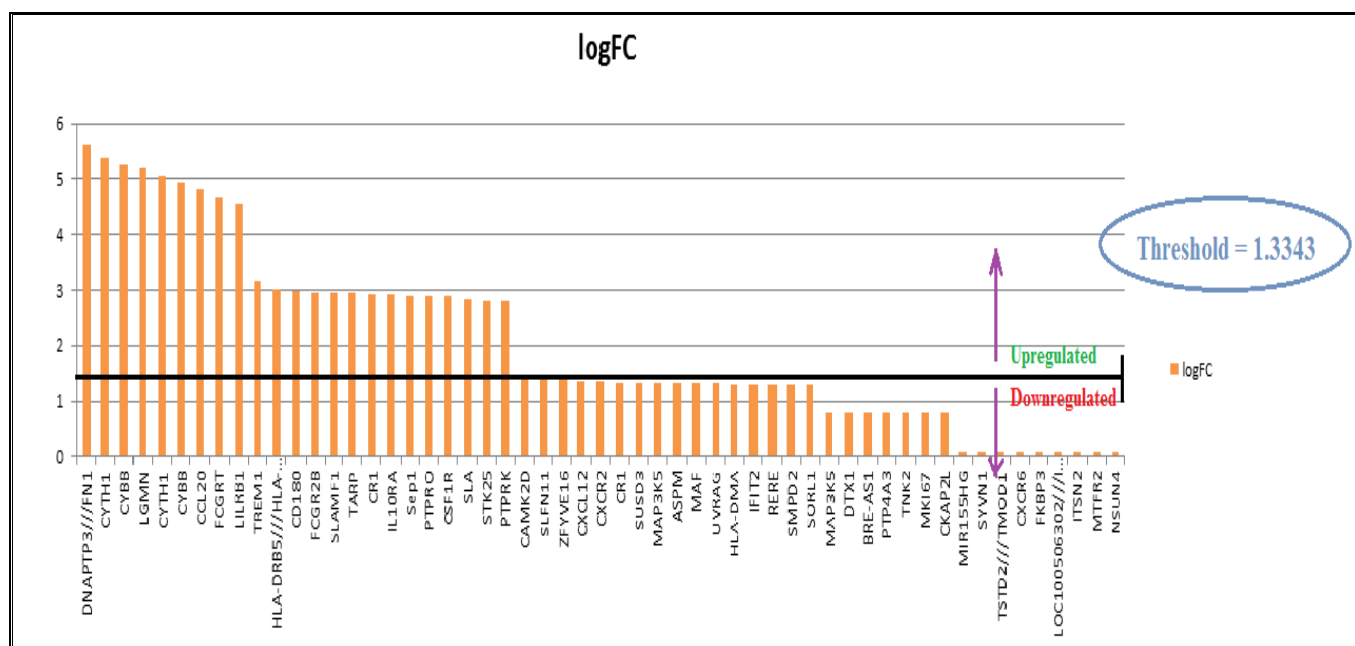


Fig17: Threshold value has been applied on Mast cells genes Fold change values.

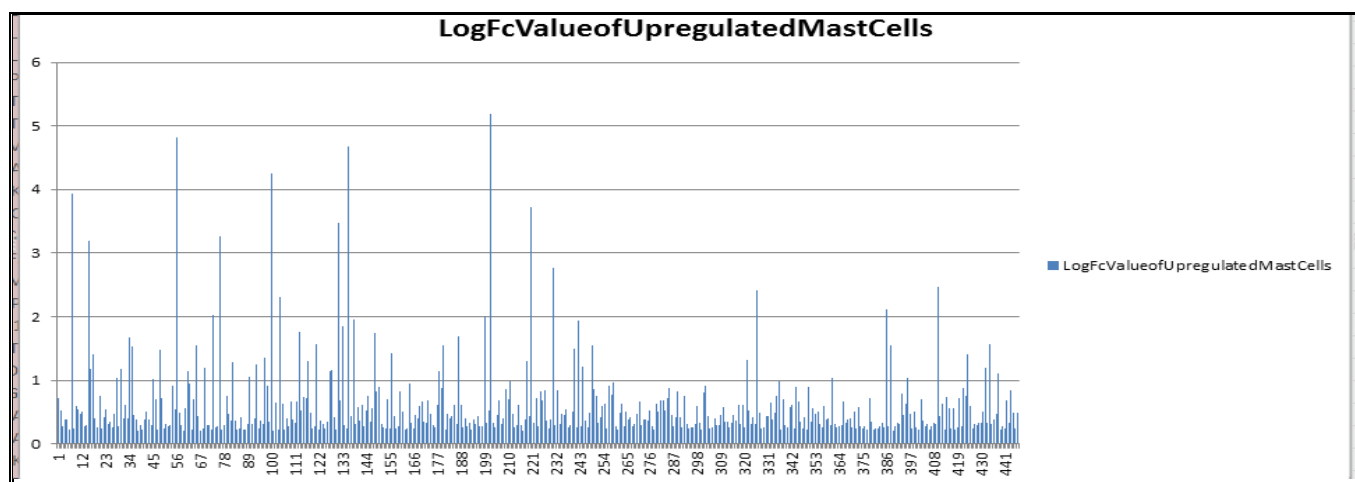
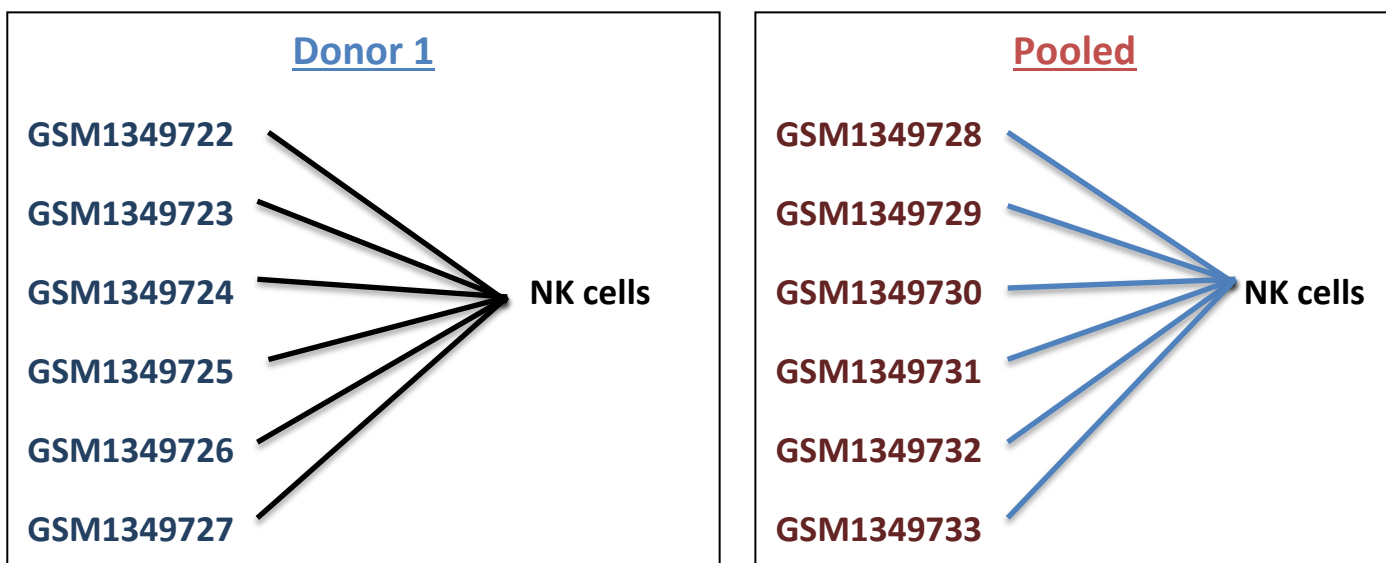


Fig18: Graphical view of Fold change of up-regulated Genes present in mast cells.

