"Identification of Important Key Regulators in Cervical Cancer Network"

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

Master of Technology
In
Bioinformatics

Submitted by

KAPIL JANGRA

(2K15/BIO/06)

Delhi Technological University, Delhi, India

Under the supervision of

Dr. ASMITA DAS

(Assistant Professor)

Department of Biotechnology

Delhi Technological University

Shahbad Daulatpur, Main Bawana Road,

Delhi – 110042, INDIA

CERTIFICATE

This is to certify that the M.Tech. Dissertation entitled "***Identification of Important Key Regulators in Cervical cancer Network",*** submitted by KAPIL JANGRA (2K15/BIO/06) in partial fulfillment of the requirement for the award of the degree of Master of Technology, 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 honoring of any other degree.

Date:

Dr. Asmita Das
(Supervisor)
Department of Biotechnology
Delhi Technological University
Shahbad Bawana Road, Delhi-110042

Prof. D. Kumar
(H.O.D)
Department of Biotechnology
Delhi Technological University
Shahbad Bawana Road, Delhi-110042

DECLARATION

I, Kapil Jangra, hereby declare that the report entitled "Identification of Important Key Regulators in Cervical Cancer Network" submitted in partial fulfillment 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 for the award of any Degree/Diploma/Associateship/Fellowship or other similar title to any candidate of any University/institution.

Date: Signature of 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. Das office was always open whenever I ran into a trouble spot or had a question about my research or writing.

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

I also wish to express my gratitude to my teachers and other non-teaching staff members for providing us their valuable teachings with constant support and guidance.

Also, my special thanks go to my beloved friends HariKrishan, Razi Khan, Mottadi Shiva, Aniruddha (PG Student, Dept. of Biotech., DTU) who helps me in understanding and seeking the programming part of my project and to all my buddies who are there with me for any kind of assistance.

Finally, I must express my very profound gratitude to my parents 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

Date :

KAPIL JANGRA

M.Tech. (Bioinformatics)

(2K15/BIO/06)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CCDB - Cervical Cancer Database

STRING - Search Tool for the Retrieval of Interacting genes

NCBI - National Centre for Biotechnology Information

GEO - Gene Expression Omnibus

Log FC - Logarithm of Fold Change value

CDK 2 - Cyclin Dependant Kinase 2

E2F1 - E2F transcription factor 1

CDKN1A- Cyclin Dependent Kinase Inhibitor 1A

CDKN2A- Cyclin Dependent Kinase Inhibitor 2A

TP53 - Tumor Protein P53

PTGS2 - Prostaglandin-Endoperoxide Synthase 2

CTNNB1 - Catenin Beta-1

ABSTRACT:

Cervical cancer is a major cause of death in women. It's a cancer of the cervix, of the female reproductive organ; cervix is a narrow opening in the uterus and joins it to the vagina. There are 3000 genes associated with cervical cancer. We created a PPI network and then studied the topological properties of the network that gave evidence that the network is scale free, the degree distribution obeys power law and have some nodes (vertices) that play a crucial role during cancer progression. So the aim of this project was to identify these nodes (also called fundamental key regulators) and the strategy used for this, is based on Barabasi-Albert model (which is a bottom up approach). The model suggested that the network we created, is evolved in nature by growth and preferential attachment concept, and have hub nodes. Finally, we concluded that there are 7 fundamental key regulators are present in our network that is dominant, which means if we are able to delete these key regulators, the whole network would be collapsed. So, by understanding the complex functionality and regulation of these fundamental key regulators, we can identify the diagnostic biomarkers and can be developed early detection techniques and therapy for cervical cancer.

The list of seven fundamental key regulators involved in cervical cancer progression is given below:

CDK2, E2F1, CDKN1A, CDKN2A, TP53, PTGS2 and CTNNB1.

2. INTRODUCTION:

Cervical cancer is cancer of the cervix, in women, the narrow opening into the uterus from the vagina. The normal "ectocervix" (the portion of the uterus extends into the vagina) is covered with flat, thin cells called squamous cells [2][8]. The "endocervix" or cervical canal is made up of another kind of cell called columnar cells. Most cervical cancers (80 to 90 percent) are squamous cell cancers. Adenocarcinoma is the second most common type of cervical cancer, accounting for the remaining 10 to 20 percent of cases [9]. Adenocarcinoma develops from the glands that produce mucus in the endocervix. The risk factors for cervical cancer are Human papilloma virus infection [7][9], Smoking, Immunosuppression, Being overweight, Long-term use of oral contraceptives (birth control pills) [1], Having multiple full-term pregnancies and Having a family history of cervical cancer [12].

There are more than 3000 genes are associated with cervical cancer progression (validated by comparing the gene expression profiles of microarray data taken from gene expression omnibus, an online database of NCBI). By using Cytoscape tool we created a PPI network for cervical cancer. The next step of this project is to prove the network we created is scale-free. Because all real network (real networks are not created, they evolved by time) exist in nature, are scale free network [56][70]. The Scale free network is defined as the network that reproduces power law degree distribution, contain hubs nodes with a very high number of links and the distribution of node linkages follows a power law. In these networks, most nodes have a few connections (very low degree) and some nodes have a large number of links (very high degree) which means the network has no scale.

Scale free network has some features (degree distribution, clustering coefficient, contains communities, sub communities, hub nodes and their reliance on these hubs) over a random network (that reproduces Poisson degree distribution instead of power law distribution) and if a network also showing these features, it must be a scale free network [51][58]. So, in order to prove our network is scale free, we have taken the help of NetworkAnalyzer, a plug-in of Cytoscape tool. The next step is to find communities, sub communities that are present in the network, by using R language [76]. The purpose of finding communities and sub communities is to find key regulators and to prove that our network is evolved in nature by growth process and preferential attachment [40][44].

Barabasi-Albert model is based on growth process (The network starts from few nodes and it grows as new nodes are added) and preferential attachment (The probability that a new node will connect to an existing node, is not uniform. The probability is proportional to degree of existing node) and it is a generative model proposed to understand the mechanism responsible for emergence of scale free network, in real world [41][71]. Finally, we concluded that there are 7 fundamental key regulators (CDK2, E2F1, CDKN1A, CDKN2A, TP53, PTGS2 and CTNNB1) present in our network that play an important role in cervical cancer progression. If we target these fundamental key regulators, the whole network will break into tiny noncommunicating islands that will be helpful in the cure of cervical cancer.

CCDB (Cervical cancer gene database), STRING (Search Tool for the Retrieval of Interacting Genes/Proteins), GEO (Gene Expression Omnibus), GEO2R (GEO-tool), Cytoscape, NetworkAnalyzer (a plug-in of Cytoscape) and R Studio (integrated development environment for R language) are the tools and databases, used in this project.

3) REVIEW OF LITERATURE:

Cervical cancer is cancer that starts in the cervix, the narrow opening into the uterus from the vagina [2]. It is a strong muscle. The normal "ectocervix" (the portion of the uterus extends into the vagina) is a healthy pink color and is covered with flat, thin cells called squamous cells. The "endocervix" or cervical canal is made up of another kind of cell called columnar cells. The area where these cells meet is called the "transformation zone" (T-zone) and is the most likely location for abnormal or precancerous cells to develop. Most cervical cancers (80 to 90 percent) are squamous cell cancers. Adenocarcinoma is the second most common type of cervical cancer, accounting for the remaining 10 to 20 percent of cases [8][9]. Adenocarcinoma develops from the glands that produce mucus in the endocervix. While less common than squamous cell carcinoma, the incidence of adenocarcinoma is on the rise, particularly in younger women. Cervical cancer is one of leading cancer among women worldwide, with an estimated 520,000 new cases and 274,000 deaths reported annually (WHO/ICO Information Centre on HPV and Cervical Cancer- HPV and cervical cancer statistics in India. 2010), but because it develops over time, **it is also one of the most preventable types of cancer [32]**. Cancer of the cervix tends to occur during midlife. Half of the women diagnosed with the disease are between 35 and 55 years of age [8][30]. It rarely affects women under age 20, and approximately 20 percent of diagnoses are made in women older than 65. For this reason, it is important for women to continue cervical cancer screening until at least the age of 70 [2][3].

3.1) Risk factors for cervical cancer:

- **Human papilloma virus infection**

  The most important risk factor for cervical cancer is infection by the human papilloma virus (HPV) [4][5]. HPV is a group of more than 150 related viruses, some of which cause a type of growth called papillomas, which are more commonly known as warts. HPV can infect cells on the surface of the skin, and those lining the genitals, anus, mouth, and throat, but not the blood or internal organs such as the heart or lungs. HPV can be spread from one person to another during skin-to-skin contact. One way HPV is spread is through sex, including vaginal, anal, and even oral sex [7]. Different types of HPV cause warts on different parts of the body [14][23]. Some types cause common warts on the hands and feet. Other types tend to cause warts on the lips or tongue. Certain types of HPV may cause warts to appear on or around the genital organs and in the anal area [9][18]. These warts may barely be visible or they may be several inches across. These are known as genital warts or condyloma acuminatum. HPV 6 and HPV 11 are the 2 types of HPV that cause most cases of genital warts. These are called low-risk types of HPV because they are seldom linked to cervical cancer. Other types of HPV are

called high-risk types because they are strongly linked to cancers, including cancers of the cervix, vulva, and vagina in women, penile cancer in men, and anal and oral cancer in men and women. The high-risk types include HPV 16, HPV 18, HPV 31, HPV 33, and HPV 45, as well as some others. There might be no visible signs of infection with a high-risk HPV until pre-cancerous changes or cancer develops [10][13].

- **Smoking**

  Women who smoke are about twice as likely as nonsmokers to get cervical cancer [1][18]. Smoking exposes the body to many cancer-causing chemicals that affect organs other than the lungs.

- **Immunosuppression**

  Human immunodeficiency virus (HIV), the virus that causes AIDS, damages the immune system and puts women at higher risk for HPV infection [11][21]. This might increase the risk of cervical cancer in women with AIDS, Another group of women at risk of cervical cancer are those taking drugs to suppress their immune response, such as those being treated for an autoimmune disease or those who have had an organ transplant.

- **A diet low in fruits and vegetables**

  Women whose diets don't include enough fruits and vegetables may be at increased risk for cervical cancer [32].

- **Being overweight**

  Overweight women are more likely to develop adenocarcinoma of the cervix [1][17].

- **Long-term use of oral contraceptives (birth control pills)**

  There is evidence that taking oral contraceptives (OCs) for a long time increases the risk of cancer of the cervix. Research suggests that the risk of cervical cancer goes up the longer a woman takes OCs, but the risk goes back down again after the OCs are stopped [1]. In one study, the risk of cervical cancer was doubled in women who took birth control pills longer than 5 years, but the risk returned to normal 10 years after they were stopped.

- **Intrauterine device use**

  A study found that women who had ever used an intrauterine device (IUD) had a lower risk of cervical cancer. The effect on risk was seen even in women who had an IUD for less than a year, and the protective effect remained after the IUDs were removed. Using an IUD might also lower the risk of endometrial (uterine) cancer.

- **Having multiple full-term pregnancies**

  Women who have had 3 or more full-term pregnancies have an increased risk of developing cervical cancer. One theory is that these women had to have had unprotected intercourse to get pregnant, so they may have had more exposure to HPV [17]. Also, studies have pointed to hormonal changes during pregnancy as possibly making women more susceptible to HPV infection or cancer growth. Another thought is that pregnant women might have weaker immune systems, allowing for HPV infection and cancer growth.

- **Being younger than 17 at your first full-term pregnancy**

  Women who were younger than 17 years when they had their first full-term pregnancy are almost 2 times more likely to get cervical cancer later in life than women who waited to get pregnant until they were 25 years or older [17].

- **Having a family history of cervical cancer**
  Cervical cancer may run in some families. If a woman's mother or sister had cervical cancer, her chances of developing the disease are 2 to 3 times higher than if no one in the family had it.

3.2) Signs and symptoms:

- Abnormal vaginal bleeding, such as bleeding after sex (vaginal intercourse), bleeding after menopause, bleeding and spotting between periods, and having longer or heavier (menstrual) periods than usual. Bleeding after douching, or after a pelvic exam is a common symptom of cervical cancer but not pre-cancer [78],[80].
- An unusual discharge from the vagina − the discharge may contain some blood and may occur between your periods or after menopause [78].
- Pain during sex (vaginal intercourse) [79].

3.3) Screening of cervical cancer:

HPV DNA test

This test determines whether the patient is infected with any of the HPV types that are most likely to cause cervical cancer. This involves collecting cells from the cervix for lab testing [35].
If the patient experiences signs and symptoms of cervical cancer, or if the Pap test revealed abnormal cells, the patient might undergo additional tests [15][16]:

- **Biopsy:** A small section of tissue is taken under general anesthetic.

- **Colposcopy:** A speculum is placed to hold the vagina open as the gynecologist looks at the cervix through a colposcope - a lighted magnifying instrument [31].
- **Cone biopsy:** A small cone-shaped section of abnormal tissue is taken from the cervix for examination [32].
- **LLETZ:** A diathermy (wire loop with an electric current) is used to remove abnormal tissue. The tissue is sent to the lab to be checked.
- **Blood tests:** measures number of blood cells, and can identify any liver or kidney problems.
- **Examination under anesthetic (EUA):** This allows the doctor to examine the vagina and cervix more thoroughly.
- **CT scan:** The patient consumes a barium drink that appears white on the scan. Just before the scan, a tampon may be placed into the vagina, and a barium liquid may be inserted into the rectum.
- **MRI:** By using high-MRI with a special vaginal coil, a technique to measure the movement of water within tissue, researchers may be able to identify cervical cancer in its early stages.
- **Pelvic ultrasound:** This is a device that uses high-frequency sound waves to create an image on a monitor of the target area.

Current therapy used for Cervical Cancer:

Chemotherapy

Chemotherapy is the use of drugs to destroy cancer cells, usually by stopping the cancer cells' ability to grow and divide. Chemotherapy is given by a medical oncologist, a doctor who specializes in treating cancer with medication [80][81].

Systemic chemotherapy gets into the bloodstream to reach cancer cells throughout the body. Common ways to give chemotherapy include an intravenous (IV) tube placed into a vein using a needle or in a pill or capsule that is swallowed (orally).

A chemotherapy regimen (schedule) usually consists of a specific number of cycles given over a set period of time. A patient may receive one drug at a time or combinations of different drugs at the same time.

Side- effects of chemotherapy:

The side effects of chemotherapy depend on the woman and the dose used, but they can include fatigue, the risk of infection, nausea and vomiting, hair loss, loss of appetite, and diarrhea. These side effects usually go away once treatment is finished [82].

Other possible long-term side effects include the inability to become pregnant and early menopause. Rarely, specific drugs may cause some hearing loss [80].

Radiation therapy

Cervical cancer is often treated along with radiation therapy. The goal of chemotherapy when given with radiation therapy is to increase the effectiveness of the radiation treatment. It can also be given to destroy cancer that is remaining after surgery, also called adjuvant therapy, or treat cervical cancer it has come back. The addition of bevacizumab (Avastin) to combination chemotherapy in patients with later stages of cervical cancer showed improvement in outcome [81][82].

Although chemotherapy can be given orally (by mouth), most drugs used to treat cervical cancer are given intravenously (IV). IV chemotherapy is either injected directly into a vein or through a thin tube called a catheter, which is a tube temporarily put into a large vein to make injections easier.

Radiation therapy is the use of high-energy x-rays or other particles to destroy cancer cells. A doctor who specializes in giving radiation therapy to treat cancer is called a radiation oncologist. Radiation therapy may be given alone, before surgery, or instead of surgery to shrink the tumor. Many women may be treated with a combination of radiation therapy and chemotherapy [80].

The most common type of radiation treatment is called external-beam radiation therapy, which is radiation given from a machine outside the body. When radiation treatment is given using implants, it is called internal radiation therapy or brachytherapy. A radiation therapy regimen (schedule) usually consists of a specific number of treatments given over a set period of time.

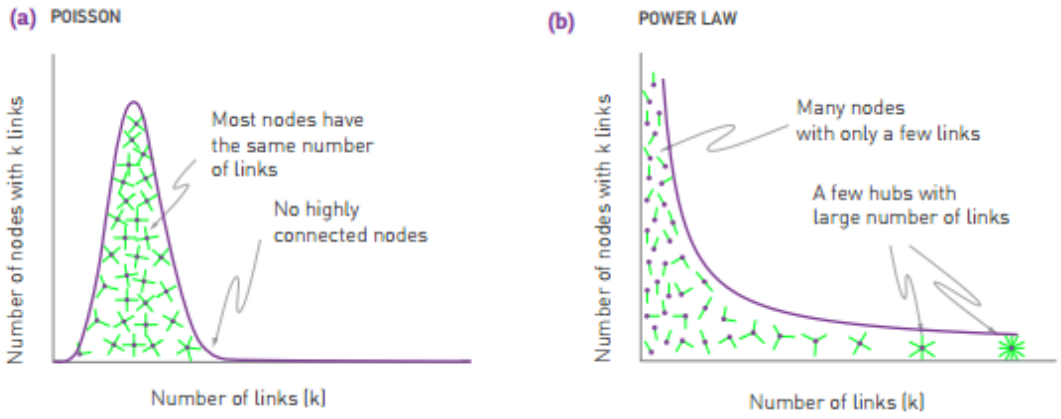
Side effects from radiation therapy

Side effects from radiation therapy may include fatigue, mild skin reactions, upset stomach, and loose bowel movements. Side effects of internal radiation therapy may include abdominal pain and bowel obstruction, although it is uncommon [81]. Most side effects usually go away soon after treatment is finished. After radiation therapy, the vaginal area may lose elasticity so some women may also want to use a vaginal dilator, which is a plastic or rubber cylinder that is inserted into the vagina to prevent narrowing.

3.4) Random network vs Scale free network:

Random networks are those networks that do not produce power law degree distribution. In such networks, a plot of the distribution of node linkages will follow a bell shaped curve, with most nodes having approximately the same number of links.

The Scale free network reproduce power law degree distribution, contain hubs nodes with a very high number of links and the distribution of node linkages follows a power law. In these networks, most nodes have a few connections (very low degree) and some nodes have a large number of links (very high degree) which means the network has no scale [51].



Source : Scale Free Network- by Albert-Laszlo and Eric Bonabeau

Most biological networks are scale-free, which means that their degree distribution follows a power law, $P(k) \sim k^{-\gamma}$, where γ is the degree exponent. The value of γ determines many properties of the system.

When ɣ < 2 = the hubs (nodes have higher degree than average) have important role in the network

ɣ>3 = the hubs are not relevant, scale free network behaves like a random one

For 2>γ>3 there is a hierarchy of hubs

3.5) Barabasi-Albert model:

Several generative models are proposed in order to understand the mechanism responsible for the evolution of scale free network, in real world. The Barabasi-Albert model is one of them and based on following two definitions:

- **Growth** - The network starts from few nodes and it grows as new nodes are added.
- **Preferential attachment** - The probability that a new node (vertex) will connect to an existing node, is not uniform. The probability is proportional to the degree of existing node [72]. The "rich get richer" process will generally favor the existing nodes, which are more likely to eventually become hubs.

For example: d(v)= the degree of node v

P(new node u connects to existing node v) = a.d(v)

a = proportionality constant

3.6) Evolutionary origin of scale-free networks:

Growth process and preferential attachment, both are responsible for the emergence of the scale-free property in complex networks [39][40]. Indeed, if a vertex (node) has higher degree, new vertices will tend to connect to it with a higher probability. This node will therefore gain new links at a higher rate than its less connected peers and will turn into a hub. Growth and preferential attachment have a common origin in protein networks that is probably rooted in gene duplication [72]. Duplicated genes produce identical proteins that interact with the same protein partners, therefore, each protein that is in contact with a duplicated protein gains an extra link.

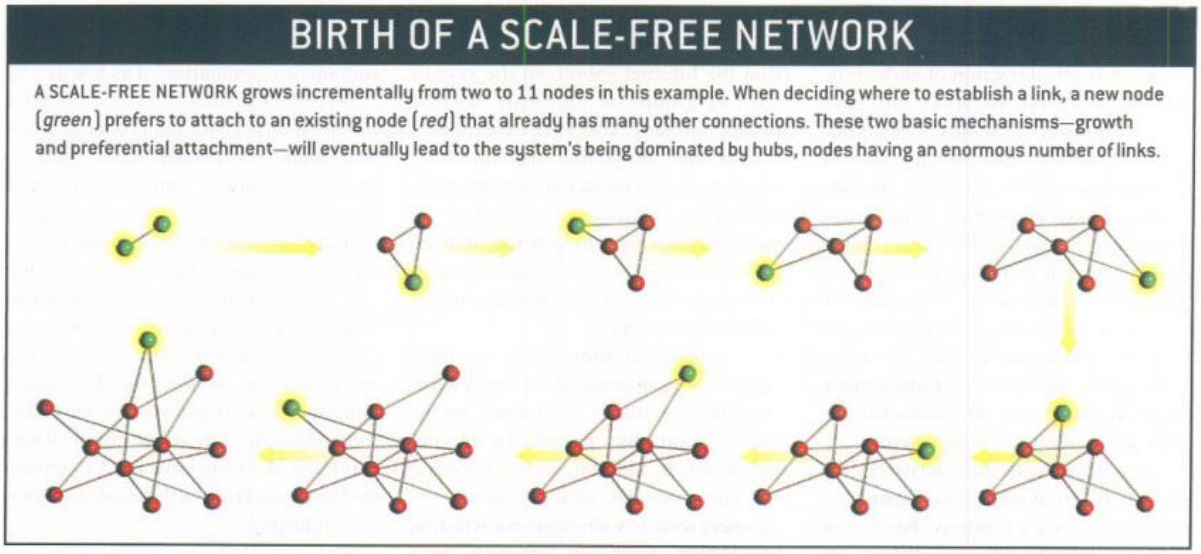


Figure 1: Birth of a scale free network

The evolution process of a real scale free network is shown here, The linkage of new node with existing nodes is based on preferential attachment and this process favor the formation of hubs in network. (Source: Scale Free Network- by Albert-Laszlo and Eric Bonabeau)

3.7) Characteristics of scale free network:

In case of random network, if a critical fraction of nodes is removed, the whole network or system breaks into tiny, non-communicating islands, but Scale free networks are robust against accidental failures, a property that is rooted in their inhomogeneous topology. Same thing happens in the case of PPI network, means if a high level of random mutations occurred in

network, still, unaffected proteins will continue to work together [51][71]. Because the random removal of nodes eliminates the low degree nodes that are much more plentiful than hubs, and their removal will not disrupt the network topology significantly.

The dependency of network topology on hubs can be advantageous in case of disease network [75]. The network's reliance on hubs provides pharmaceutical researchers with new strategies for selecting drug targets, potentially leading to cures for disease [38].

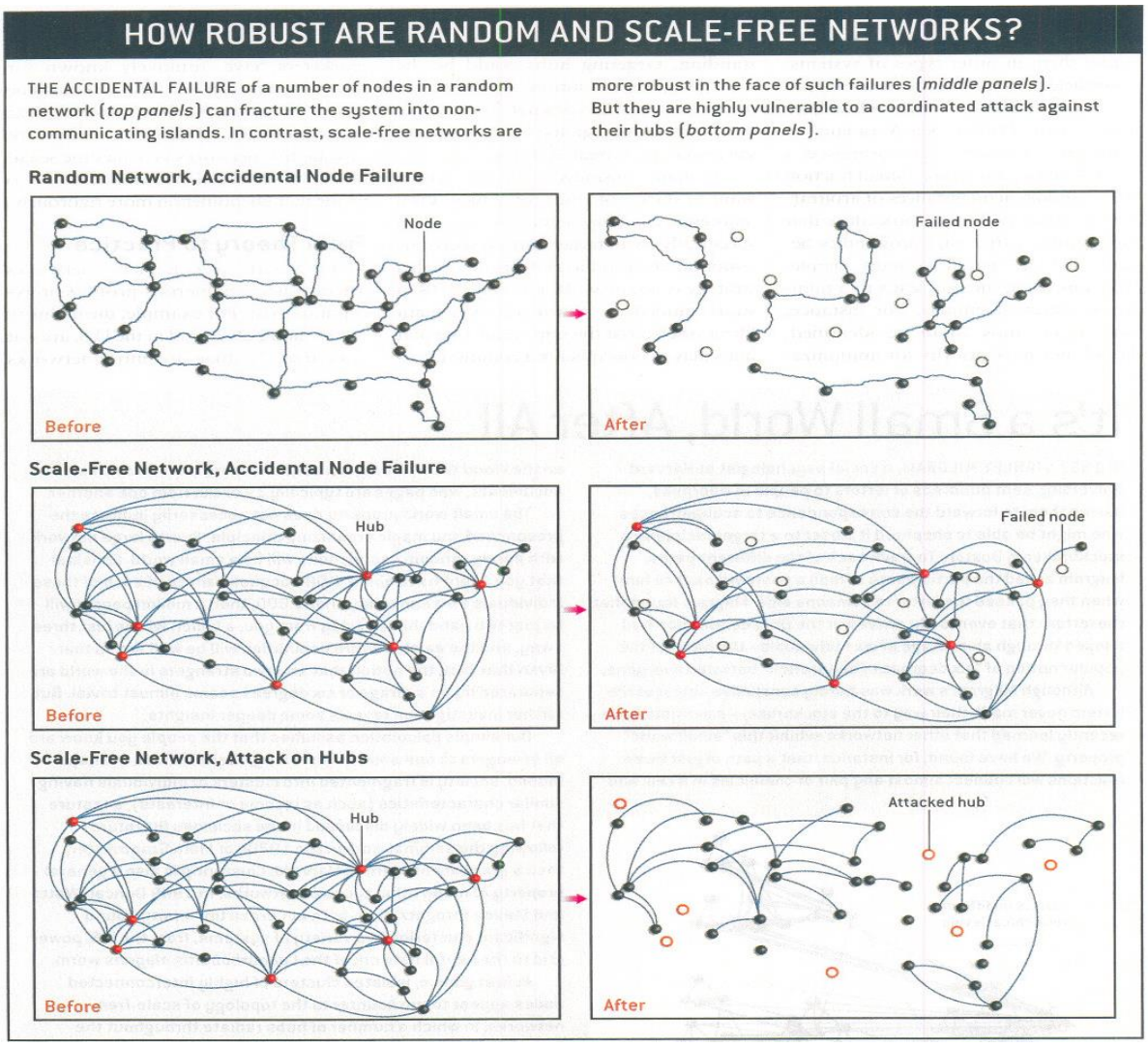


Figure 2: Random vs Scale free network

Scale free network is robust against accidental failures but they are very vulnerable to deliberate attacks and sabotage. (Source: Scale Free Network- by Albert-Laszlo and Eric Bonabeau)

4) TOOLS AND DATABASES USED:

- Cervical Cancer Database (CCDB)_ _ _ _ _ _ _ _ _ _ (Gene Source)
- STRING Database_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _(interaction)
- Geo (Gene Expression Omnibus) Database_ _ _ _ _(for validation)
- Cytoscape_ _(construction of gene network)
- Network Analyzer(Plug-in) _ _ _ _ _ _ _ _ _ _ _ _ _(to study topological properties)
- R language (R studio) _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _(for finding communities)

4.1) CCDB: Cervical Cancer Gene

Cervical cancer gene database is a manually curated database (available online), compiled 537 genes that are associated with different stages of cervical carcinogenesis. Despite a large women population that is affected from this malignancy, no other database on cervical cancer associated gene is present except CCDB. So, CCDB is designed to provide information that may be novel therapeutic treatments for a major cause of cancer within the population of women. In CCDB each gene linked with cervical cancer causation processes such as mutation, methylation, gene amplification, change in expression level and polymorphism, as given in many published literature. The archive of the database contains detailed information about each gene like architecture (exon-intron structure), location, function, sequences (mRNA/CDS/protein), gene ontology, interacting partners, homology to other eukaryotic genomes, structure, and links to other public databases, thus linking CCDB with external data. Manually curated literature references also provided to embrace the presence of the gene in the database (the total number of references is 192).In addition, CCDB procures information on microRNA changed in cervical cancer as well as search facility for querying, several browse options and an online tool for sequence similarity search.

Link for CCDB: http://crdd.osdd.net/raghava/ccdb/

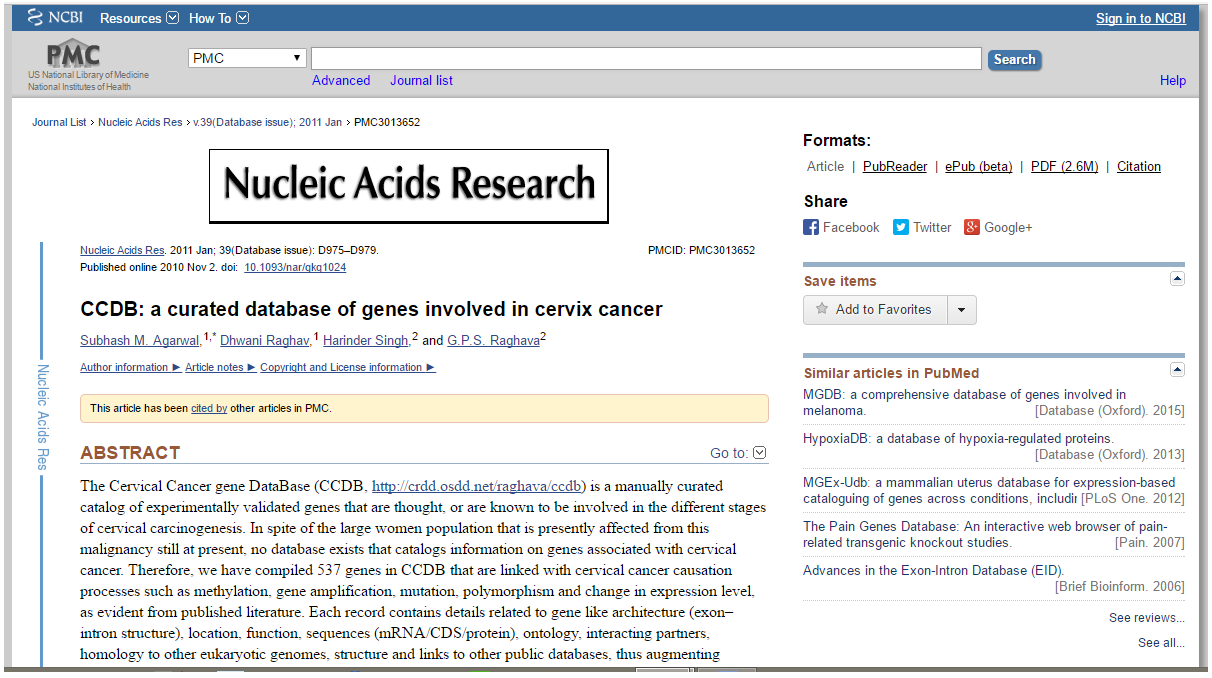


Figure 3. PubMed publication for CCDB (cervical cancer gene database) PubMed id: PMC3013652

Figure 4. Homepage of cervical cancer gene database

4.2) STRING: (Search Tool for the Retrieval of Interacting Genes/Proteins)

In order to understand cellular processes at system-level as well as at molecular level, Protein-protein interaction networks play an important role. These networks can be utilized for annotating structural, functional and evolutionary properties or proteins by filtering and assessing functional genomics data. Exploring the interaction networks that are already predicted, opens a new way for future experimental research and procure cross-species predictions for efficient interaction mapping.