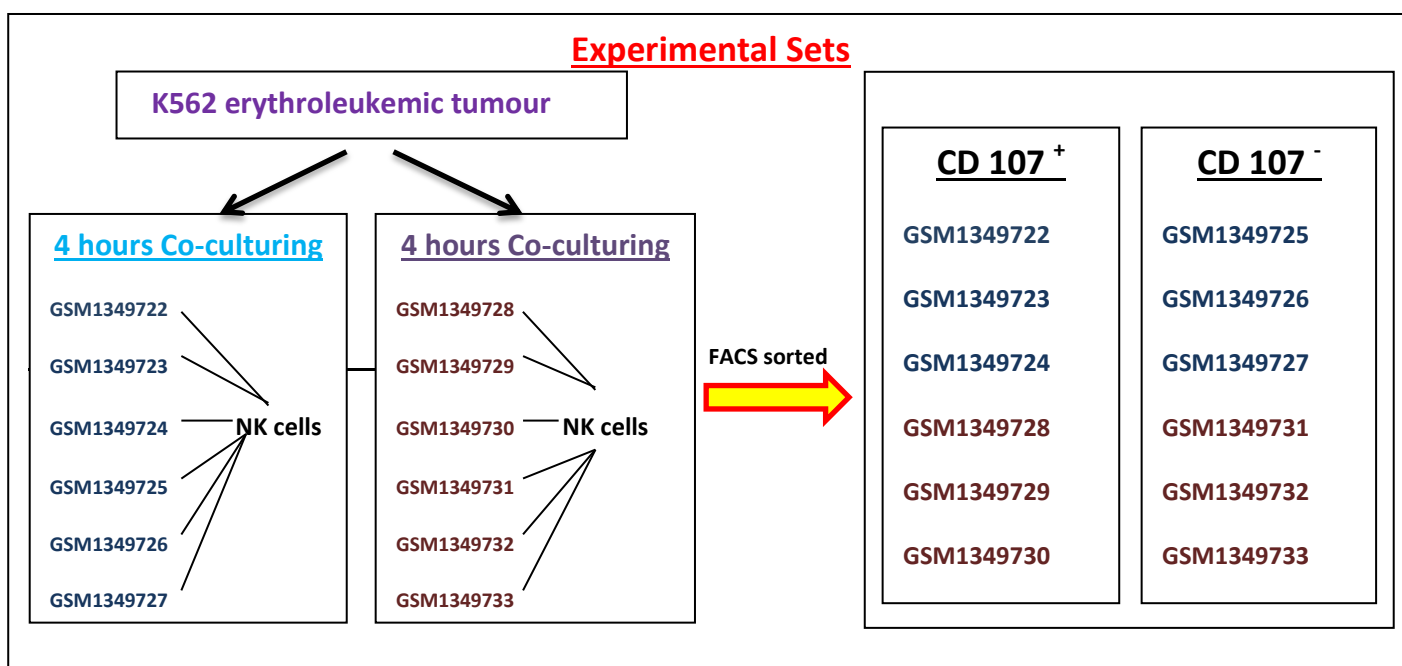
Stimulation condition of Natural Killer cells Microarray data:

Natural killer cells microarray data was collected from Gene expression omnibus datasets (GEO) with the help of GEO accession id (**GSE55977**). Genome wide expression profiling of human NK cells stimulated with K562 erythroleukemic tumour cells after four hours of NK-tumour co-culture. Responding NK cells were compared to non-responding NK cells, represented by display of CD107 on the NK cell surface following cytotoxic granule release.

Human peripheral blood NK cells were taken from Pooled and Donor1 and NK cells were co-cultured with tumour target cell line K562 for 4 hours.



NK cells were stimulated with K562 erythroleukemic tumour cells treatment in the following manner.



The same procedures like

- 1) Box plot generation of Non- Normalised genes expression data of NK cells.
- 2) Q-Q plot generation of Non- Normalized genes expression data of NK cells.
- 3) Simple Normalization (Fig10 Formula was applied) done to normalize the NK cells genes expression data.
- 4) Box plot generation of Normalised genes expression data of NK cells.
- 5) Q-Q plot generation of Normalized genes expression data of NK cells.
- 6) Analysis of Fold change in genes expression of NK cells upon stimulation.
- 7) Setting of Threshold for NK cells expression data.

Have been followed in case of Natural killer cells microarray data Normalization, Fold change value calculation and setting of Threshold.

```
library(limma)
library(readxl)
n <- read_excel("C:/Users/my/Desktop/NK.xlsx")
n2 <- n$`logFC`
boxplot(n2)
quantile(n2,probs = c(0,0.25,0.5,0.75,1))

  0%      25%      50%      75%      100%
-0.0461  0.0831  0.2360  0.5780  9.7600

Threshold : 2.1222
```

Fig19: Threshold value was calculated from the Fc value using Quantile Function.

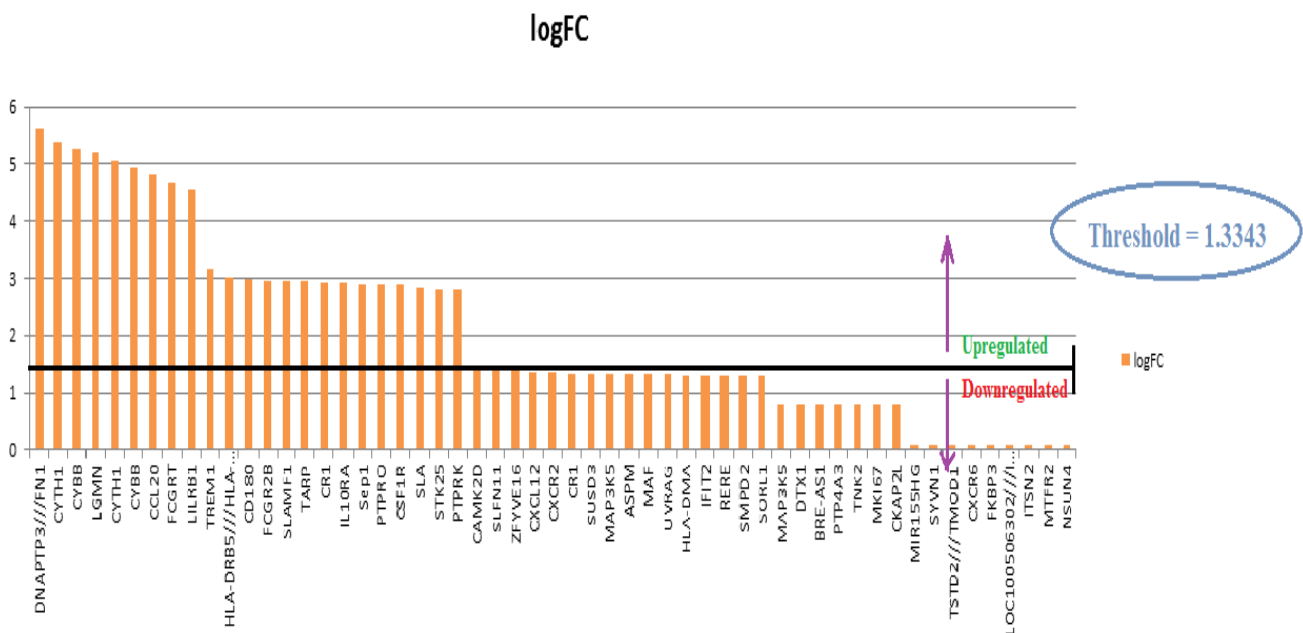


Fig20: Threshold value has been applied on NK cells genes Fold change values.

LogFcValueOfUpregulatedNk Cells

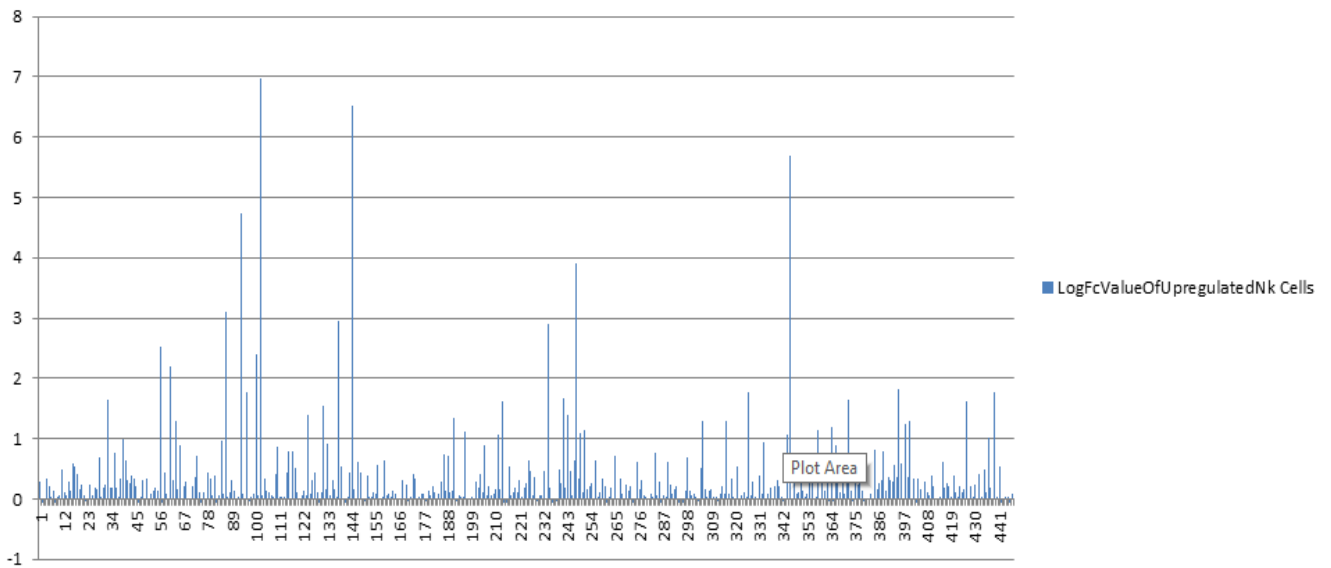


Fig21: Graphical view of Fold change value of up-regulated Genes present in natural killer cells.

4.5 Analysis of Common Genes between Mast cells and NK cells

After sorting the up-regulated genes of mast cells and natural killer cells a java program was developed to find out the common up-regulated genes between mast cells and natural killer cells.

```
package javaapplication77;

import java.io.FileInputStream;
import java.io.FileNotFoundException;
import java.io.IOException;
import javax.swing.JFileChooser;
import org.apache.poi.xwpf.extractor.XWPFWordExtractor;
import org.apache.poi.xwpf.usermodel.XWPFDocument;

public class GenesComparison {

    String Store,String Store1;

    public void GeneList1() throws IOException{

        int i; String str1;String str2;int count5=0; int count6=0;

        try{

            JFileChooser Chooser = new JFileChooser();

            int returnvalue = Chooser.showOpenDialog(null);

            if(returnvalue==JFileChooser.APPROVE_OPTION){

                XWPFDocument document = new XWPFDocument(new FileInputStream(Chooser.getSelectedFile()));

                XWPFWordExtractor extract = new XWPFWordExtractor(document);

                Store = extract.getText();

                for(i =0; i<=170;i++){

                    str1=Store.split("\n")[i];

                    str2= str1;    count5++;

                    count6=count5;

                } }

        }catch(FileNotFoundException jk){ }catch(Exception ml){}}
```

```

public void GeneList2() throws IOException{
    int j; int count1=0; String str3; String str4;
    try{
        JFileChooser Chooser = new JFileChooser();
        int returnvalue = Chooser.showOpenDialog(null);
        if(returnvalue==JFileChooser.APPROVE_OPTION){
            XWPFDocument document = new XWPFDocument(new FileInputStream(Chooser.getSelectedFile()));
            XWPFWordExtractor extract = new XWPFWordExtractor(document);
            Store1=extract.getText();
        }
    }catch(FileNotFoundException jk){

    }catch(Exception fg){

    } }

    String Str1; String Str2;String Str3; String Str4; String Str5;
    int i; int j; int Count=0; int Count1=0;int Count2=0; int y=0; String CommonGene="A"; int count=0; int count1=0; int k;
    public void CompareGenes(){
        try{
            new Thread(){
                public void run(){
                    for(;;){
                        for(i=0;i<=170;i++){
                            Str1=Store.split("\n")[i];
                            Str2=Str1;
                            for(j=0;j<=34107;j++){
                                Str3=Store1.split("\n")[j];
                                Str4=Str3;
                                if(Str2.equals(Str4)){

```

```

        Count++;

        System.out.println(Str2+" MATCHED "+Str4);

        for(k=j+1;k<=34107;k++){

            Str5=Store1.split("\n")[k];

            if(Str4.equals(Str5)){

                count++;

            }else{ } }else{

            } Count1++;}

        Count2++;

        System.out.println(Count+" = Number Of Total matched Gene");

        System.out.println(count+" = Number Of Duplicate matched Gene");

    }

    System.out.println(Count+" = Number Of Total Matched Gene");

    System.out.println(count+" = Number Of Duplicate Matched Gene");

} }

}.start();

}catch(Exception hj)} } }

public static void main(String []args) throws IOException{

    GenesComparison m = new GenesComparison();

    m.GeneList1();

    m.GeneList2();

    m.CompareGenes();

}}

```

Fig22: Identification of Up-regulated and Down-regulated genes common to Mast cells and NK cells Signalling pathway.

To find out the common genes between mast cells and natural killer cells in java program first read the excel file which contained the up-regulated genes list. To read excel file XWPFDocument package was imported and after reading both excel file the Genes list was stored in a string variable 'Store' which contained mast cells up-regulated genes and another string variable 'Store1' which contained natural killer cells. For extracting the genes list from excel file XWPFWordExtractor package was imported. After extracting the gene list, it was stored in 'Store' and 'Store1' variables.

CompareGenes() method was created in java program to compare the genes between up-regulated mast cells and up-regulated natural killer cells. If Common genes were found between two variables then it printed the name of the gene which was matched. In CompareGenes() method Split() method was used to split when it found a new line then it stored string which contained one gene name into 'Str1' variable. This 'Str1' variable was present in a 'for loop' because a large number of genes were present in 'Store' variable. This 'for loop' ran continuously still the last gene came then it entered in second 'for loop' which was for 'Store1' variable which contained natural killer cells genes list and here also split() method was used to store a single gene at a time in 'Str3'. Then in this second 'for loop' one condition was applied if Str2 equals to Str4 then program printed the matched gene name. Here 'equals ()' method was used for matching the gene name. After the complete run of the program, we got the matched up-regulated genes between mast cells and natural killer cells.

After running java program 447 up-regulated common genes were found between mast cells and natural killer cells.

UpregulatedMatchedGenes	LogFcValueOfUpregulatedNk Cells	LogFcValueofUpregulatedMastCells
ABCA1	0.297	0.717
ABL1	-0.0539	0.52
ABL2	-0.0693	0.277
ACP2	0.33	0.379
ACTB	0.214	0.383
ACTG1	0.0422	0.231
ACVR1	0.132	3.94
ADAM28	-0.0557	0.246
ADK	0.0251	0.591
ADORA2A	0.0545	0.539
AES	0.49	0.473
AFF2	0.114	0.504
AFMID	0.0552	0.281
AGPAT2	0.299	0.288
AIF1	0.136	3.19
AKT3	0.589	1.17
ALDH5A1	0.547	1.4
ALG10B	0.402	0.407
ANAPC5	0.154	0.255
ANAPC7	0.241	0.745
ANKRD49	0.0699	0.234
APTX	-0.0336	0.409
ARHGAP11A	-0.017	0.535
ASB13	0.234	0.309

Fig23: 447 Up-regulated Genes were found after Running Java Program, Common between Mast Cells and NK cells.

In this diagram the common genes and their respective Fold change values in mast cells and natural killer cells was retrieved with the help of java program. So from this diagram we could easily detect a particular gene's Fold change value in natural killer cells and also in mast cells. TLR2, TLR6, MYD88, IRAK1, IRAK4, TRAF6, IRF5 genes were found in common gene list. It means that these

following genes are common between mast cells and natural killer cells and these following genes are involved in MYD88 pathway.

4.6 Study of Common Genes in Interacting patterns of Mast cells and NK cells

From 447 common up-regulated genes between Mast cells and NK cells, only 8 genes were selected because (TLR2, TLR6, MYD88, TRAF6, IRAK1, IRAK4, IRF5 and IL-8) these 8 genes are present in MYD88 pathway. Multiple proteins list such as TLR2, TLR6, MYD88, TRAF6, IRAK1, IRAK4, IRF5 and IL-8 was uploaded in String database to check the interactions between these genes.