STRING is a biological database of known and predicted protein interactions. The database is freely accessible and updated regularly. The interactions include physical (direct) and functional (indirect) associations and derived from four resources: Genomic context, High-throughput experiments, (conserved) co-expression and previous knowledge. So, STRING extracted interaction data from these sources and quantitatively integrates for a large number of organisms, and exchanges information between these organisms where applicable.

The latest version 10.5 contains information on about 9.6 million proteins from more than 2000 organisms. The database has been developed by a consortium of academic institutions that include: Novo Nordisk Foundation Centre for Protein Research, European Molecular Biology Laboratory (EMBL), University of Copenhagen (UCPH), SIB Swiss Institute of Bioinformatics, TU Dresden (abbreviated as TUD) from German, and University of Zurich.

Link for STRING Database: https://string-db.org/

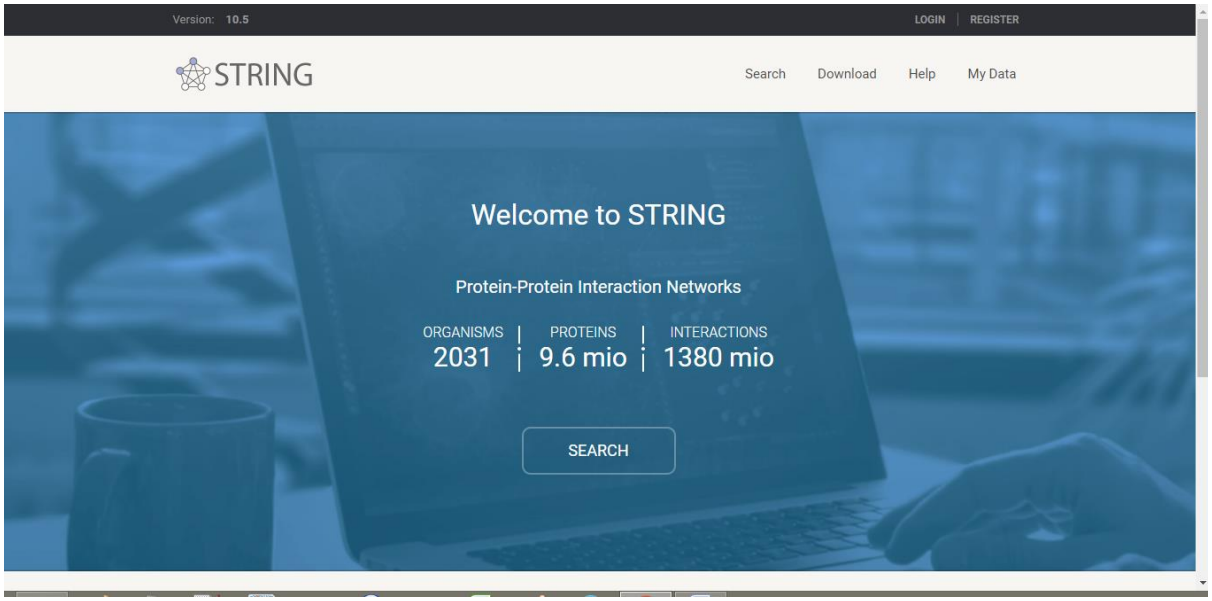


Figure 5. Homepage of STRING database

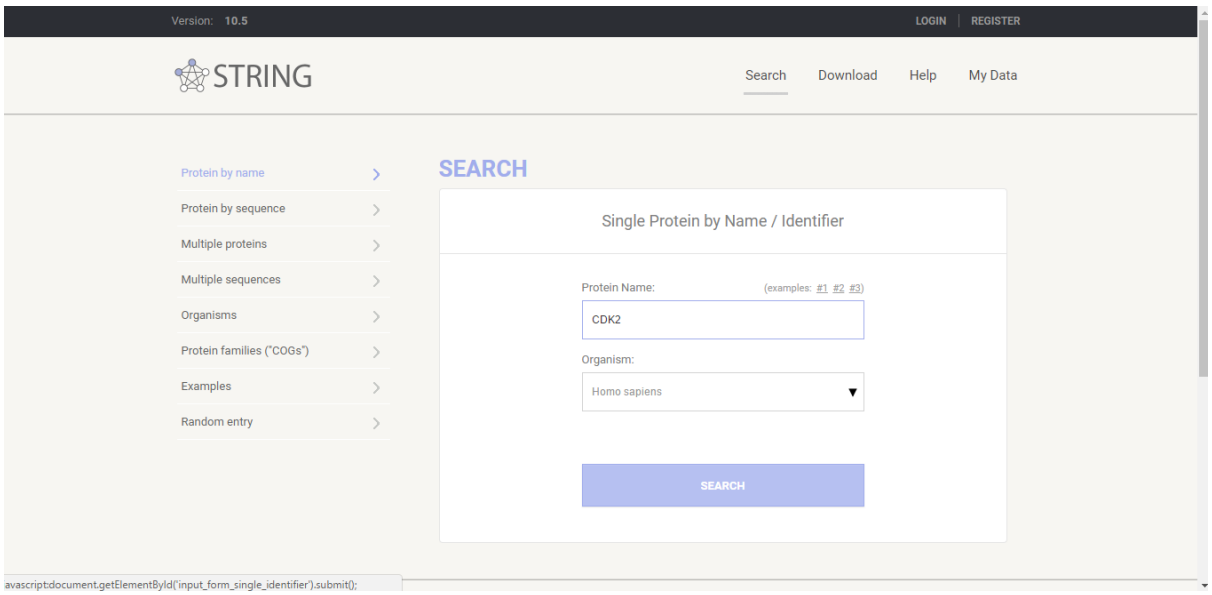


Figure 6. Type the name of protein in first box, say CDK2 (an example), and select the organism of your interest ,say homo sapiens (example) in drop down menu and then click on search button.

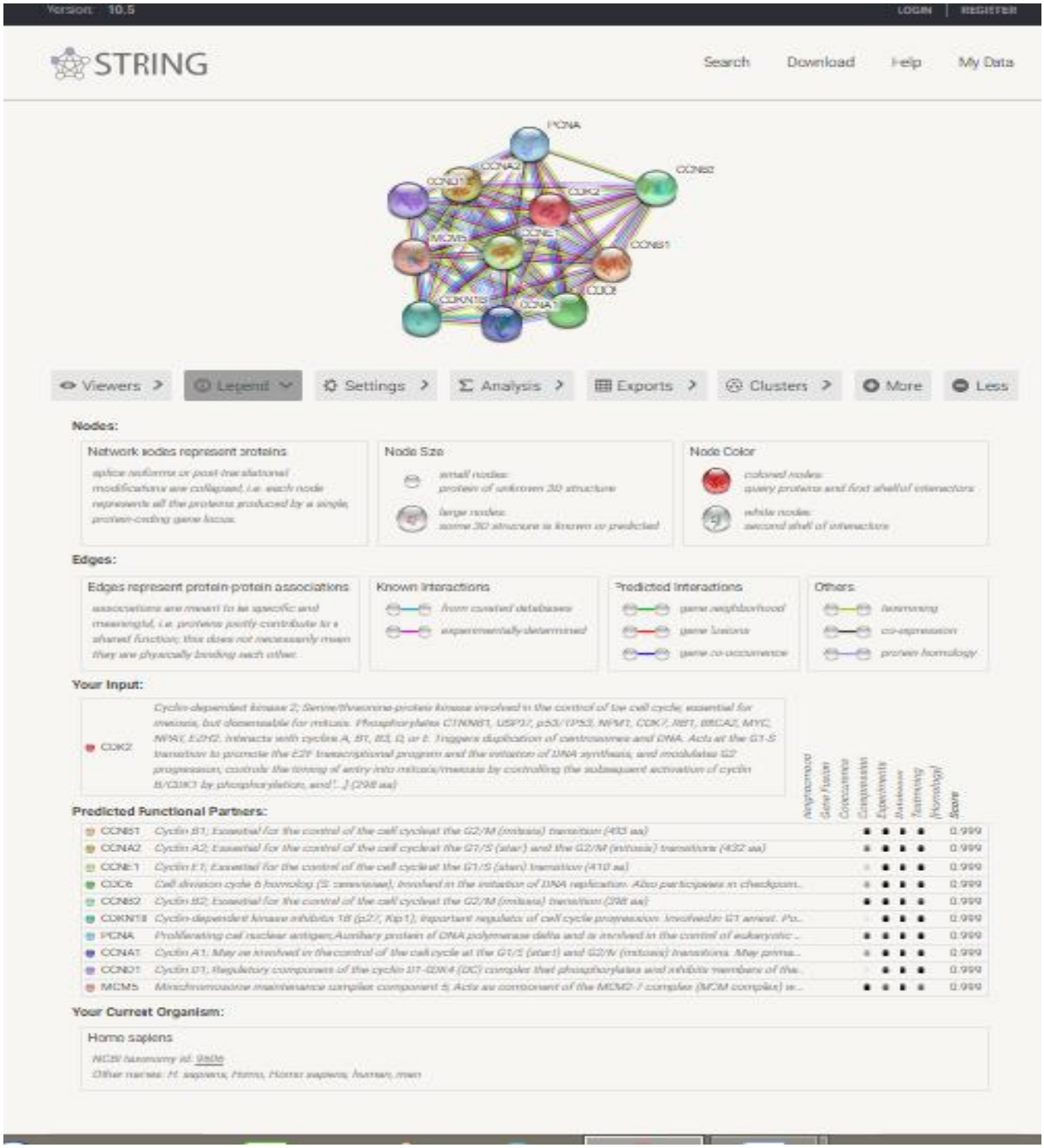


Figure 7. Result showing a picture of interacting partners with gene symbol and list of interacting partners of CDK2 (the protein of our interest)

4.3) GEO: Gene Expression Omnibus:

Expression Omnibus (GEO, created in 2000) at National Center for Biotechnology Information (NCBI) is a freely available repository that contains microarray data, next-generation sequencing (NGS), and other forms of high-throughput functional genomics data sets. In this database, various tools are provided in order to help users for query and download experiments and curated gene expression profiles. The 90% data in GEO database based on gene expression profiles that look into a broad spectrum of biological themes like disease, development, evolution, immunity, ecology, toxicology, and metabolism. Remaining 10% data in GEO about other categories of functional genomics and epigenomic studies covering those that examine genome methylation, the structure of chromatin, copy number variations in the genome and genome-protein interactions. Data in GEO is a collection of original research submitted by the scientific community with a journal agreement that requires data to be made available in a public repository, and the purpose of this to facilitate evaluation of results, reanalysis and full access to all parts of the study. The architecture of database in such way that one easily go through all parts of a study including raw data files, processed data and descriptive metadata that are indexed, cross-linked and searchable. The database is highly flexible that it can handle different style of both processed and raw data in an MIAME- (Minimum Information About a Microarray Experiment) supportive infrastructure that facilitates fully annotated submissions.

At present, GEO stores more than a billion individual gene expression measurements that are derived from over hundred organisms, submitted by more than 1500 laboratories all over the globe addressing a wide range of biological phenomena. Several user-friendly web-based interfaces and apps have been developed that enable effective exploration, query, and visualization of these data, at the level of individual genes or entire studies.

The main goals of GEO are to provide robust, flexible raw data library in which to efficiently store a wide variety of high-throughput functional genomic data sets, simple submission procedures, and user-friendly interface.

Link for GEO database: http://www.ncbi.nlm.nih.gov/geo/.