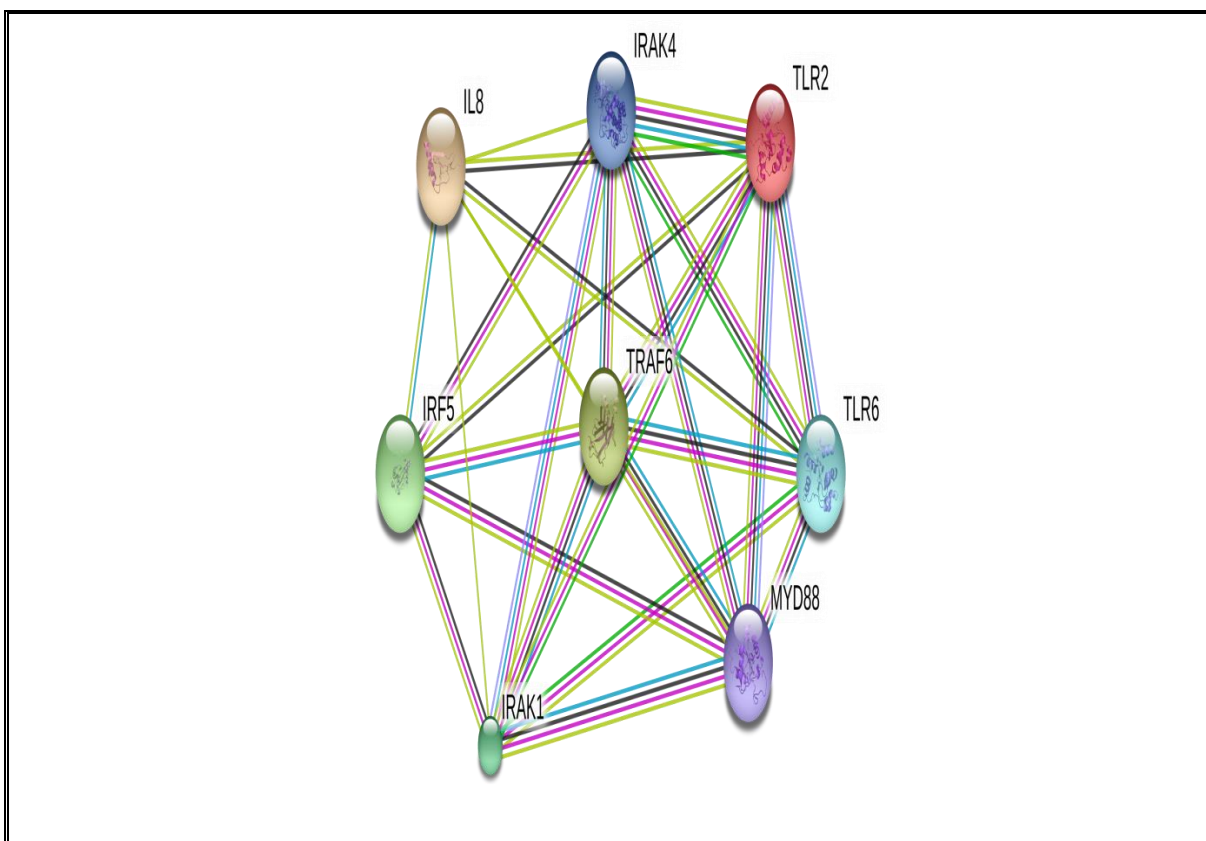


Fig24: These Common Proteins are interacted with each other and present in MyD88-dependent Pathway.

Both TLR2 and TLR6 interact with MYD88 and then MYD88 also interacts with IRAK1 and IRAK4. MYD88 strongly interacts with IRAK4. Then IRAK4 and IRAK1 interact with TRAF6. This TRAF6 interacts with IRF5. Then IRF5 transcribes DNA and IL-8 is produced.

1	Protein1	Protein2	Protein1_external_id	Protein_external_id	Interaction_score
2	MYD88	TRAF6	9606.ENSP0000040139	9606.ENSP000003378	0.999
3	MYD88	IRAK4	9606.ENSP0000040139	9606.ENSP000003906	0.999
4	IRAK1	TRAF6	9606.ENSP0000035899	9606.ENSP000003378	0.998
5	MYD88	IRAK1	9606.ENSP0000040139	9606.ENSP000003589	0.997
6	MYD88	TLR6	9606.ENSP0000040139	9606.ENSP000003713	0.996
7	IRAK4	TRAF6	9606.ENSP0000039065	9606.ENSP000003378	0.986
8	IRAK4	IRAK1	9606.ENSP0000039065	9606.ENSP000003589	0.982
9	MYD88	TLR2	9606.ENSP0000040139	9606.ENSP000002600	0.979
10	IRAK4	TLR2	9606.ENSP0000039065	9606.ENSP000002600	0.975
11	TLR6	TLR2	9606.ENSP0000037137	9606.ENSP000002600	0.974
12	TRAF6	TLR2	9606.ENSP0000033785	9606.ENSP000002600	0.973
13	TLR6	TRAF6	9606.ENSP0000037137	9606.ENSP000003378	0.968
14	IRF5	TRAF6	9606.ENSP0000034977	9606.ENSP000003378	0.933
15	IRAK1	TLR2	9606.ENSP0000035899	9606.ENSP000002600	0.916
16	IL8	TLR2	9606.ENSP0000030651	9606.ENSP000002600	0.904
17	IRF5	IL8	9606.ENSP0000034977	9606.ENSP000003065	0.855
18	MYD88	IRF5	9606.ENSP0000040139	9606.ENSP000003497	0.849
19	MYD88	IL8	9606.ENSP0000040139	9606.ENSP000003065	0.843
20	IRAK1	IL8	9606.ENSP0000035899	9606.ENSP000003065	0.751
21	TLR6	IRAK1	9606.ENSP0000037137	9606.ENSP000003589	0.711

Fig25: Protein-Protein Interaction and Their Interaction Score.

In above table the interaction between proteins and their respective ‘ENSP’ id is shown and for a particular interaction between two proteins the interaction score is also shown. Most of the Protein-Protein interaction, the interaction score of is very high. The interactions of these proteins are shown because these proteins are present in MYD88-dependent pathway in natural killer cells. Through this pathway various kind of inflammatory cytokines are secreted such as IL-8, IL-12, and IL-2 etc. These cytokines can activate Natural killer cells and can change multiple cellular responses such as cytokine/chemokine production, cytotoxicity (source: **Genomic expression profiling of NK cells in health and disease Fuyan Wang^{1,2}, Zhigang Tian¹ and Haiming Wei¹ 2015**).

4.7 Design a Mast cell TLR mediated pathway

The genes such as TLR2, TLR6, MYD88, TIRAP, IRAK1, IRAK4, TRAF6 and IRF5 are common between mast cells and natural killer cells. From the microarray data of mast cells activation via Fc (epsilon) RI (GEO accession id is GSE3271) and gene expression analysis of tumour stimulated human natural killer cells (GEO accession id is GSE55977) these genes are found out. These genes are present in TLR receptor mediated natural killer cell pathways and the interactions between all of these genes are found in MYD88-dependent pathway.

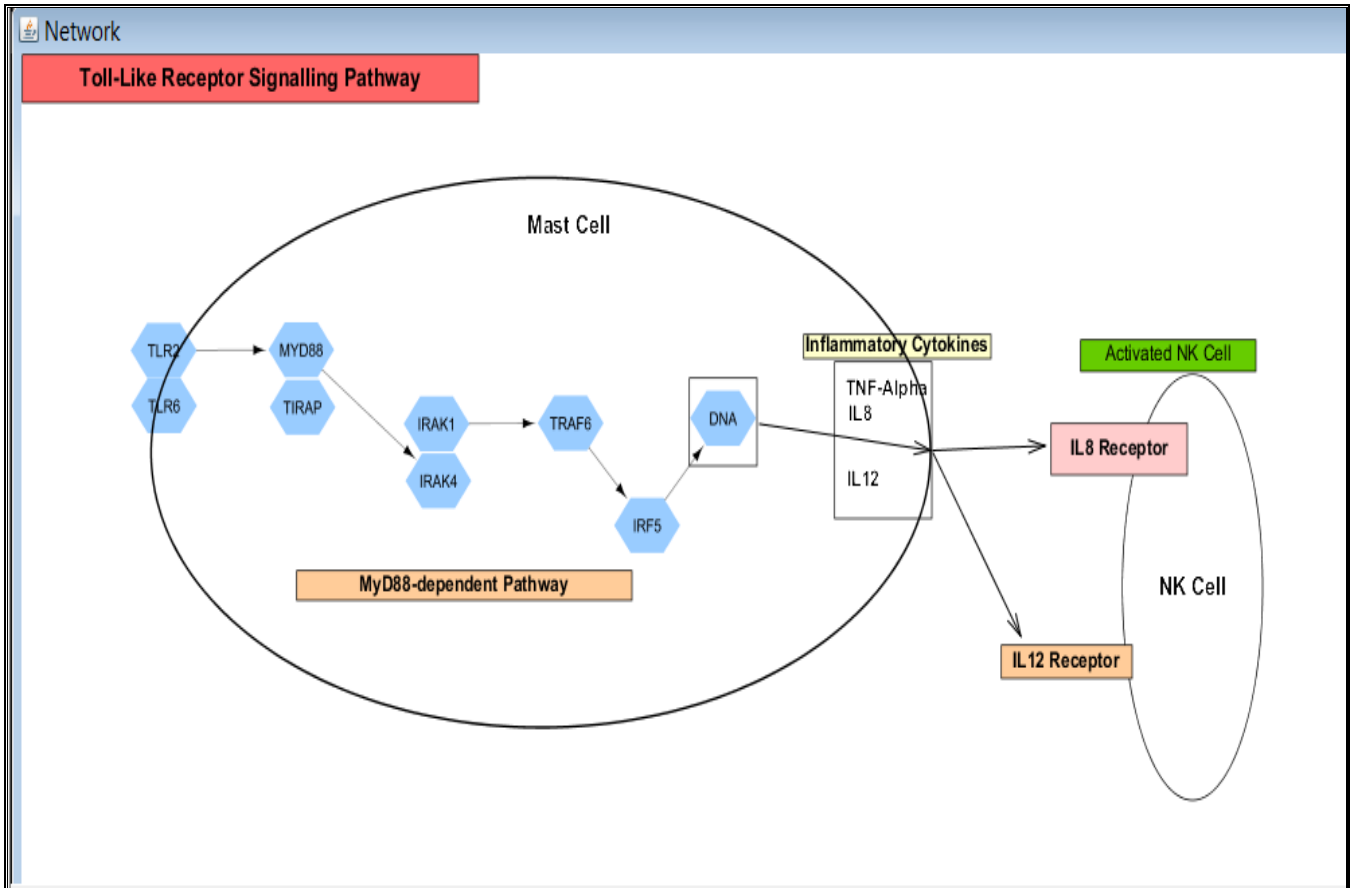


Fig26: MyD88-dependent Pathway Mast cells Produce Inflammatory Cytokines

If these following genes are common between mast cells and natural killer cells and if these genes are involved in TLR receptor mediated killer cells signalling pathway, then these genes are also involved in mast cells TLR receptor mediated pathway and in mast cells through this signalling pathway different kinds of inflammatory cytokines are released such as IL-8, IL-2 etc. These cytokines can also activate the natural killer cells.

	A	B	C	D	E	F	G	H
1	ID	adj.P.Val	P.Value	t	B	logFC	Gene.symbol	Gene.title
2	68	0.831536	5.26E-01	6.68E-01	-6.87572	1.59E-01	IL12A	interleukin 12A
3	648	0.870048	6.08E-01	5.37E-01	-6.9613	1.19E-01	IL12B	interleukin 12B
4	4675	0.870777	6.09E-01	5.36E-01	-6.96216	1.21E-01	IL2	interleukin 2
5	13889	0.974839	9.07E-01	1.21E-01	-7.11408	5.50E-02	IL4	interleukin 4

Fig27: Microarray data of Mast cells where Interleukins are released during Mast cells activation

	A	B	C	D	E	F	G	H
1	ID	adj.P.Val	P.Value	t	B	logFC	Gene.symbol	Gene.title
2	ILMN_1684349	2.02E-01	2.29E-02	-2.63127	-3.75418	2.45E-01	IL2RB	interleukin 2 receptor subunit beta
3	ILMN_2108699	5.09E-01	1.64E-01	-1.52617	-5.36177	1.89	IL2RA	interleukin 2 receptor subunit alpha
4	ILMN_1652185	7.49E-01	4.33E-01	-0.81252	-6.28608	2.74E-01	IL4R	interleukin 4 receptor

Fig28: Microarray data of Natural killer cells where Interleukins receptors are expressed in natural killer cells

So from the microarray data of mast cells activation via Fc (epsilon) RI (GEO accession id is GSE3271) cells the Fold change value of IL2, IL4 and gene expression analysis of tumour stimulated human natural killer cells (GEO accession id is GSE55977) the Fold change value of IL12A, IL12B, IL2RB, IL2RA and IL4R is taken. If TLR2-TLR6 MYD88 pathway mediated pathway is present in mast cells then IL-2, IL-8, IL-4 interleukins (Fig 17) are released from the mast cells when Interferon regulatory factor 5 (IRF5) binds with DNA (MYD88-dependent pathway) and with the help of these interleukins the natural killer cells are activated because the fold change value of IL12A, IL12B, IL2RB, IL2RA and IL4R interleukin receptors (Fig 18) are much higher in natural killer cells. So IL2 receptor beta or alpha which presents on natural killer cells receives IL2 and IL4 receptor which presents on natural killer cells receives IL4 which are secreted from mast cells. It is the possible way through which natural killer cells are activated by mast cells.

4.8 Analysis of Mast cells degranulation and NK cells receptor pathway for prediction of synergistic role of Mast and NK cells in Tumour Regression

These following pathways were exported from KEGG pathway.