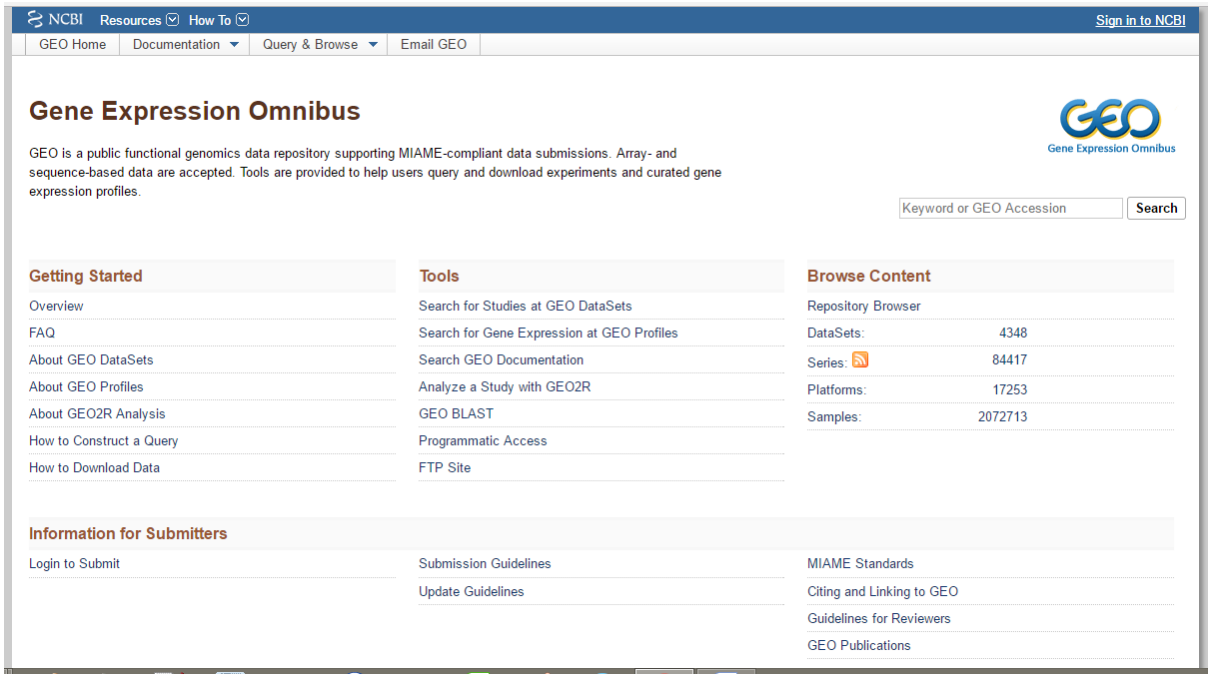


Figure 8. Homepage of gene expression omnibus database

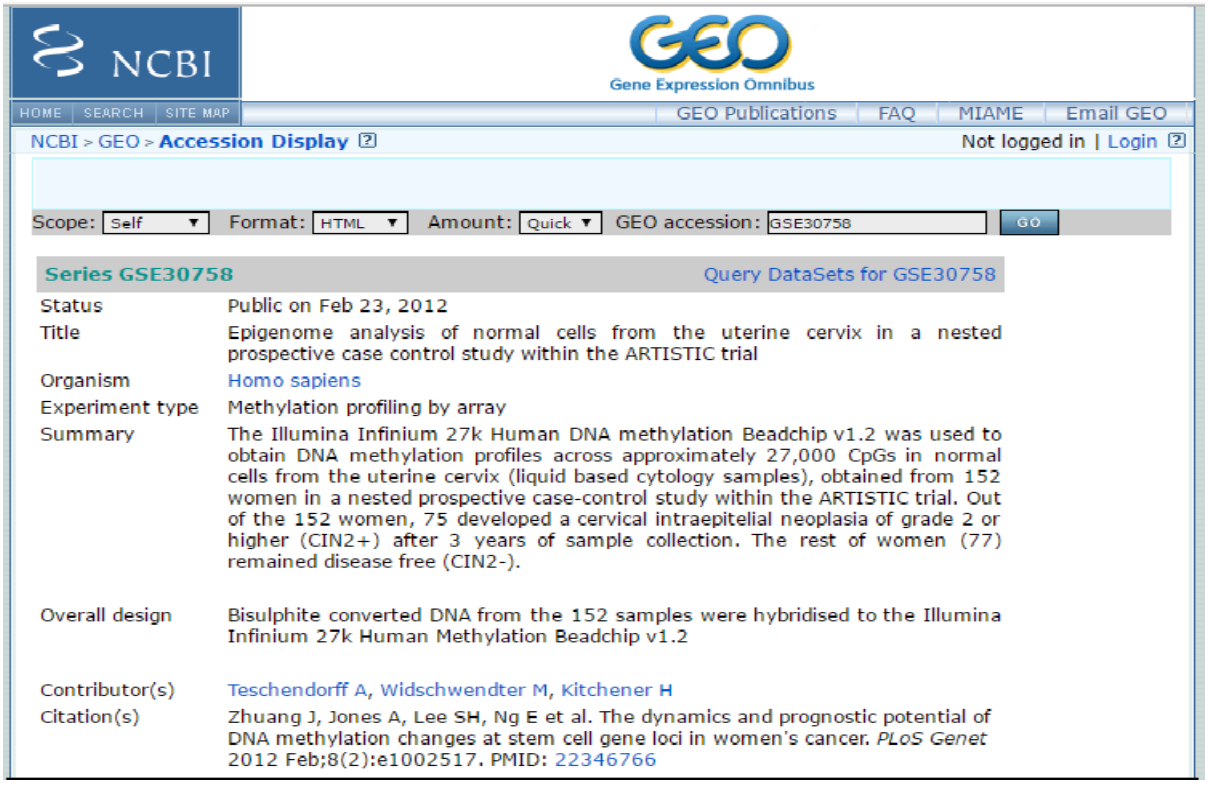


Figure 9. Result obtained after typing GEO accession id (in this case the id is GSE30758) in the search box at the homepage of GEO. From here, one can easily download the micro array data with the help of GEO2R tool.

4.4) GEO2R (GEO-tool):

GEO2R is a web tool of Gene Expression Omnibus database that permits users to compare two or more groups of samples in a GEO series in order to find those genes that are differentially expressed across experimental conditions. Results are obtained in table form, in which genes are arranged according to their significance. The comparison is done by GEO2R tool with the help of two R packages of Bioconductor project: GEOquery and limma.

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 GEOquery R package analyzes GEO data into R data structures so that it can be used by other R packages. The limma (Linear Models for Microarray Analysis) R package has become one of the most widely used statistical tests for identifying differentially expressed genes.

4.5) Cytoscape:

Cytoscape is an open source software platform in bioinformatics that is used for molecular interaction network visualization and linked with high-throughput expression data and other molecular states into a consolidated conceptual framework. Cytoscape is a most reliable tool when used in conjunction with large databases of protein-protein, protein-DNA and genetics interactions. Cytoscape was developed the first time at the Institute for Systems Biology in Seattle in 2002.

Initially, Cytoscape was designed for biological network visualization and research, but at present, it becomes a general platform for complex network studies. Cytoscape provides a large number of apps(known as a plug-in) for data analysis, integration, and visualization. These apps are freely available on Cytoscape app store. Plug-ins can be developed by anyone, using the Cytoscape open API that is based on JAVA language.

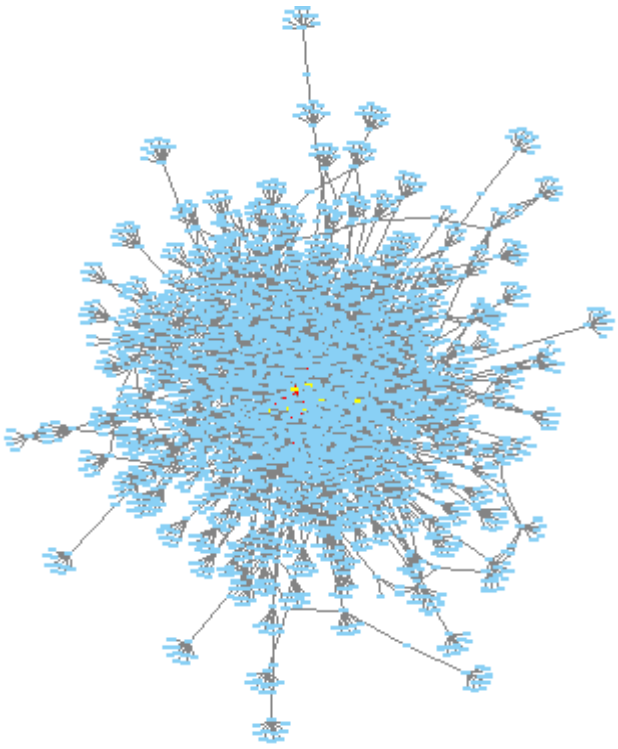
Cytoscape's software core provides basic functionality to layout and queries the complex network, to visually embrace the network with expression profiles, phenotypes, and other molecular states; and to link the network to different databases of functional annotations.

The latest version of Cytoscape is 3.5.1 and one can easily download it from World Wide Web link: "http://www.cytoscape.org/".

Figure 10.

- Cervical Cancer Gene Network

An example of PPI network created by Cytoscape



4.6) Network Analyzer:

NetworkAnalyzer is a very useful Cytoscape plugin that requires no dextrous knowledge in graph theory from the user. By using NetworkAnalyzer a perspicuous set of topological properties and centrality measures for both directed as well as undirected networks, which includes the number of nodes, edges, and connected components, the network diameter, radius, density, centralization, heterogeneity, clustering coefficient and the characteristic path length can be obtained. Besides topological parameters, NetworkAnalyzer plots charts of the degree distribution of nodes, neighborhood connectivities, average clustering coefficients and shortest path lengths.

NetworkAnalyzer generates the intersection, union, and difference of two networks. It gives the facility for extraction of connected components as separate networks and the removal of self-loops.

To run NetworkAnalyzer, select **Tools → NetworkAnalyzer → Network Analysis → Analyze Network**.

4.7) R programming language:

R is a programming language, developed at Bell Laboratories, the USA by John Chambers and colleagues provide an environment or platform for statistical analysis and graphics. R comes under GNU project that is similar to S language. Maximum codes are written for S, easily run in R. R covers a broad area of statistical (linear and nonlinear modeling, classical statistical tests, time-series analysis, classification, clustering) and graphical field and highly flexible.

the main advantage of using R is the ease with which well-designed publication-quality graphs can be obtained, that includes mathematical symbols and different types of formulae where required and user have full control.

R is available as free software under the terms of the Free Software Foundation's GNU General Public License in source code form and it designed so well that it can be easily compiled and run on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

RStudio is a freely available IDE (integrated development environment) for R language and developed by ColdFusion and Hadley Wickham, at present who is Chief Scientist at RStudio. It is available in two editions: one for desktop, where the program runs locally as a regular desktop application and second is RStudio Server, that allows accessing RStudio using a web browser but it is compatible with remote Linux server. RStudio is written in C++ and uses the Qt framework for its graphical user interface. The beta-version of RStudio was officially out in February 2011 and version 1.0 released on 1 Nov 2016.

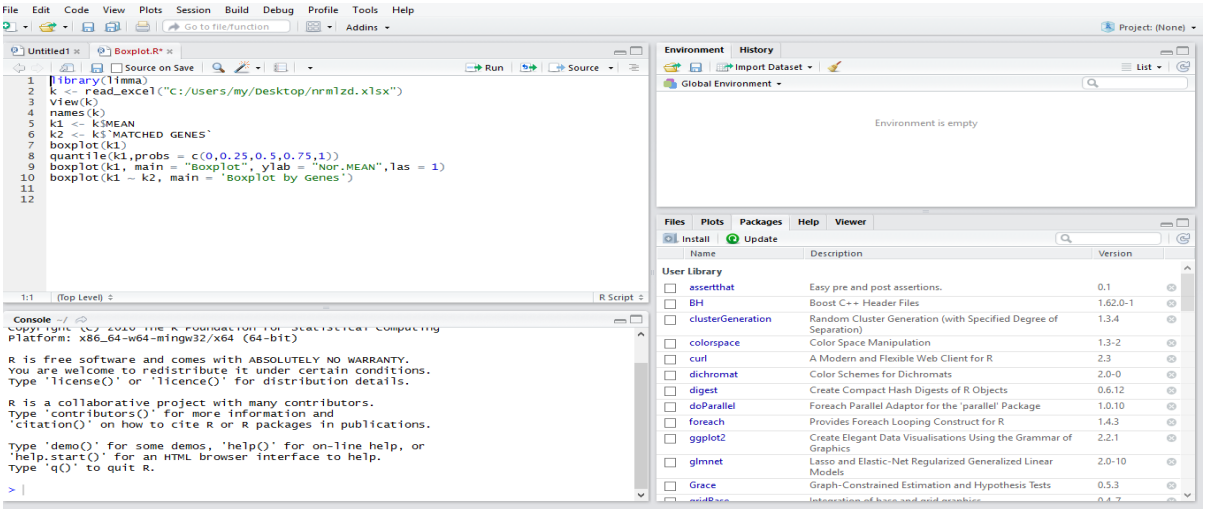
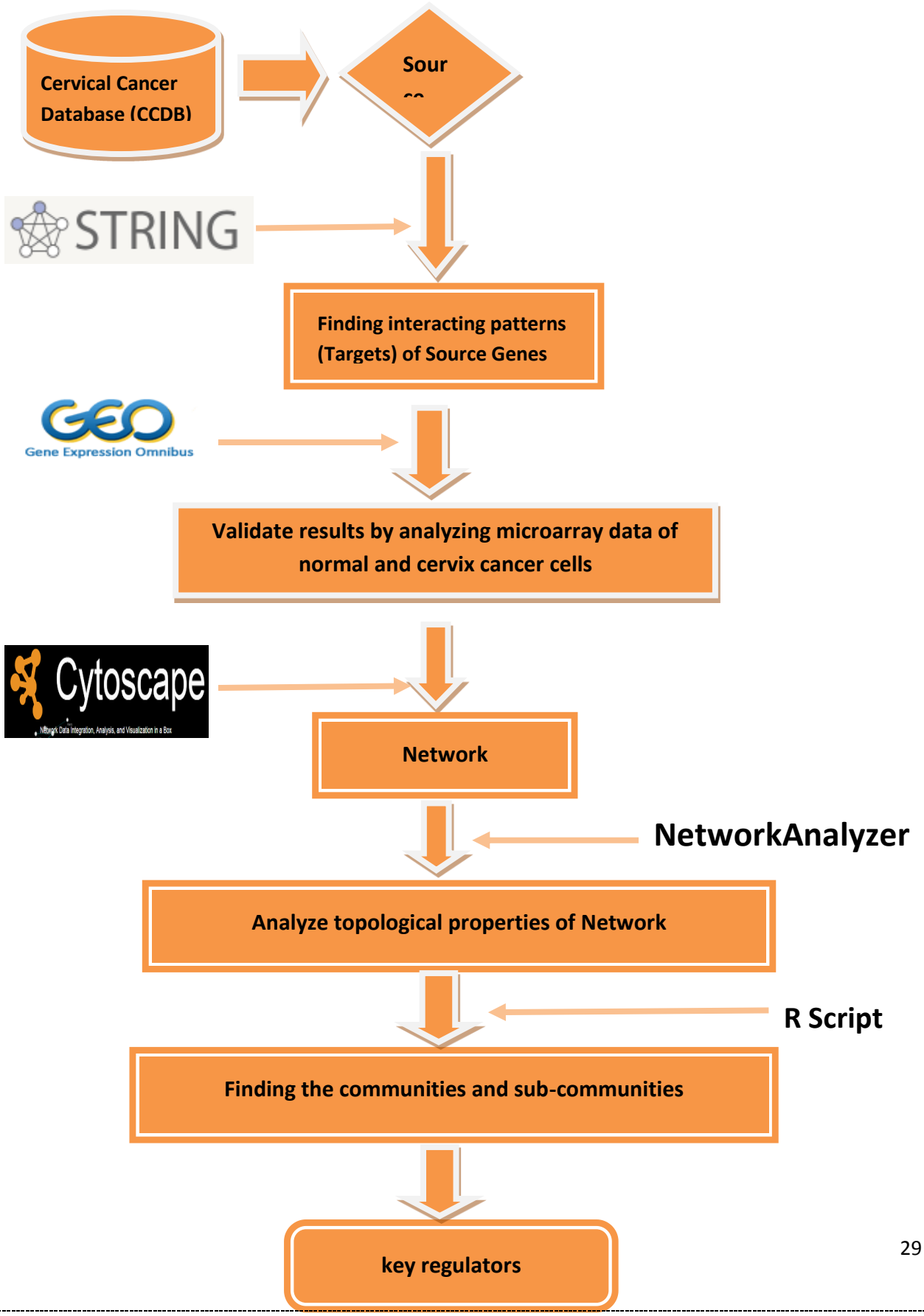


Figure 11. Integrated Development Environment (RStudio) for R programming language.

5. METHODOLOGY:



Step1.

Prepared a gene-list that is associated with cervical cancer from CCDB (cervical cancer database).In CCDB, the total number of non-redundant genes is 537 and a total number of references is 192.

Step2.

By using online tool STRING, retrieved the interacting genes for each gene present in gene-list. (Taking only those genes that have a score greater than 0.90 and also experimentally verified.)

Step3.