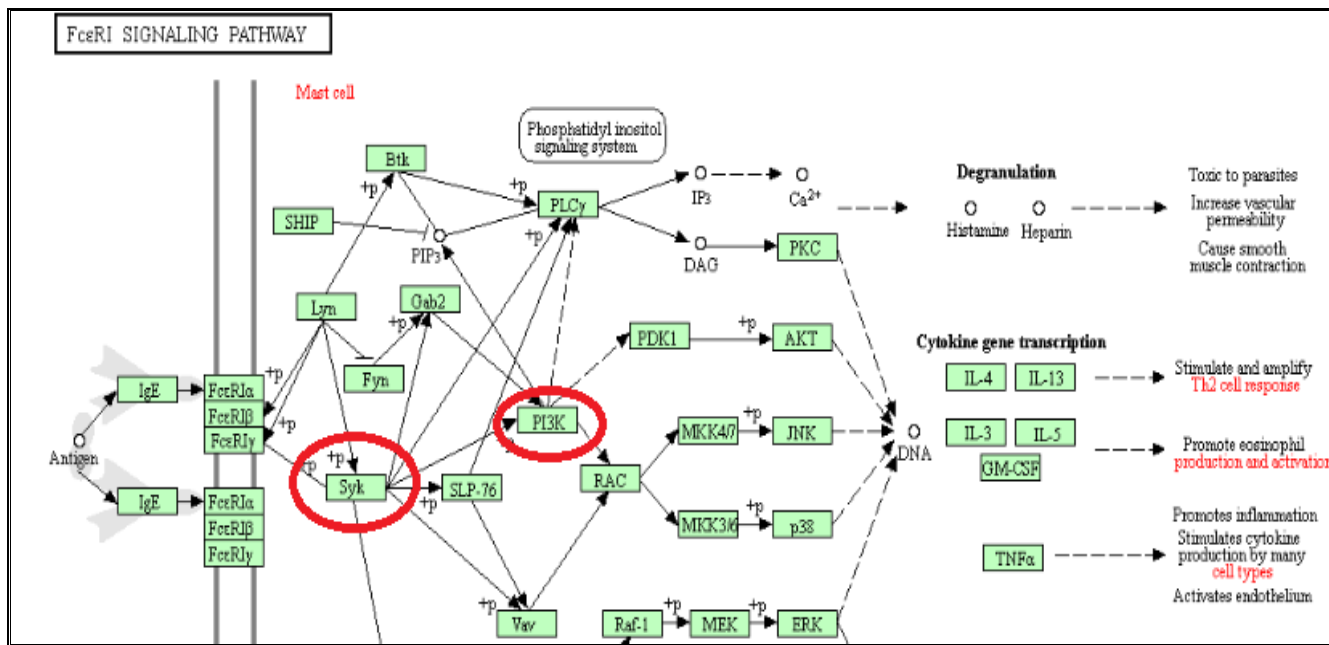


Fig29: Mast cells Degranulation pathway

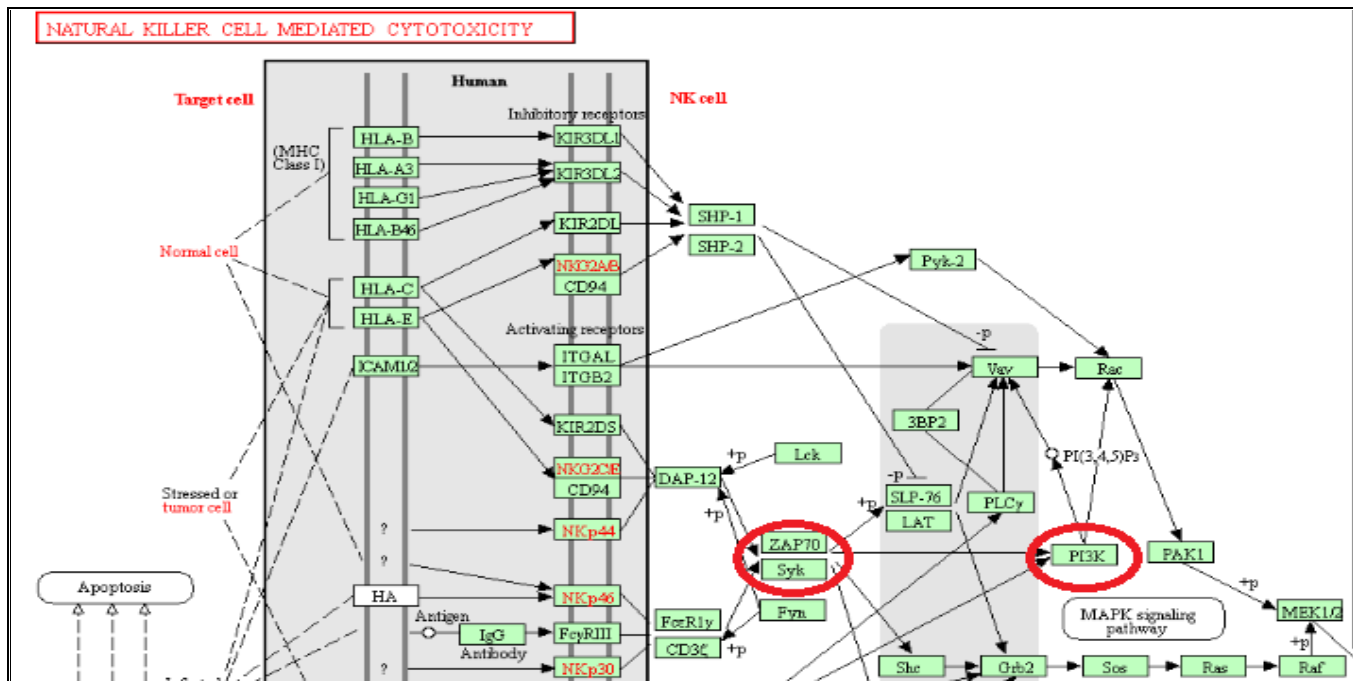


Fig 30: Natural killer cells receptor mediated pathway

In case of mast cells degranulation pathway (Histamine releasing) when IgE binds with FcεRI that time FcεRI phosphorylates syk which then activates PI3K. After that the PI3K activates RAC and after interacting with different genes various cytokines (IL-4, IL-5 and IL-13) are released after transcribing the cytokine genes.

Degranulating mast cells at tumour sites can also activate NK cells via histamine release. Histamine can induce NK cell activity by signalling.

But in case of natural killer cells receptor mediated pathway, the activating receptors such as NKp44, NKp30 and NKp46. The spleen associated tyrosine kinase (Syk) is activated when ZAP70 interacts with it and after that this complex (Syk and ZAP70) activates Phosphoinositide-3-kinase (PI3K) and then PI3K activates Ras-related C3 botulinum toxin substrate (RAC) and then different cytokines (such as IFN-gamma, IL-4, IL-2, etc.) are released.

Thus the both of these above pathways are very important in Tumour inhibition.

In Mast cells Degranulation pathway and Natural killer cells receptor mediated pathway, both of these pathway syk is present but it activates in different way. In case of mast cells FcεRI phosphorylates syk and in case of natural killer cells syk is activated when ZAP70 interacts with it. So, in Mast cells Degranulation pathway and Natural killer cells receptor mediated pathway, both of these pathways spleen associated tyrosine kinase (Syk) is present in upstream.

4.9 Role of TLR in NK mediated Tumour therapy

These following pathways are exported from KEGG pathway.

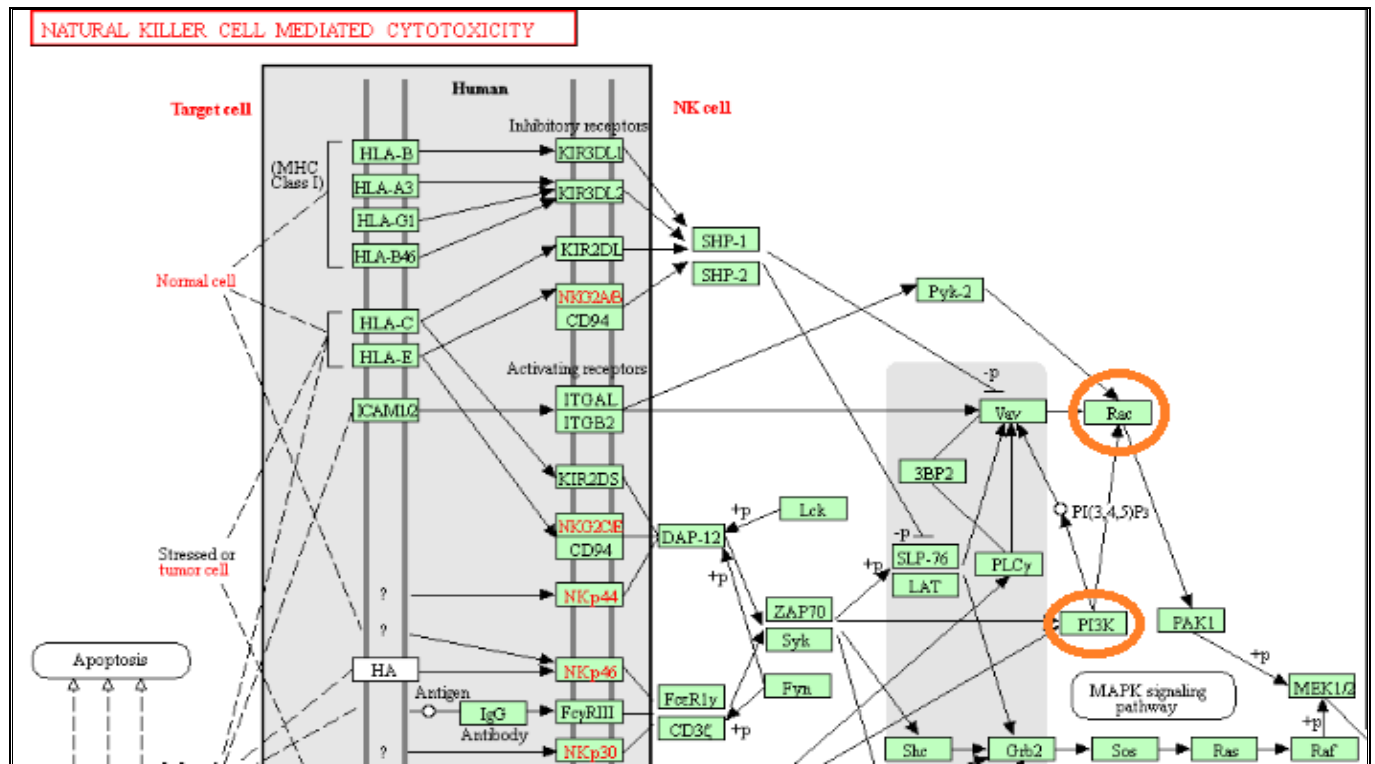


Fig31: Natural killer cells receptor mediated (Nkp⁴⁴, Nkp³⁰) pathway.

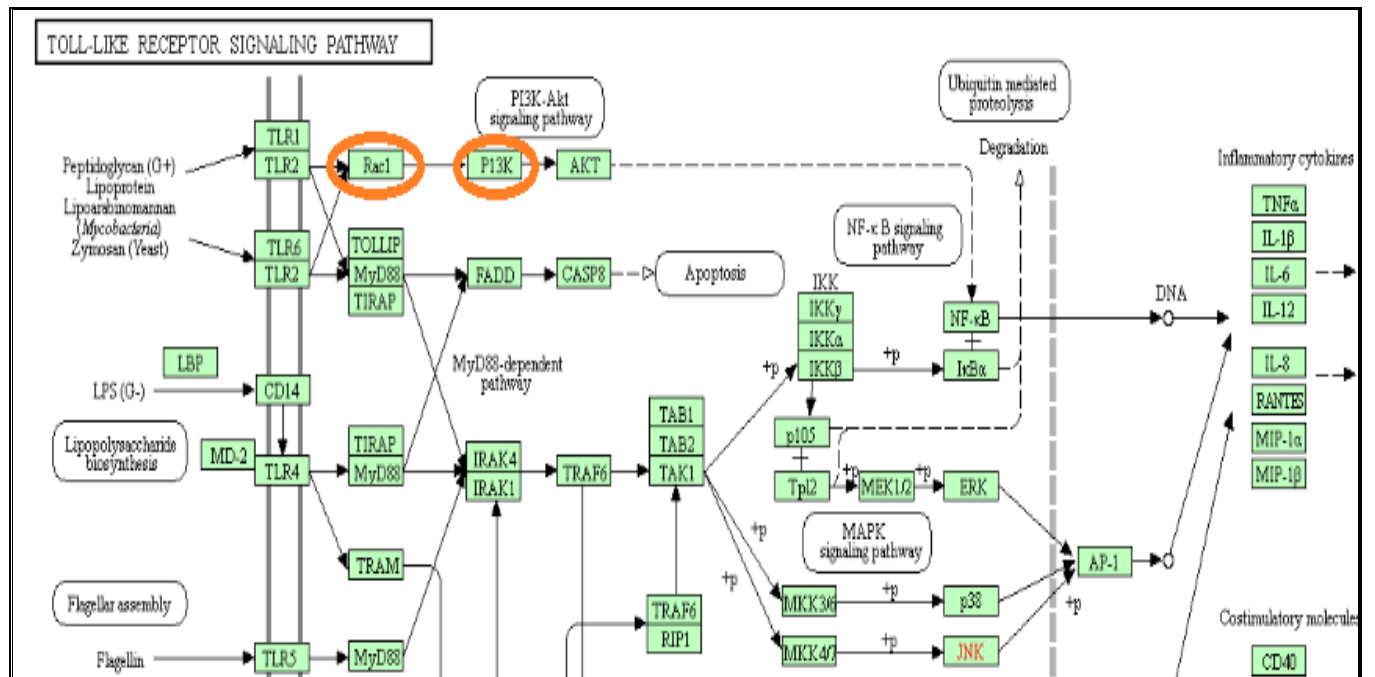


Fig32: Natural killer cells TLR mediated pathway (TLR2).

In natural killer cells receptor mediated pathway where the activating receptors such as NKp44, NKp30 and NKp46 are present. In this pathway PI3K activates Rac and then Rac activates PAK1 and this way different cytokines are released.

In natural killer cells TLR mediated pathway Rac activates PI3K and then PI3K activates AKT and then AKT activates NF-κB and after transcribing DNA different kinds of cytokines are released from NK cells.

In both of these pathways PI3K and Rac are present, so both of these genes are common between these two pathways. But in natural killer cells receptor mediated pathway PI3K activates Rac and in natural killer cells TLR mediated pathway Rac activates PI3K.

In case of natural killer cells receptor mediated pathway both Rac and PI3K genes are present in downstream but in natural killer cells TLR mediated pathway both of these genes are present in upstream.

4.10 Predictive TLR mediated Combinatorial Tumour therapy

TLR2 is present in mast cells as well as natural killer cells. Different kinds of ligands are present for TLR2 such as – lipoprotein, lipoteichoic acid, lipophosphatidylserine from gram (+) bacteria and Zymosan from yeast. Synthetic ligands such as Pam3CSK4. Using these ligands we activate mast cells expressed TLR2 and natural killer cells expressed TLR2. In both of these cells TLR2 is expressed in the surface of the cells. So using above mentioned TLR ligands we easily enhanced the activation of both of these cells which can help in the inhibition of tumour progression. Using these TLR2 ligands we can stimulate the tumour killing process.

4.11 Mast cells secreted cytokines activate natural killer cells

From mast cells microarray data Fold change value of this CXCL8 gene is calculated.

	A	B	C	D	E	F	G	H
1	ID	adj.P.Val	P.Value	t	B	logFC	Gene.symbol	Gene.title
2	15618	0.970678	8.87E-01	-1.49E-01	-6.99874	9.68E-02	CXCL8	C-X-C motif chemokine ligand 8

Fig33: Fold change value of CXCL8 gene in mast cells

Mast cells are activated by TLR2 and secretes CXCL8 chemokine and different inflammatory cytokines such as IL-2, IL-4, IL-8 and TNF α etc.

In microarray data of mast cells Fold change value of chemokine CXCL8 is 0.968 which helps in the recruitment of natural killer cells in tumour site.

Different cytokines such as IL-2, IL-4, IL-8 and TNF α also helps in activation of natural killer cells. These different kinds of cytokines are produced in TLR mediated pathway of mast cells.

CONCLUSION

Microarray data proves beneficial in many regards. Microarrays have made a significant contribution to science both because they can survey a large number of genes quickly or study a small sample size.

Microarrays have been used to assay gene expression within a single sample or to compare gene expression in two different cell types or tissue samples, such as healthy and diseased tissue. TLR-2 is known to be present on NK and Mast cells. Present microarray data of Mast cells stimulation with Anti-human IgE antibody (GEO accession id: **GSE3271**) and human NK cells stimulation with K562 erythroleukemic tumour cells (GEO accession id is **GSE55977**) lead to TLR-2 signalling up-regulation.

Hence it is an interesting finding that pro-inflammatory stimulation by Mast cells and receptor-mediated stimulation of NK cells lead to TLR2 signalling up-regulation. Our work proposes the use of Mast cells and NK cells synergy in TLR2 mediated signalling. Based on the microarray data we can conclude that different cytokines such as

IL-2

IL-4

IL-8

TNF- α

Show significant increase in expression upon IgE mediated Mast cells degranulation. These cytokines are also known for significant inflammatory response. Natural killer cells has been shown to express different cytokines receptors such as

IL12A

IL12B

IL2RB

IL2RA

IL4R.

Hence there is a direct Mast cells mediated NK cells stimulation through cytokines.

TLRs present on NK cells- **TLR-1, TLR-2, TLR-4, TLR-5, and TLR-6.**

TLRs present on Mast cells- **TLR-2, TLR-4, TLR-6, and TLR-8.**

Analysis of KEGG pathway suggests that TLR-2 signalling through MYD88, different genes such as MYD88, TIRAP, IRAK1, IRAK4, TRAF6 and IRF5 are activated in both NK and Mast cells. These different genes (MYD88, TIRAP, IRAK1, IRAK4, TRAF6 and IRF5) are present in MYD88 signalling of NK cells and also present in mast cells microarray data with high Fold change values. So that means upon stimulation of Anti-IgE on Mast cells different cytokines such as IL-2, IL-4, IL-8 and TNF-alpha are released etc. These cytokines (IL-2, IL-4, IL-8 and TNF-alpha) binds to the receptors (such as IL12A, IL12B, IL2RB, IL2RA and IL4R) of natural killer cells and activates the natural killer cells. So here the TLR-2 is the key molecule through which we can activate both mast cells and natural killer cells in in-vitro condition using synthetic ligands such as Pam3CSK4 which is recognized by TLR-2.

In combinatorial therapy, inflammatory response of Mast cells degranulation and TLR-mediated stimulation can be predicted to boost Natural killer cells cytolytic activity in a synergistic manner.

From our study we can predict some key stimulants for both Mast cells activation and TLR stimulation that can act as adjuvants for Natural killer cells activation.

Future Perspectives

- (i) Immunotherapeutic approaches targeting NK cells and mast cells should prove useful in inducing more effective immune responses to improve treatments.
- (ii) The targeting of TLRs is to induce cytokines, which are well validated in the diseases, and indeed successfully being targeted in the clinic.
- (iii) Targeting only one TLR (TLR2) we can activate Mast cells as well as NK cells.
- (iv) Immunotherapy combining the use of TLR agonists together with chemotherapy or radiation therapy seems to be a promising approach.

REFERENCE

1. Rasmussen T, Jensen JF, Ostergaard N, Tanner D, Ziegler T, Norrby PO. On the mechanism of the copper-catalyzed cyclopropanation reaction. *Chemistry* (2002) **8**(1):177–84.
2. Siraganian RP. Mast cell signal transduction from the high-affinity IgE receptor. *Curr Opin Immunol* (2003) **15**(6):639–46. doi:10.1016/j.coi.2003.09.010
3. Sibilano R, Frossi B, Pucillo CE. Mast cell activation: a complex interplay of positive and negative signaling pathways. *Eur J Immunol* (2014) **44**(9):2558–66. doi:10.1002/eji.201444546
4. Kalesnikoff J, Galli SJ. New developments in mast cell biology. *Nat Immunol* (2008) **9**(11):1215–23. doi:10.1038/ni.f.216
5. Iwaki S, Tkaczyk C, Metcalfe DD, Gilfillan AM. Roles of adaptor molecules in mast cell activation. *Chem Immunol Allergy* (2005) **87**:43–58.
6. Hitomi T, Zhang J, Nicoletti LM, Grodzki AC, Jamur MC, Oliver C, et al. Phospholipase D1 regulates high-affinity IgE receptor-induced mast cell degranulation. *Blood* (2004) **104**(13):4122–8. doi:10.1182/blood-2004-06-2091
7. Shalit M, Levi-Schaffer F. Challenge of mast cells with increasing amounts of antigen induces desensitization. *Clin Exp Allergy* (1995) **25**(9):896–902. doi: 10.1111/j.1365-2222.1995.tb00033.x
8. Karre K, Ljunggren HG, Piontek G, Kiessling R. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature*.1986; **319**(6055):675–678. [PubMed: 3951539]
9. Saleh A, Davies GE, Pascal V, et al. Identification of probabilistic transcriptional switches in the Ly49 gene cluster: a eukaryotic mechanism for selective gene activation. *Immunity*.2004 Jul; **21**(1):55–66. [PubMed: 15345220]
10. Davies GE, Locke SM, Wright PW, et al. Identification of bidirectional promoters in the human KIR genes. *Genes Immun*.2007 Apr; **8**(3):245–253. Epub 2007 Feb 22. [PubMed: 17315044]
11. Lanier LL. Up on the tightrope: natural killer cell activation and inhibition. *Nat Immunol*. 2008; **9**(5):495–502. [PubMed: 18425106]
12. Braud VM, Allan DS, O’Callaghan CA, et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature*. 1998; **391**(6669):795–799. [PubMed: 9486650]
13. Vance RE, Kraft JR, Altman JD, Jensen PE, Raulet DH. Mouse CD94/NKG2A is a natural killer cell receptor for the nonclassical major histocompatibility complex (MHC) class I molecule Qa-1(b). *J Exp Med*. 1998; **188**(10):1841–1848. [PubMed: 9815261]
14. Norrby K. Mast cells and angiogenesis. *APMIS* (2002) **110**(5):355–71.