Validated the genes (present in gene-list as well as extracted online with the help of STRING tool) by aid of micro array data (for validation, two micro array data are used, one micro array data is for normal cervix cells (GEO ID = GSE30758) and other for cervical cancer cells (GEO ID = GSE27469). Here, we compared the expression level of genes in the normal cell and in the cancer cell.)

Step4.

The network is formed by Cytoscape tool. Initially, there are 3241 nodes and 5241 edges in the cervical cancer gene network. After removing all self-loops and duplication, the total number of nodes is 3099 and 4932 edges.

Step 5.

Studied the topological properties of the network (degree distribution, cluster coefficient, betweenness centrality, neighborhood connectivity) by using a network analyzer, a plug-in of Cytoscape tool. All properties should follow a power law.

Step6.

With the help of R language, we found the communities of the network at a different level. We find the communities until we got single motif. (Motif is a basic building block of a biological network, it contains 3 nodes and 3 edges, in a closed triangular shape.) Initially, we found 15 communities at level 1 and going further we got 6 motifs in cervical cancer associated gene network at level 4.

Step7.

After finding the number of motifs, we checked the influence of nodes that are present in the motif. On the basis of above examination, we conclude that there are 7 genes (TP53, CTNNB1, PTGS2, CDK2, E2F1, CDKN2A, CDKN1A) in the network, which behave as fundamental key-regulators in cervical cancer associated gene network.

Step8.

Study the effects on the network and its topological properties after removing these key-regulators and hubs (hubs are the nodes that have degree greater than average).

6. RESULTS AND DISCUSSION:

Source Gene:

- The total number of genes associated with cervical cancer (as mentioned in Cervical Cancer Gene Database) is 537 and list of genes is given below

Gene List No. 1

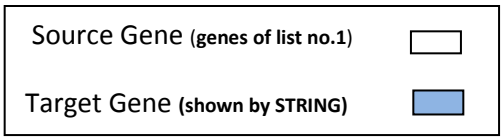
A2M	CDH4	ENG	IL18	MMP14	PTGS2	STK31
ACADVL	CDK2	EPB41L3	IL1A	MMP2	PTK6	STX16
ADAM9	CDK6	EPB41L4B	IL1RN	MMP3	PTPN6	SULF1
ADRM1	CDKN1A	EPCAM	IL32	MMP9	*PTTG1*	SULT2B1
AEBP1	CDKN1B	EPS8	IL6	MRPL11	PWP2H	SYCP2
AHCY	CDKN2A	ERBB2	IL8	MRPS2	PYCARD	SYK
AIM2	CDKN2B	ERBB4	ILF2	MRPS23	RAF1	TARS
AKT1	CEACAM1	EREG	IMP3	MSH2	RALY	TBC1D1
ALDH7A1	CEACAM5	ESR1	INPP5E	MSLN	RAPGEF3	TBX3
ALOX12	CEACAM6	ESR2	IPO5	MSN	RARA	TCEA2
ALOX12B	CEACAM7	EXO1	IRF3	MSX1	RARB	TERT
ANG	CEBPB	EZH2	ITGA1	MT1G	RARG	TFF3
ANLN	CELA2A	FANCA	ITGA4	MUL1	RARRES3	TFPI2
ANXA1	CELSR1	FANCF	ITGA6	MYC	RASSF1	TFRC
ANXA2	CES1	FASTKD3	ITGAV	MYCL1	RBL1	TGFA
ANXA4	CHFR	FGF1	ITGB1	MYCN	RBL2	TGFB1
ANXA8	CITED2	FGF2	ITGB2	MYH9	RBP1	TGFB2
APC	CKS2	FGF7	ITGB8	MYOD1	RCL1	TGFBR1
APLP2	CLDN1	FGFR2	ITM2C	NAE1	RGS1	TGM3
APOC1	CLEC2B	FHIT	IVL	NCL	RHCG	THBD
APOD	CLIC3	FLNA	JAK1	NDRG1	RNASEN	THBS1
APOL2	CNOT1	FLT1	JUNB	NDRG4	RNF114	THBS2
AQP3	COL1A1	FOXL2	KCNC4	NDUFS6	RPL37	TIMP1
ARHGAP12	CPE	FUS	KCNJ4	NEBL	RPL39L	TIMP2
ARHGAP8	CPEB1	GAPDH	KIAA1324	NEK1	RPRM	TIMP3
ASXL1	CPEB3	GARS	KIT	NKX6-1	RPS19	TKTL1
ATIC	CR1	GATA-3	KITLG	NME1	RPSA	TLR2
ATP9A	CRABP2	GDF15	KLF3	NME2	RRAD	TLR4
B4GALT5	CRCT1	GFOD2	KLK10	NNT	RRAS	TLR8
BAG1	CRISP3	GJA1	KLK11	NOL4	RRM2	TMC6
BANF1	CRNN	GLI1	KLK12	NOL7	RTN3	TMCO6
BASP1	CRYAB	GLTP	KLK13	NPTX1	RUNX1	TMEM132A
BAT2L2	CSE1L	GPX3	KLK2	NQO1	RUNX3	TMPRSS11D
BCL2	CSF-1	GRB2	KLK7	NRAS	S100A10	TMPRSS11E
BCL2L1	CSF1R	GRM4	KLK8	NRM	S100A7	TNF
BCL2L2	CTGF	GSS	KRAS	NTF3	S100A8	TNFRSF10C
BICD2	CTNNB1	GSTP1	KRT1	NTN4	S100A9	TOB1
BIRC5	CTSC	GTF2F2	KRT10	OAS3	S100P	TOP2A
BLU	CTSF	GTF2H4	KRT13	OAT	SAA1	TP53
BNIP3	CTSH	GYS2	KRT15	ONECUT1	SCAMP1	TP73

Gene List No 1. to be continued……

BRAF	CTSL1	HBA2	KRT17	ORAOV1	SCD	TPX2
BRCA1	CTSS	HDLBP	KRT18	OSMR	SCEL	TRA2A
BRD9	CUEDC1	HIC1	KRT19	OSR1	SDC1	TREM2
BRIX1	CWH43	HIN-1	KRT3	P2RX4	SDHA	TRIM29
BST2	CXCR2	HIST1H1C	KRT4	PA2G4	SEMG2	TRIP13
BXDC2	CYBA	HIST2H2BE	KRT6B	PAIP1	SEPP1	TSPAN3
BZW2	CYR61	HLA-C	KRT7	PAPD7	SERPINA3	TUBB2
C13ORF18	DAP3	HLA-DQB1	KRT8	PAX1	SERPINB5	TWIST1
C15orf48	DAPK1	HLA-DRA	KRTHA3B	PCDH10	SFN	TYK2
C1QBP	DDIT3	HLA-DRB3	LAMA5	PCNA	SFRP1	TYMS
C20orf20	DDOST	HLA-DRB5	LARP7	PCNP	SFRP2	UFD1L
C20orf30	DDX27	HLTF	LDHA	PDCD1	SFRP4	UPK1A
C3	DEK	HMGB1	LGALS7	PDCD4	SFRP5	VEGFA
C4orf41	DHFR	HOPX	LGR5	PDGFRA	SIPA1	VEGFB
C5orf22	DLC1	HOXA1	LHFPL4	PDIA3	SIX1	VEGFC
C5orf28	DMC1	HOXB6	LMX1A	PDK2	SKP2	VHL
CACYBP	DNAJC9	HOXC10	*LSM3*	PEBP1	SLC3A2	VPS35
CADM1	DNMT1	HPGD	LSMD1	PECAM1	SLC7A6OS	VSIG4
CALCA	DNMT3L	HRAS	LY6D	PGR	SLURP1	VWF
CAMKK2	DSG1	HS3ST2	LYPLA1	PHB	SMAD2	WT1
CAMKV	DSG2	HSP90AA1	MAL	PIGF	SMAD4	XRCC1
CCNA1	DTL	HSPB2	MAP3K11	PIGT	SOCS1	YIF1A
CCNB1	DUSP1	HSPB8	MBNL2	PLAGL1	SOX1	ZBED2
CCND1	DUSP26	HSPD1	MCM2	PLAT	SP100	ZNF217
CCND2	DUT	HUWE1	MCM3	PLOD2	SPA17	CDH3
CCNE1	E2F1	HYAL1	MCM4	PLP2	SPARC	ENDOU
CCR2	E2F2	ICOS	MCM5	PLSCR1	SPINK5	IL10
CD28	E2F3	IER5	MDM2	POFUT1	SPP1	MMP13
CD44	E2F4	IFI16	ME1	POP5	SPRR1A	PTEN
CD63	E2F5	IFI6	MEF2A	POU2F3	SPRR1B	STK11
CD83	ECT2	IFITM1	MET	PPL	SPRR2A	STAU1
CDA	EDN1	IFNG	MGMT	PPP1R3C	SPRR2B	PSMB9
CDC40	EDN2	IGF1	MLH1	PPP2R1B	SPRR3	MMP12
CDH1	EDN3	IGF2	MLLT4	PRDM5	ST14	IGFBP6
CDH11	EDNRA	IGFBP2	MMP1	PRR11	STARD10	EMP1
CDH13	EFNA1	IGFBP3	MMP10	PRSS2	STAT1	CDH2
CDH15	EGFR	IGFBP5	MMP11	PRSS3	STAT3	

End of Gene List No.1

In order to construct network, all target genes should be known, so that a complete network of genes associated with cervical cancer can be obtained.



Target gene:

- By using online tool STRING (Search Tool for Retrieval of Interacting Genes/Proteins), we have found all the interacting partners of each gene present in list 1. Now, the total number of genes is 3241. Only those genes are included in the list whose interacting score greater than 0.90.
- The list of 3241 genes is given below

Source	Target genes									
A2M	TGFB1	APP	KLK3	SERPINA1	ALB	PLG	SERPINF2	SERPING1	AMBP	APOA1
ACADVL	HADHA	EHHADH	HADHB	ECHS1	ACAA2	ACAA1	CPT2	FOXA2	HADH	GCH1
ADAM9	MAD2L2	SNX9	COL17A1	TNF	UBC	REG4	TAC1	SH3GL2	ITGAV	MAD2L1
ADRM1	UCHL5	PSMD7	PSMD4	PSMD14	PSMD5	UBC	PSMD1	PSMC5	PSMD2	PSMC4
AEBP1	PTEN	TCF3	SHH	PCSK2	MYOCD	TAGLN	SMTN			
AHCY	UBC	BHMT	MTR	CBS	MAT1A	MAT2A	MTRR	XIAP	CTH	DNMT1
AIM2	PYCARD	CASP1	NLRP3	DDX58	EIF2AK2	NLRP1	LAMP1	SQSTM1	BECN1	NLRC4
AKT1	MTOR	FOXO1	RICTOR	NOS3	HSP90AA1	MDM2	FOXO3	ILK	PTEN	GSK3B
ALDH7A1	GAD1	AASS	AGXT2L2	GAD2	EHHADH	ABAT	ACSS1	ACSS2	CHDH	HIBADH
ALOX12	PLA2G2A	PLA2G4A	PTGS2	PTGS1	PLA2G6	PLA2G10	CYP2B6	CYP4F3	CYP4F2	CYP4A11
ALOX12B	PLA2G4A	PTGS1	PTGS2	PLA2G2F	PLA2G3	PLA2G5	PLA2G10	ALOX15	PLA2G4E	PLA2G4B
ANG	RNH1	VEGFA	RNASE8	TARDBP	RNASE6	SOD1	FUS	FST	SELV	TP53
ANLN	RACGAP1	CD2AP	ECT2	MCM5	KIF23	SEP07	SH3KBP1	MYL2	MYL9	MYL12B
ANXA1	FPR1	UBC	S100A11	FPR2	FPR3	GNA15	GNA11	APP	TRPM7	NR3C1
ANXA2	S100A10	PLG	UBC	CDC42	TP53	S100A4	PCSK9	HSP90AA1	ERBB2	SLPI
ANXA4	NFKB1	UBC	TP53	LAMP2	ANXA1	ANXA3	ANXA7	MPST	DNAJB11	PADI4
ANXA8	SRPX2	HLF	AES	TLE6	TLE2	COBLL1	TCF3	GRIN1	TFPT	TM4SF1
APC	CTNNB1	AXIN1	GSK3B	ARHGEF4	CSNK1E	BTRC	SIAH1	DVL1	AXIN2	CTBP1
APLP2	APBB1	PSEN1	MAPK8	UBC	APBB3	MAPK9	PRNP	APBA3	MASP2	KAT5
APOC1	APOA1	APOA2	APOC4	PEPD	APOC3	APOE	APOC2	APOB	EXOC1	LPL
APOD	LCAT	APOA1	APOA2	NOTCH4	NOTCH3	JAG1	ZG16	RARA	LCN1	CSTB
APOL2	UBC	SP1	COMT	AR	OR2AK2	CHRM4	PRODH	TEX33		
AQP3	AQP2	STX4	SLC14A2	EXOC3	EXOC4	GK	INS	AVP	PPARG	KLF4
ARHGAP12	RHOA	RAC3	RAC2	CDC42	RHOU	RHOG	RHOV	RHOH	MCF2	ARHGDIA
ARHGAP8	RHOA	CTTN	CDC42	RAC3	RHOC	RHOU	RHOB	ARHGEF6	FGD1	AKAP13
ASXL1	BAP1	EZH2	SUZ12	EED	UBC	EZH1	CBX2	RARA	BMI1	PCGF2
ATIC	GART	PFAS	PAICS	ADSL	SHMT2	SHMT1	MTHFD1	UBC	PPAT	ADSS
ATP9A	MON2	PTGS2	UBC	PTGS1	ARL1	TMEM30C	TMEM30A	TMEM30B	UBE3D	MYH14
B4GALT5	GCNT1	MUC12	MUC5B	MUC5AC	MUC4	MUC15	MUC2	MUC6	MUC20	MUC19
BAG1	HSPA8	HSPA4	BCL2	STUB1	RAF1	NR3C1	AR	ESR1	JUN	PARK7
BANF1	VRK1	ACTL6A	ANKLE2	EMD	SMARCA4	VRK2	VRK3	CRX	PPP2CA	PPP2R2A
BASP1	ATP5A1	WT1	UBC	MARCKS	GAP43	HSPA5	EFTUD2	TRNT1	UBXN1	INSR
BAT2L2	CSNK2A1	UBC	HUWE1	RPL38	HINFP	AES	PMM2	ABHD16A	PMM1	NLRP9
BCL2	BCL2L11	BAD	BECN1	TP53	BIK	BAX	BID	BAK1	FKBP8	MAPK8
BCL2L1	BCL2L11	BIK	APAF1	BID	BAX	BAK1	BAD	BBC3	BECN1	TP53
BCL2L2	BCL2L11	MCL1	BIK	BAD	AKT1	BAX	BID	PABPC1	MAPK7	TP53BP2
BICD2	RAB6A	NEK9	DYNC1H1	EXD1	DCTN2	DYNC1I1	DCTN1	PLK1	PAFAH1B1	NEK8
BIRC5	AURKB	INCENP	CDCA8	CDK1	BIRC2	CASP9	AURKA	PLK1	DIABLO	XPO1
BLU	UBE2V2	UBE2V1	RNF8	TRAF6	UBC	OTUB1	STAMBPL1	STUB1	PARK2	MAP3K7
BNIP3	BCL2	BNIP3L	RHEB	HIF1A	MAP1LC3B	TMEM11	ARNT	FOXO3	HRAS	TNF
BRAF	MAP2K1	MAP2K2	RAF1	HRAS	NRAS	YWHAQ	RAP1A	KSR1	MAPK3	PRKCA
BRCA1	TP53	ATM	ESR1	H2AFX	CHEK2	RBBP8	UIMC1	FAM175A	PALB2	BAb RD1
BRD9	NEO1	SS18	DPF1	BCL7C	DPF3	SS18L1	NOVA2	C2orf62	NOVA1	TRIP12
BRIX1	EBNA1BP2	RPF2	NIP7	WDR12	RPF1	MKI67IP	NOP2	PWP1	RBM28	PES1
BST2	BTRC	TAB2	ARHGAP	IRF7	UBC	IRF1	PLVAP	TRIM5	TFRC	APOBEC3G

Gene List No.2 to be continued.....