15. Bulfone-Paus S, Bahri R. Mast cells as regulators of T cell responses. *Front Immunol* (2015) **6**:394. doi:10.3389/fimmu.2015.00394
16. Goto Y, Kurashima Y, Kiyono H. The gut microbiota and inflammatory bowel disease. *Curr Opin Rheumatol* (2015) **27**(4):388–96.
17. Harker JA, Godlee A, Wahlsten JL, et al. Interleukin 18 coexpression during respiratory syncytial virus infection results in enhanced disease mediated by natural killer cells. *J Virol*. 2010; **84**(8): 4073–4082. [PubMed: 20130064]
18. Culley FJ. Natural killer cells in infection and inflammation of the lung. *Immunology*. 2009; **128**(2):151–163. [PubMed: 19740372]
19. Anthony DA, Andrews DM, Chow M, et al. A role for granzyme M in TLR4-driven inflammation and endotoxemia. *J Immunol*. 2010; **185**(3):1794–1803. [PubMed: 20585036]
20. Colonna M, Jonjic S, Watzl C. Natural killer cells: fighting viruses and much more. *Nat Immunol*. 2011; **12**(2):107–110. [PubMed: 21245897]
21. Miyake T, Satoh T, Kato H, et al. I κ B ζ is essential for natural killer cell activation in response to IL-12 and IL-18. *Proc Natl Acad Sci USA*. 2010; **107**(41):17680–17685. [PubMed: 20876105]
22. Van Domselaar R, Philippen LE, Quadir R, Wiertz EJ, Kummer JA, Bovenschen N. Noncytotoxic inhibition of cytomegalovirus replication through NK cell protease granzyme M-mediated cleavage of viral phosphoprotein 71. *J Immunol*. 2010; **185**(12):7605–7613. [PubMed: 21059895]
23. Muntasell A, Magri G, Pende D, Angulo A, Lopez-Botet M. Inhibition of NKG2D expression in NK cells by cytokines secreted in response to human cytomegalovirus infection. *Blood*. 2010; **115**(25):5170–5179. [PubMed: 20393128]
24. Slavuljica I, Busche A, Babic M, et al. Recombinant mouse cytomegalovirus expressing a ligand for the NKG2D receptor is attenuated and has improved vaccine properties. *J Clin Invest*. 2010; **120**(12):4532–4545. [PubMed: 21099111]
25. Terme M, Ullrich E, Delahaye NF, Chaput N, Zitvogel L. Natural killer cell-directed therapies: moving from unexpected results to successful strategies. *Nat Immunol*. 2008; **9**(5):486–494. [PubMed: 18425105]
26. Lanier LL. NK cell recognition. *Annu Rev Immunol*. 2005; **23**:225–274. [PubMed: 15771571]
27. Bryceson YT, March ME, Ljunggren HG, Long EO. Activation, coactivation and costimulation of resting human natural killer cells. *Immunol Rev*. 2006; **214**:73–91. [PubMed: 17100877]
28. Metcalf D. The colony-stimulating factors and cancer. *Nat Rev Cancer*. 2010; **10**(6):425–434. [PubMed: 20495576]
29. da Silva EZ, Jamur MC, Oliver C. Mast cell function: a new vision of an old cell. *J Histochem Cytochem* (2014) **62**(10):698–738. doi: 10.1369/0022155414545334

30. Xiang Z, Block M, Löfman C, Nilsson G. IgE-mediated mast cell degranulation and recovery monitored by time-lapse photography. *J Allergy Clin Immunol* (2001) **108**(1):116–21. doi:10.1067/mai.2001.116124
31. Basu S, Fenton MJ. Toll-like receptors: function and roles in lung disease. *Am J Physiol Lung Cell Mol Physiol* (2004) **286**(5):L887–92. doi:10.1152/ajplung.00323.2003
32. Bryceson YT, March ME, Ljunggren HG, Long EO. Activation, coactivation and costimulation of resting human natural killer cells. *Immunol Rev.* 2006; **214**:73–91. [PubMed: 17100877]
33. Wadler S, Schwartz EL. Antineoplastic activity of the combination of interferon and cytotoxic agents against experimental and human malignancies: a review. *Cancer Res.* 1990; **50**(12):3473–3486. [PubMed: 1692761]
34. Carson WE, Ross ME, Baiocchi RA, et al. Endogenous production of interleukin 15 by activated human monocytes is critical for optimal production of interferon- γ by natural killer cells in vitro. *J Clin Invest.* 1995; **96**(6):2578–2582. [PubMed: 8675621]
35. Zhang H, Cui Y, Voong N, et al. Activating signals dominate inhibitory signals in CD137L/IL-15 activated natural killer cells. *J Immunother.* 2011; **34**(2):187–195. [PubMed: 21304401]
36. Giorgia Gri,¹ Barbara Frossi,¹ Federica D’Inca,¹ Luca Danelli,¹ Elena Betto,¹ Francesca Mion,¹ Riccardo Sibilano,¹ and Carlo Pucillo¹. Mast Cell: An Emerging Partner in Immune Interaction. *Front. Immunol.* 25 May 2012; doi: [10.3389/fimmu.2012.00120](https://doi.org/10.3389/fimmu.2012.00120). [PMC free article]

Interplay of Mast Cells and Natural Killer Cells

ORIGINALITY REPORT

7% SIMILARITY INDEX	5% INTERNET SOURCES	4% PUBLICATIONS	2% STUDENT PAPERS
-------------------------------	-------------------------------	---------------------------	-----------------------------

PRIMARY SOURCES

1	Anderson, Ronald, Gregory Tintinger, Riana Cockeran, Moliehi Potjo, and Charles Feldman. "Beneficial and Harmful Interactions of Antibiotics with Microbial Pathogens and the Host Innate Immune System", Pharmaceuticals, 2010. Publication	1%
2	bmcgenomics.biomedcentral.com Internet Source	1%
3	www.jimmunol.org Internet Source	1%
4	Submitted to British International School, Jakarta Student Paper	<1%
5	java.worldbestlearningcenter.com Internet Source	<1%
6	www.ncbi.nlm.nih.gov Internet Source	<1%
7	"New Findings from University of Melbourne Update Understanding of Nucleic Acids Research (limma powe", Life Science	<1%

Weekly, June 16 2015 Issue

Publication

8	Wang, Fuyan, Zhigang Tian, and Haiming Wei. "Genomic expression profiling of NK cells in health and disease", European Journal of Immunology, 2014. Publication	<1%
9	blog.professional-webworkx.de Internet Source	<1%
10	Debika Bhattacharya. "Expression of Toll-Like Receptors on Breast Tumors: Taking a Toll on Tumor Microenvironment", International Journal of Breast Cancer, 2012 Publication	<1%
11	www.cse.scu.edu Internet Source	<1%
12	Cheng, Min, Yongyan Chen, Weihua Xiao, Rui Sun, and Zhigang Tian. "NK cell-based immunotherapy for malignant diseases", Cellular and Molecular Immunology, 2013. Publication	<1%
13	dreampharmaceuticals.com Internet Source	<1%
14	r-forge.r-project.org Internet Source	<1%
15	Nabil Belacel. "Clustering Methods for Microarray Gene Expression Data", Omics A Journal of Integrative Biology, 12/2006	<1%

Publication

16	www.onmoso.com Internet Source	<1%
17	forums.sun.com Internet Source	<1%
18	Li, Xiu-zhen, Lin Lv, Xu Zhang, Kenneth Yongabi Anchang, Auwalu Yusuf Abdullahi, Liqing Tu, Xiaohu Wang, Lijun Xia, Xiu-Xiang Zhang, Weili Feng, Chunxia Lu, Shoujun Li, and Zi-Guo Yuan. "Recombinant canine adenovirus type-2 expressing TgROP16 provides partial protection against acute <i>Toxoplasma gondii</i> infection in mice", <i>Infection Genetics and Evolution</i> , 2016. Publication	<1%
19	jexp.ru Internet Source	<1%
20	www.curevents.com Internet Source	<1%
21	Laprairie, Robert B., Amina M. Bagher, Sophie V. Precious, and Eileen M. Denovan-Wright. "Components of the endocannabinoid and dopamine systems are dysregulated in Huntington's disease: analysis of publicly available microarray datasets", <i>Pharmacology Research & Perspectives</i> , 2015. Publication	<1%