			44							
BXDC2	EBNA1BP2	RPF2	NIP7	WDR12	RPF1	MKI67IP	NOP2	PWP1	RBM28	PES1
BZW2	UBC	EIF4G3	EIF4G1	BZW1	PSTPIP1	CNBP	LZTFL1	ABHD14A	GINS3	EIF6
C13ORF18	BECN1	KRTAP19-4	KRTAP13-4	KRTAP22-1	KRTAP19-6	KRTAP21-2	KRTAP11-1	ARMC7	GDAP1L1	KRTAP13-2
C15orf48	NDUFA4									
C1QBP	KNG1	MMP14	SRSF1	F12	NUP153	PRCP	C1QA	F11	NCAPD2	SERPING1
C20orf20	MORF4L1	DMAP1	MORF4L2	MEAF6	EPC1	KAT5	VPS72	ACTL6A	YEATS4	BRD8
C20orf30	UBC	CXorf56	EMC7	UTP23	C4orf46	TIGD2	C1orf109	CNEP1R1		
C3	CFH	CD46	CR1	VSIG4	ITGAM	CR2	CFI	CFD	CFP	CFHR1
C4orf41	TRAPPC3	TECPR1	UBC	TRAPPC12	TRAPPC8	VPS18	TRAPPC4	C5orf44	NCKIPSD	TRAPPC2L
C5orf22	UBC	ELOF1	ELF1	ACAT1						
C5orf28	PIANP	KIAA0947								
CACYBP	SIAH1	CTNNB1	S100A6	SKP1	TBL1X	UBC	GSK3B	AXIN1	DACT1	APC
CADM1	EPB41L3	CRTAM	MPP3	RAC1	CASK	PVRL3	MPP6	CADM3	TIAM1	DLG3
CALCA	CALCRL	RAMP1	IAPP	ADM	CALCR	PTH	ADM2	SCT	ADCYAP1	NPS
CAMKK2	CAMK4	PRKAG1	PRKAG2	CALM1	PRKAG3	PRKAB1	PRKAA2	PRKAA1	CAMK2B	CAMKK1
CAMKV	PPP3R2	CHP1	TESC	CIB1	CHP2	CIB4	PPP3R1	F11	ZNF576	RHBDL3
CCNA1	CDK2	CDK1	CDKN1A	CDC25A	CDKN1B	CCNB1	E2F1	CDT1	FZR1	RB1
CCNB1	CDK1	CDK2	CDC20	FZR1	ANAPC11	CDKN1A	PLK1	UBC	CDC27	PTCH1
CCND1	CDK1	CDK2	CDC20	FZR1	ANAPC11	CDKN1A	PLK1	UBC	CDC27	PTCH1
CCND2	CDK4	CDK6	CDKN1A	RB1	CDK2	CDK5	CDKN1B	CDKN2A	MYC	STAT5A
CCNE1	CDK2	RB1	CDKN1B	CDKN1A	FBXW7	CDK1	UBC	HIST1H1A	E2F3	CDC25A
CCR2	CCL2	CCL11	CCL8	CCL13	CCL5	CCL7	CXCL12	CXCL10	CCL16	CCR5
CD28	CD80	PIK3R1	GRB2	GRAP2	VAV1	LCK	CD86	ITK	CD247	IGSF11
CD44	SPP1	ERBB2	MMP9	EZR	SELE	NF2	HMMR	MSN	LCK	ICAM1
CD63	SRC	TIMP1	UBC	ITGB1	SCARB2	SYT7	LAMP1	PTK2	LAMP2	NPY
CD83	TNF	FOXP3	LMNA	DDX41	CD40	CD86	CD80	IL18	ITGAX	CCR7
CDA	TYMP	DCK	UPP1	UCKL1	PNP	UPP2	UCK2	NT5E	NT5C2	TK1
CDC40	MAGOH	DHX38	PRPF19	PRPF18	PRPF8	SNRPA1	SLU7	SNRNP200	SF3B1	NCBP2
CDH1	CTNNB1	CTNND1	CBLL1	CTNNA1	EGFR	PSEN1	SRC	UBC	VCL	SNAI1
CDH11	CTNND1	CDH2	CDH13	CDH17	CDH3	CDH15	CTNNA1	CDH4	CDH1	CDH5
CDH13	CDH17	ADIPOQ	CDH8	CDH5	CDH11	CDH10	CDH15	CDH9	CDH7	CDH3
CDH15	CTNND1	CTNNB1	CTNNA1	CDH9	CDH7	CDON	CDH13	CDH17	CDH5	CDH8
CDH2	CTNNB1	CTNND1	CTNNA1	AXIN1	GRIA2	GJA1	PTPN1	LRP5	MMP9	FGFR1
CDH3	CTNNA1	CTNND1	CTNNB1	CDH1	CDH13	CDH17	CDH5	CDH8	CDH11	CDH12
CDH4	CTNNA1	CDH2	CTNNB1	CTNND1	CDH17	CDH13	CDH8	CDH7	CDH11	CDH18
CDK2	CDKN1A	CCNE1	CDKN1B	CCNA2	CCNB1	CCNA1	RB1	CDC25A	CCNE2	RBL1
CDK6	CCND3	CCND1	CDKN2A	RB1	CDKN2C	CDKN2D	CDKN1B	CCND2	CDKN1A	CDKN2B
CDKN1A	CDK2	TP53	CDK4	CCND1	PCNA	CCNE1	SKP2	CCNB1	CCND3	CCND2
CDKN1B	CDK2	CDK4	CCND1	SKP2	CCNE1	CCND3	CDK6	UBC	CCNA2	AKT1
CDKN2A	CDK4	CDK6	MDM2	TP53	MYC	NPM1	CDKN1A	CCND1	UBC	RB1
CDKN2B	CDK4	CDK6	SMAD4	MYC	SMAD3	CDKN1B	CDKN1A	SMAD2	CCND1	ZBTB17
CEACAM1	CEACAM8	CEACAM6	FLNA	CD209	PTPN11	SOX9	VEGFA	ZAP70	LCK	CRH
CEACAM5	VPS33B	VIPAS39	ADRM1	VPS33A	HNRNPM	STX12	SELE	CD209	STX8	MCMDC2
CEACAM6	CEACAM8	CEACAM1	SRC	MCMDC2	XKR9	CCDC3	KRTAP3-1	LACTB2		
CEACAM7	CD209	CYP2A13	CYP2A7							
CEBPB	ESR1	CEBPA	NFKB1	SMARCA2	RELA	IL6	MYB	RPS6KA1	IL8	SMAD4
CELA2A	CPA1	APLP2	PNLIPRP1	CELA3A	CELA3B	PNLIP	CPB1	NT5C2	CPA2	GFI1
CELSR1	VANGL2	FZD6	VANGL1	PRICKLE1	FZD3	ESR1	PRICKLE2	STMN1	DVL1	DVL3
CES1	CES2	CES3	CES5A	CES4A	TFDP2	TFDP1	E2F1	APOE	IL6	FOSL1
CHFR	UBE2D2	AURKA	UBC	PARP1	UBE2D1	PLK1	HLTF	APLF	UBE2N	USP7
CITED2	EP300	PPARA	HIF1A	CREBBP	NRIP1	EPAS1	NCOA1	ARNT	APOA1	ELK1
CKS2	CDK1	CDK2	CCNB1	CCNA2	CCNB2	CDK3	UBC	CCNA1	SKP2	CCNF
CLDN1	TJP1	TJP2	TJP3	INADL	F11R	MPDZ	CD81	MMP14	CLDN5	PVRL3
CLEC2B	CLEC3A	MT-CYB	HCCS	LYZ	MRC1	NPR2	PRDM10	ALB	NPPC	NKRF
CLIC3	MAPK15	CFTR	GABRP	DNM1	CLCNKB	CLIC4	CLIC2	CLCNKA	CLIC5	CLIC6
CNOT1	CNOT2	RQCD1	CNOT4	CNOT7	CNOT8	CNOT6L	CNOT6	CNOT3	DDX6	C2orf29
COL1A1	ITGA2	COL7A1	ITGB1	LEPRE1	P4HB	COL1A2	FN1	SP7	SPARC	COL4A5
CPE	INS	ARF6	KIF5B	CPB2	PCSK1	CPB1	PSMA6	GCG	PCSK2	CCK
CPEB1	AURKA	PLK1	NGDN	PABPC1L	BICC1	PAPD4	PARN	PUM2	EIF4E1B	AURKC

Gene List No.2 to be continued…….

CPEB3	CNOT7	TOB1	CNOT6L	CNOT6	CNOT8	CNOT1	CNOT10	RQCD1	TNKS1BP1	C2orf29
CR1	C3	C4A	C4B	CFI	CD46	C5AR1	IL8	CR2	CDC42	TGM1
CRABP2	RARS	CCND3	UBE2I	UBC	ALDH1A1	IFRD1	HPSE2	CNNM1	INS	CREB1
CRCT1	UBAP2L	C1orf43	LCE6A	KPRP	LCE3E	C8orf58	S100A7A	KIAA1614	GOLGA6L9	LCE3D
CRISP3	A1BG	KNG1	MTDH	TMPRSS2	DIP2B	HPN	HOPX	AMACR		
CRNN	C10orf10	KPRP	WWOX	S100A7A						
CRYAB	CRYBB2	CRYGC	HSF4	VEGFA	FBXO4	TP53	DRD2	ROCK2	CRYBB3	MAPK14
CSE1L	RAN	UBC	KPNA4	KPNA1	UNC45A	USP1	TP53AIP1	RRM2B	KPNA6	TP53I3
CSF-1	CSF1R	CBL	ITGB3	ITGAV	HRAS	BCAR1	PTPN1	PDGFRB	CD163	KIT
CSF1R	CBL	CSF1	SPI1	GRB2	MYB	IL34	SHC1	ITGAV	SOCS1	PIK3R1
CTGF	WWTR1	YAP1	TGFB1	VEGFA	EGFR	PPARA	SMAD3	SMAD2	MMP3	TGFBR2
CTNNB1	AXIN1	GSK3B	TCF7L2	LEF1	CDH1	CDH2	CDH5	PSEN1	AR	CTNNA1
CTSC	BMP4	NOG	HSP90B1	GLT6D1	CD81	GZMA	CTSL1	ANO6	CDIPT	CST7
CTSF	CD74	CTSS	IFI30	HLA-DRB5	HLA-DRB1	CST4	ATG4B	GABARAPL1	GABARAPL2	M6PR
CTSH	INS	SFTPB	CSTB	CSTA	CTSD	CST4	LGMN	CST3	BGLAP	C1QTNF6
CTSL1	CD74	CSTA	IFI30	CTSS	HLA-DRB1	HLA-DRB5	GOPC	SERPINB3	FOXO1	CST3
CTSS	CD74	CTSD	LGMN	SERPINB4	IFI30	UNC93B1	TLR9	B2M	TLR7	TLR3
CUEDC1	UBC	TOM1L2	EIF3D	FECH	C19orf57	TCEAL3	TOM1	CALB1	C11orf1	ANKRD13D
CWH43	MOGS	PGAP2	PGAP1	A2ML1	PGAP3	CERS5	DNASE1L1	ATP12A	CLGN	CANX
CXCR2	IL8	CXCL1	PPBP	CXCL2	CXCL5	CXCL3	CXCL6	CXCL12	ARRB1	PPP2CA
CYBA	NCF2	NCF1	NOX1	NCF4	NOX3	RAC1	NOXO1	NOX4	CYBB	DUOX1
CYR61	MIS12	ZWINT	ITGAV	ITGB5	FOS	JUN	CTNNB1	MKL1	SRF	RHOA
DAP3	FADD	MRPS2	MRPS10	PFAS	MRPS5	TNFRSF10A	HIBCH	MRPS6	MRPS15	METTL17
DAPK1	UNC5B	FADD	BECN1	UBC	CALM2	HOXB7	HSP90AA1	PDCD6	CASP3	NTN1
DDIT3	ATF4	MAPK14	ATF2	BCL2	ATF3	JDP2	ATF6	CEBPB	CEBPG	MYC
DDOST	RPN2	DAD1	RPN1	STT3B	STT3A	SEC61G	TUSC3	MAGT1	SSR3	PRKCSH
DDX27	KIAA0020	NOP2	WDR12	RBM28	DDX54	PES1	DDX56	FTSJ3	RPF2	BRIX1
DEK	KAT2B	UBC	CSNK2A1	RELA	HIST1H4A	GSK3B	SRRM1	TP63	PRDX6	U2AF2
DHFR	TYMS	SHMT1	SP1	FPGS	HSPD1	MTHFD1	PCNA	SHMT2	MTHFD1L	SPR
DLC1	TRAF3IP3	RHOA	RHOC	RHOB	TNS1	TENC1	CDC42	TNS3	PLCD1	RAC3
DMC1	RPA1	MRE11A	BLM	RAD50	BRCA2	TOP3A	FANCM	RAD54B	RAD51	ATM
DNAJC9	FAM149B1	UBC	HSPA8	HSPA4	HYOU1	HSPA14	HSPA1L	HSPA1A	HSPA1B	HSPA12A
DNMT1	HDAC1	DNMT3B	PCNA	DNMT3A	UHRF1BP1	EZH2	HDAC2	USP7	EHMT2	SMARCA5
DNMT3L	DNMT3A	DNMT3B	RSL24D1	HDAC1	TDG	RPL24	PHPT1	DNMT1	HIST2H2AC	TRDMT1
DSG1	PKP1	PKP2	CASP3	DDX41	PKP3	KLK5	MET	JUP	DSP	KRT1
DSG2	PKP2	CASP3	JUP	UBC	PKP3	DSC2	DSP	KRT74	TMEM43	CTNND2
DTL	DDB1	CUL4A	CUL4B	CDKN1A	PCNA	SETD8	UBC	CDT1	FBXW5	TP53
DUSP1	MAPK14	MAPK3	MAPK1	MAPK8	MAPK12	UBC	TP53	JUN	MAPK11	ATF2
DUSP26	AK2	PES1	HSF4	TP53	MAPK12	MAPK6	MAPK7	MAPK14	MAPK4	MAPK15
DUT	PPARA	TYMS	DTYMK	DCTD	RRM1	ENSG00000260851	UNG	RFC3	TK1	PCNA
E2F1	RB1	TFDP1	TFDP2	RBL1	CCNA2	UBC	TOPBP1	SP1	KAT2B	CCNA1
E2F2	RB1	TFDP1	TFDP2	RBL1	CCNE1	CCNE2	CCNA2	MYC	CDK2	CDK1
E2F3	RB1	TFDP1	TFDP2	RBL1	CCNE1	CCNE2	CCNA2	CDK2	CREBBP	CDK4
E2F4	RBL1	RBL2	RB1	TFDP1	TFDP2	CDK2	LIN9	HDAC1	CCND1	CDK1
E2F5	RBL1	RBL2	TFDP1	TFDP2	CDK2	EP300	CCNE2	CREBBP	SMAD3	CCNE1
ECT2	RACGAP1	KIF23	PLK1	RHOA	RHOB	RHOC	RHOG	CDC42	RAC1	RAC3
EDN1	EDNRA	EDNRB	AKT1	AVP	JUN	FOS	EGFR	PPARA	GATA4	HIF1A
EDN2	EDNRA	EDNRB	MCHR1	GNRH1	KNG1	EDN3	AVP	EDN1	TAC1	APP
EDN3	EDNRB	EDNRA	GNRH1	AVP	KNG1	NMB	CCK	OXT	MLN	HCRT
EDNRA	EDN1	EDN2	EDN3	MMP1	AGT	GNA11	KNG1	GNRH1	AVP	TAC1
EFNA1	EPHA2	EPHA4	EPHA1	EPHA5	EPHA7	EPHA6	EPHA3	EPHB3	ARHGEF15	EPHB1
EGFR	SHC1	GRB2	STAT3	CBL	PLCG1	EPS15	CBLB	PTPN1	SRC	HSP90AA1
EMP1	LOH12CR1	SLC34A3	MYO1A	CYP27B1	PTH	CDC42EP2	SMIM3	SLC34A1	FGF23	CACNG6
ENDOU	STS	TPK1	PPEF1	LGALS13	HSPB11	HEBP2	PP632	NUTF2	IGFBP1	TCEAL1
ENG	TGFBR1	TGFB1	ACVRL1	TGFBR2	ACVR1	SMAD4	BMP2	ACVR2A	SMAD3	HIF1A
EPB41L3	MAGOH	CNTNAP2	CADM1	YWHAB	YWHAG	YWHAH	YWHAQ	CNTNAP1	TJP2	CASK

Gene List No.2 to be continued…….

EPB41L4B	TAAR1	TP53	PIN1	RPRD1A	METTL21C	KLHL3	MPP5	PCCA		
EPCAM	CLDN7	GPNMB	TP53	MYC	TSPAN8	CCND1	NFKB1	CDH1	PTPRC	PROM1
EPS8	BAIAP2	ABI1	SOS1	USP6NL	EGFR	CDC42	RAB5A	SRC	ABL1	RAC1
ERBB2	GRB2	HSP90AA1	UBC	SHC1	ERBB3	EGFR	CTNNB1	PTPN11	ERBB2IP	SRC
ERBB4	NRG1	NRG2	SHC2	STAT5A	GRB2	SHC1	NRG4	CBLB	DLG4	ERBB2
EREG	EGFR	HBEGF	ERBB4	ERBB3	NRG4	NRG2	ERBB2	NRG1	AREG	EGF
ESR1	NCOA3	SRC	SP1	BRCA1	EP300	NCOA2	NCOA1	CREBBP	NRIP1	HDAC1
ESR2	NCOA3	NCOA1	ESR1	NCOA2	SRC	MED1	NOS3	FOS	NR0B1	NCOR2
EXO1	MSH2	MLH1	PCNA	PMS2	MRE11A	RECQL	WRN	RAD50	RBBP8	BLM
EZH2	EED	SUZ12	YY1	DNMT1	DNMT3A	HDAC1	RBBP4	DNMT3B	BMI1	CDK1
FANCA	FANCG	FANCC	FANCF	FANCL	FANCM	FANCE	C1orf86	FANCD2	TOP3A	FANCB
FANCF	FANCG	FANCA	FANCE	FANCC	FANCM	FANCL	APITD1	FANCD2	TOP3A	HES1
FASTKD3	ACYP2									
FGF1	FGFR1	FGFR2	FGFR4	FGFR3	S100A13	FGF23	HSPG2	SYT1	GRB2	FGF2
FGF2	FGFR1	FGFR4	FGFR2	FGFR3	HSPG2	SDC1	SDC4	GPC1	SDC2	NANOG
FGF7	FGFR2	FGFR3	FGFR1	EGFR	FGFR4	FGFBP1	SHH	FGF2	FGF6	FGF1
FGFR2	FGF1	FGF7	FGF4	FGF8	FGF10	FGF6	FGF3	FGF2	GRB2	FGF9
FHIT	SRC	HSPD1	UBE2I	DNMT3B	FDXR	CDKN1A	JUN	HSPE1	CCND1	AKT1
FLNA	RALA	ITGB7	ARHGAP24	CEACAM1	FBLIM1	PAK1	FILIP1	CDC42	CAV1	ITGB2
FLT1	VEGFA	PGF	VEGFB	NRP1	HIF1A	NRP2	SHC1	EPAS1	PLCG1	PTPN11
FOXL2	NR5A1	SMAD3	CYP19A1	LATS1	DDX20	SOX9	RSPO1	UBE2I	SIRT1	LHX4
FUS	SF3A2	YBX1	SRSF2	HNRNPA3	PTBP1	SRRM1	HNRNPU	SRSF4	HNRNPK	POLR2A
GAPDH	UBC	ENO1	PGK2	RPL13A	PGK1	HTT	BPGM	ALDOA	TP53	ALDOB
GARS	IARS	AARS	NARS	QARS	EPRS	RARS	ASNS	DARS	GPHN	UBA1
GATA-3	ZFPM2	NKX2-5	MEF2C	HAND2	SRF	NFATC4	BMP4	SMAD4	EDN1	ZFPM1
GDF15	TP53	EGR1	ATF3	AKT1	IL6	FOS	HAMP	MAPK8	HIF1A	SP1
GFOD2	GNE	HLTF	UBC	B4GALNT1	OXSM	DCXR	ZNF12	CHMP2A	IMPDH2	ARPC3
GJA1	TJP1	UBC	CDH2	MAPK7	SRC	JUN	MAPK1	UBQLN4	MAPK3	CTNNB1
GLI1	SUFU	PTCH1	STK36	SHH	FOXA2	SMO	IHH	PTCH2	SAP18	HDAC1
GLTP	PLEKHA3	UBC	CEL	AP4B1	CTSA	ANKRD6	PMPCB	SEC63	VEPH1	DMRTB1
GPX3	SOD3	GSS	GSTT1	GSTO1	GSTT2B	GSTO2	GGT1	GSTM3	TNF	GSTP1
GRB2	SOS1	SHC1	EGFR	CBL	GAB2	LCP2	ERBB2	FRS2	IRS1	GAB1
GRM4	APP	GNAT3	HTR1A	CNR1	ADORA1	HRH3	NPY	RGS4	HTR1B	DRD2
GSS	GCLC	GSR	GCLM	GGCT	GGT7	GPX1	GPX2	GPX3	GPX6	GGT1
GSTP1	MAPK8	GSTT1	CYP1A1	CYP2E1	EPHX1	GSTT2B	CYP1B1	PRDX6	CYP1A2	GSTO1
GTF2F2	GTF2F1	POLR2E	POLR2G	GTF2B	POLR2H	POLR2B	SUPT5H	POLR2A	POLR2C	CTDP1
GTF2H4	ERCC3	GTF2H3	ERCC2	GTF2H1	MNAT1	GTF2H5	GTF2H2C	GTF2H2	H2AFV	H2AFZ
GYS2	PYGL	PYGB	GBE1	GYG1	GYG2	UBC	AGL	UGP2	EPM2A	PYGM
HBA2	HBB	AHSP	HBD	HBG2	HP	HBE1	HBA1	CYB5R3	SCN2A	INS
HDLBP	VPS25	TTC4	VPS35	VPS36	DNAJC2	CTCF	UPF1	GNB2L1	VPS26A	MOGAT3
HIC1	SIRT1	TP53	CTBP1	MTA1	CTBP2	ARID1A	RBBP4	PHF19	E2F1	NFKBIB
HIN-1	MARCO	SCGB1A1	HOXB5	RASSF1	CRABP1	ENSG00000257184	HOXA9	RARB		
HIST1H1C	MAGOH	UBC	CDK1	HIST1H2AD	H1FX	TP53	HNRNPUL1	KRT16	GSTA3	GSTT1
HIST2H2BE	H2AFB2	H2AFB3	HIST1H2AI	HIST1H2AL	HIST1H2AK	HIST1H2AG	HIST1H2AA	HIST1H2AM	HIST1H2AD	H2AFZ
HLA-C	CD8A	UBC	B2M	KIR2DL3	KIR2DL1	CALR	ERAP1	NCAM1	KLRD1	KLRC1
HLA-DQB1	HLA-DQA1	CIITA	CD4	TOR1A	INS	HLA-DRB5	CD74	TNFRSF25	CD8A	BRD2
HLA-DRA	MS4A1	CD74	HLA-DRB5	HLA-DRB1	UBC	HLA-DMA	MBP	36951	PRPF40B	CD63
HLA-DRB3	HLA-DQA1	CIITA	CD4	TOR1A	INS	HLA-DRB5	CD74	TNFRSF25	CD8A	BRD2
HLA-DRB5	CD74	OAS3	ICAM1	OAS1	IRF5	CD3E	IRF9	OASL	MT2A	CIITA
HLTF	UBC	RAD18	CHFR	TCEB1	SHPRH	HIST2H2AC	RECQL	BLM	POLR2I	WRN
HMGB1	TLR4	TLR2	TP53	KAT2B	TBP	NFKB1	AGER	CREBBP	RELA	SMARCC1
HOPX	SRF	TBX21	HDAC2	HSP90AA1	EPC1	HSPA4	TSPO	HADH	AMBP	UBL4A
HOXA1	FGF3	PBX1	CDH1	FOXQ1	STAT3	PBX2	PKNOX1	CHN1	EFEMP2	DFNA5
HOXB6	CREBBP	TRIM37	PKNOX1	FGF2	PBX1	PBX2	HOXB7	ITGB3	CLK3	GRN
HOXC10	MLL4	MLL2	TBX4	PBX3	INTS12	PHF1	PHF19	MTF2	PBX1	PBX2
HPGD	PTGR1	HPGDS	LTC4S	ALOX12	ALOX5AP	PTGS2	SLCO2A1	MMP9	NFIL3	HDAC9
HRAS	RAF1	RALGDS	SOS1	PIK3CA	RASA1	BRAF	EGFR	ZHX2	RIN1	SRC
HS3ST2	GLCE	GPC2	GPC5	GPC4	GPC1	GPC6	GPC3	SDC1	SDC2	SDC4
HSP90AA1	SUGT1	STIP1	PTGES3	CDC37	AHSA1	PPP5C	ERBB2	AKT1	HIF1A	UBC

Gene List No.2 to be continued…….

HSPB2	MAPKAPK2	HSF1	MAPKAPK5	AKT1	MAPK14	JUN	DAXX	HSPBAP1	TP53	RHOA
HSPB8	BAG3	HSPB7	HSPB1	HSPB2	KHDRBS1	CCND1	MAP3K7	BMPR1A	HSPB3	MN1
HSPD1	HSPE1	UBC	HSPA9	HSP90AA1	DHFR	HSPA5	GAPDH	HSPA8	HSP90AB1	SF3A2
HUWE1	UBC	MCL1	UBE2D1	MYC	UBE2L3	RANGAP1	MAX	USP9X	UBL4A	TP53
HYAL1										
ICOS	ICOSLG	CD40LG	STAT3	CTLA4	P85B	PIK3R5	PIK3CD	PIK3CG	PIK3CB	PIK3CA
IER5	CDC25B	MPDZ	PDZK1IP1	MYCL1	LCMT2	CCT8	TAL1	TAL2	ETS1	CCT7
IFI16	TP53	TMEM173	UBC	DDX58	ID2	TP53BP1	BRCA1	RUNX2	UBTF	ID1
IFI6	ISG15	MX1	IFIT1	MX2	IFIT3	IFITM1	IFIT2	OAS1	IRF7	IFITM3
IFITM1	CD81	CD19	CR2	IFI6	IFIT1	MX1	IFIT3	IFI27	ISG15	IFI35
IFNG	IFNGR1	STAT1	TNF	TBX21	IRF1	IL18R1	IL2	IFNGR2	IL18	TGFB1
IGF1	IGFBP3	IGFBP1	IGFBP4	IGFBP5	IGF1R	IGFBP2	INS	IRS1	INSR	TGFB1
IGF2	IGFBP3	IGF2R	IGFBP1	IGFBP2	IGF1R	IGFBP5	IGFBP6	GPC3	IGF2BP2	VEGFA
IGFBP2	IGF1	IGF2	INS	MIIP	LEP	PTEN	TP53	FGF2	IL6	NFKB1
IGFBP3	IGF1	IGF2	IGFALS	PLG	TP53	MAPK8	ADAM12	TF	MMP2	F2
IGFBP5	IGF1	IGF2	IGFALS	PAPPA2	PAPPA	FHL2	VTN	THBS1	SERPINE1	PTH
IGFBP6	IGF2	IGF1	SOAT2	CSTA	KRT2	IGFALS	TF	SP1	FZD8	WFIKKN1
IL10	IL10RA	IL10RB	GATA3	JUN	IL6	IL4	IL2	FOXP3	IL8	STAT6
IL18	CASP1	IL1B	IL18R1	IL18BP	IL18RAP	IFNG	NFKB1	IL13	IL10	IL6
IL1A	IL1B	IL1R1	IL1R2	IL1RN	IL1RAP	IL6	S100A13	CSF2	MYD88	FOS
IL1RN	IL1R1	IL1A	IL1R2	MYD88	IL1B	IL1RAP	IRAK3	IRAK2	TOLLIP	HNRNPD
IL32	PRTN3	PRKCD	IL8	AKT1	IL1B	STAT3	HSPG2	PRKCE	CD80	UBC
IL6	IL6R	IL6ST	STAT3	NFKB1	CEBPB	JUN	OSM	STAT1	SOCS3	FOS
IL8	CXCR2	RELA	CXCR1	DARC	NFKB1	CXCL12	JUN	CEBPB	EGFR	FOS
ILF2	MAGOH	ILF3	UBC	PRKDC	XRCC5	DHX9	SF3A2	ADAR	EIF4A3	YBX1
IMP3	IMP4	BYSL	MPHOSPH10	MRPS11	RPS14	RPS18	KRR1	PWP2	RPS3	WDR75
INPP5E	SACM1L	PI4KA	PI4KB	PIP5K1A	PIK3CB	PIK3CG	PTEN	PIK3CA	PIK3C2B	PIK3CD
IPO5	NUP98	UBC	RAN	RPL23A	KPNB1	RANBP1	NUP153	IPO4	RAG2	SENP1
IRF3	TBK1	CREBBP	EP300	UBC	JUN	MAVS	TMEM173	RELA	IFIT1	ATF2
ITGA1	ITGB1	LAMA1	SEMA7A	TLN1	COL13A1	ITGB2	COL1A1	ITGB5	COL4A1	LAMA3
ITGA4	ITGB7	ITGB1	PXN	YWHAZ	ADAM28	MADCAM1	CD81	FN1	VCAM1	JAM3
ITGA6	ITGB1	ITGB4	PLEC	CD151	COL17A1	CD9	LAMA5	GRB2	ITGB5	ITGB6
ITGAV	ITGB3	FN1	ITGB5	ITGB1	ITGB6	PTK2	CYR61	SPP1	CD47	EDIL3
ITGB1	ITGA5	ITGA8	ITGA6	ITGA4	ITGA7	ITGA9	ITGA2	ITGA1	ITGAL	ILK
ITGB2	ITGAX	ICAM1	ITGAL	ICAM2	ITGAM	ITGAD	ITGA8	TLN1	ICAM3	ITGA5
ITGB8	ITGAV	ITGA7	ITGA8	ITGA9	ITGA5	ITGA2B	ITGA6	TGFBR1	ITGAL	ITGA2
ITM2C	UBC	RIT2	FURIN	BACE1	TCF4	ELSPBP1	RPSA	NDUFAF3	SLC7A5	NPAT
IVL	LOR	SP1	FLG	KRT16	FOSL1	JUND	JUNB	EP300	KRT1	JUN
JAK1	STAT3	STAT5A	STAT6	STAT1	STAT5B	IL2RB	PTPN11	IRS1	STAT2	SOCS3
JUNB	FOS	FOSL1	MAPK8	FOSL2	BATF	FOSB	SMAD3	MAPK9	EP300	NOS3
KCNC4	KCNH2	KCNAB2	KCNAB1	KCNAB3	KCNH1	KCNH6	KCNH3	KCNQ4	KCNH8	KCNH7
KCNJ4	DLG4	DLG1	KCNJ2	LIN7A	CASK	KCNJ14	TAX1BP3	GNAQ	DLG2	LIN7C
KIAA1324	HACE1	C4orf27	PLBD1	ISOC1	FAM110B	SPEF2	UFSP2	NABP1	RGPD3	C9orf152
KIT	GRB2	PIK3R1	PTPN11	STAT1	EPOR	SOCS6	STAT5B	PIK3CG	DOK1	
KITLG	KIT	EPOR	EPO	GDF9	JAK2	GRB2	MATK	MYB	SHC1	STAT3
KLF3	CTBP2	WEE1	CTBP1	FHL3	GATA1	ZNF516	CBX2	CBX4	LHX8	XRN1
KLK10	STK11IP	NFS1	XPO1	KRT10	LAP3					
KLK11	HP	HPR	CRLF2	GJB4	PKP3	TSLP	AGL	ZDHHC20	LAP3	DSC3
KLK12	ACPT	HPR	LAP3	SPINK6						
KLK13	SIGLEC9	SERPINF2	SERPINA1	A2M	LAP3					
KLK2	AR	NCOA2	KDM4C	CARM1	SRC	GNB2L1	MMP8	ACPP	SERPINA3	SERPINB6
KLK7	CDH1	A2ML1	SPINK5	GLIS1	SPRR1B	DSG1	DSC1	FLG	CDSN	SPINK9
KLK8	SPINK9	DEFA3	DEFA4	CAMP	DEFB1	SPINK6	SPINK5	PRR15	TMEM62	LAP3
KRAS	RAF1	EGFR	RALGDS	PIK3CA	SOS1	BCL2	TP53	PTPN11	MAP2K1	ARAF
KRT1	KRT10	MAGOH	KRT2	CTNNB1	LEF1	LOR	IVL	UBC	RASD2	SUMO2
KRT10	KRT2	KRT1	MAGOH	CCND1	CDK9	NSD1	CXCL14	UBC	CMA1	PI3

Gene List No.2 to be continued…….

KRT13	IVL	SLC4A3	KRT7	FLG						
KRT15	UBC	PRL	TELO2	HR	CCDC120	AMOTL2	KRT6A	ARC	C1orf216	LDOC1
KRT17	JUN	FOS	UBC	CCHCR1	IL22	KRT6A	UCHL1	USP1	CALR	ILF3
KRT18	KRT8	UBC	CASP6	YWHAZ	USP32	USP10	YWHAQ	TCHP	EGR1	DEDD
KRT19	MAGOH	USP32	USP10	TP53	MAPK14	MAPK11	ANXA2	KLF4	MEIS1	SIRT2
KRT3	KRT14	KRT5	UBC	LMOD1	HP1BP3	HNRNPU	ITGAV	ABCG2	KRT4	AP2M1
KRT4	UCHL1	NSMAF	USP21	MT3	MT4	USP25	USPL1	MT1A	USP1	KRT3
KRT6B	UBC	EIF4A3	KRT14	TNFAIP6	MEGF11	MEGF6	PEAR1	DKK2	MEGF10	SCARF2
KRT7	UBC	SLC4A3	CDX2	CDX1	NKX2-1	FST	MME	MUC2	PIP	MUC1
KRT8	KRT18	MAPK14	UBC	PPL	MAPK11	TP53	USP32	USP10	TDGF1	TCHP
KRTHA3B	CT45A5	RASL10A	DEFB103B	PADI3						
LAMA5	LAMC1	DAG1	ITGA6	ITGB1	LAMB1	LAMB2	LAMB3	ITGB4	ITGA3	LAMA1
LARP7	CDK9	MEPCE	HEXIM1	RRS1	LYAR	CCNT1	FBL	NHP2	RPP38	NHP2L1
LDHA	UBC	HIF1A	ACLY	ARNT	EP300	PKM	MYC	CREB1	JUN	CREBBP
LGALS7	TP53	MMP9	BCL2	USP38	LGALS7B	USP15	TUBG1	USP1	WIPI1	CCDC134
LGR5	RSPO1	RSPO2	LRP6	ZNRF3	RSPO3	RSPO4	RNF43	UBC	UBB	UBA52
LHFPL4	TSPAN4	PLEKHA8	TOM1L1	SEC63	DSTN	CA2	ENPEP			
LMX1A	NR4A2	WNT7A	MSX1	FGF2	WNT1	ATOH1	FGF8	FGF13	OTX2	LDB1
LSM3	LSM5	LSM6	LSM7	LSM4	LSM1	LSM2	NAA38	SNRPE	SNRPD3	PATL1
LSMD1	NAA35	NAA30	NAA11	NAA15	NAA16	UBC	NAA10	N4BP2L2	RPL26	SLC7A6OS
LY6D	CLPS	NDUFAF3	TCF3	ELSPBP1	ITM2C	KLF2	CD59	EIF4A2	SLC7A5	NPAT
LYPLA1	PLA2G4A	LPCAT1	LPCAT2	PLA2G2A	PLA2G12A	PLA2G12B	PLA2G4B	PLA2G2E	ENSG00000169970	LPCAT4
MAL	DIAPH1	ACTA1	SRF	HNF1A	USP7	MYD88	SRC	LYN	PTK2	SYK
MAP3K11	CDC42	MAPK8IP1	MAP2K7	RAC1	MAP4K1	MAPK8IP2	UBC	MAPK8	NF2	MAPK8IP3
MBNL2	CNBP	CCT3	CELF1	TNNT2	CELF2	PAPD5	RBFOX2	CEBPD	DMPK	RBFOX3
MCM2	MCM5	CDC7	MCM4	MCM3	MCM7	MCM6	ORC1	CDC45	CDC6	ORC5
MCM3	MCM5	MCM7	MCM2	MCM4	MCM6	CDC45	MCMBP	CDK2	ORC5	DBF4
MCM4	MCM7	MCM2	MCM3	MCM6	MCM5	CDC7	MCMBP	ORC1	DBF4	CDK2
MCM5	MCM3	MCM2	MCM7	MCM6	MCM4	CDC7	MCMBP	ORC1	CDC45	CDC6
MDM2	TP53	CDKN2A	USP7	AKT1	RB1	RPL11	RPL5	SUMO1	UBE2D1	MDM4
ME1	MDH1	PKM	LDHC	LDHB	FH	PC	MDH2	LDHAL6B	LDHAL6A	EHHADH
MEF2A	MAPK14	HDAC9	MAPK7	MYOD1	MYOG	HDAC5	EP300	MYF6	HDAC4	HDAC7
MET	HGF	CBL	GRB2	GAB1	UBC	CTNNB1	CDH1	PTPN1	PTPN11	PLXNB1
MGMT	TRIP12	TP53	NRAS	KRAS	HRAS	UBC	MLH1	IL24	HIF1A	MYB
MLH1	MSH2	MSH6	PMS2	BLM	MSH4	EXO1	MSH3	MLH3	BRIP1	PCNA
MLLT4	PVRL2	TJP1	PVRL1	HRAS	RAP1A	PVRL3	PVRL4	RRAS	RAP1B	PTPN11
MMP1	TIMP1	BSG	ETS1	JUN	FOS	FOSL1	CTSG	SDC1	TIMP2	ACAN
MMP10	TIMP2	TIMP1	MMP9	BCAN	MMP2	MMP19	MMP1	MMP7	MMP8	PRSS1
MMP11	FURIN	MMP9	MMP2	MMP19	MMP8	MMP7	MMP1	CD40	S100A4	TIMP1
MMP12										
MMP13	TIMP1	RUNX2	TIMP2	TP53	ETS1	POSTN	CYR61	IL1B	TIMP3	S100B
MMP14	TIMP2	TIMP1	EPAS1	FURIN	LAMB3	SP1	LUM	MMP2	C1QBP	ADI1
MMP2	TIMP2	TIMP3	COL18A1	TIMP1	JUN	TP53	THBS1	SRC	TGFB1	VEGFA
MMP3	TIMP1	TNC	RAC1	RAC3	MAPK8	TIMP3	SNAI1	TIMP2	CTGF	SDC4
MMP9	TIMP1	PLG	THBS1	JUN	CD44	TIMP2	NFKB1	FOSL1	SRC	TGFB1
MRPL11	MRPL16	MRPL2	MRPL1	MRPL13	MRPL15	MRPL4	MRPL22	RPL10A	MRPL17	MRPL24
MRPS2	MRPS16	MRPS9	MRPS5	MRPL13	MRPS10	MRPS12	MRPL22	MRPL1	MRPS14	MRPS15
MRPS23	MRPS25	QARS	MRPS27	MRPS18B	MRPL54	PFDN5	MRPS14	MRPS22	MRPL32	FAM91A1
MSH2	MLH1	PMS2	EXO1	MSH6	MSH3	PCNA	ERCC1	PMS1	TP53	XRCC6
MSLN	MUC16	NDUFB5	ZBTB43	RALA	NDUFB9	RNMTL1	LAMTOR3	WNT5A	WFDC2	MPV17
MSN	LRRK2	ROCK1	SLC9A3R1	ROCK2	RHOA	UBC	SPN	MAPK8	CD44	ICAM3
MSX1	TBP	PAX9	BMP4	FGF8	WNT1	BMP2	WNT4	MAGED1	LHX2	MYOD1
MT1G	CDH1	SPINK7	NQO1	APRT	MT4	GSR	PLA2G2A	CRABP1	PTPN1	ENSG00000269375
MUL1	UBC	VPS35	NR6A1	UBE2D2	MFN1	MFN2	UBE2D1	UBE2D3	UBE2E3	DNM1L
MYC	ZBTB17	CDKN2A	MAX	FBXW7	SUPT3H	PFDN5	EP300	GSK3B	SKP2	KAT5
MYCL1	MAX	RLF	C1QBP	PPIE	PPT1	REEP5	MXD1	TRIT1	TP53	COL9A2
MYCN	MAX	NDRG1	NTRK1	CDKN2A	TRRAP	AURKA	SIRT1	SP1	CCND1	PTEN
MYH9	MYL6	MYH10	MAGOH	MYL9	UBC	MYL12B	S100A4	MYL12A	ACTB	RHOA

Gene List No.2 to be continued…….

MYOD1	EP300	KAT2B	TCF3	MEF2C	MEF2A	CREBBP	MYB	ID1	SMAD3	JUN
NAE1	UBA3	NEDD8	UBE2M	APP	NCSTN	PSEN1	APLP2	PPP1R13B	UBE2F	PSEN2
NCL	NPM1	HSP90AA1	MYC	ARFGEF1	TP53	MDM2	TOP1	FBL	RPL11	RPL3
NDRG1	TP53	MYC	MYCN	HIF1A	PTEN	CDH1	BCL2L11	ARNT	CUL2	CD82
NDRG4	PTEN	ALCAM	GNB1	GNB2	GNB4	GNB3	MYCN	C2orf65	MYOM2	GMFG
NDUFS6	NDUFA2	NDUFAB1	NDUFS2	NDUFS8	NDUFS7	NDUFS4	NDUFA5	NDUFB10	NDUFA6	NDUFV2
NEBL	ZYX	MYPN	HAND2	DAZAP1	XIRP2	ACTN2	ACTN1	ACTB	XIRP1	NFIB
NEK1	DYNC2H1	VDAC1	XRCC5	VHL	KIF3A	FEZ2	FEZ1	MAP1LC3A	ATRX	TSC2
NKX6-1	FOXA2	ONECUT1	NEUROG 3	HNF1A	ISL1	MAFA	GCG	FGF10	OLIG2	PAX6
NME1	TIAM1	RAC1	SET	DNM1	ARF6	CDC42	UBC	APEX1	ANP32A	DNM2
NME2	MYC	NME1	POLR1B	DTYMK	UBC	GUK1	CMPK2	RRM1	POLR2B	PKM
NNT	NADSYN1	IDH2	NMNAT3	NMNAT1	NMNAT2	ENPP1	ENPP3	BST1	NUDT12	CD38
NOL4	SNAPC5	CTBP1	CTBP2	PAF1						
NOL7	MYC	UBC	THBS1	MRTO4	UTP6	PSME3	UTP11L	RPL7L1	NOP16	MKI67IP
NPTX1	OPA1	NIT2	RCN2	SLITRK2	TBCD	NADKD1	TCOF1	HDC	IMMP2L	SLITRK1
NQO1	ODC1	ENSG00000 255439	GGCX	TP53	UBC	APRT	ADK	MT1X	MT1H	MT1G
NRAS	PIK3CA	RAF1	EGFR	BRAF	PIK3CG	ERBB2	PIK3CD	ERBB3	RALGDS	SOS1
NRM	TMEM192	C14orf166	TCTE1	DSE	PDXP	IHH	TST	RCN1	NEFH	NHS
NTF3	NTRK3	NTRK2	NTRK1	BDNF	NGFR	CREB1	SHC1	FGF2	GRB2	PLCG1
NTN4	DCC	UNC5A	NEO1	KRTAP5-9	UNC5B	PRKAA2	DSCAM	ALOX12	UNC5D	LAMC1
OAS3	MX1	IFIT1	MX2	RNASEL	TRIP12	IFIT3	IRF7	ISG15	IFIT2	IFI27
OAT	OTC	ALDH18A1	ARG2	ARG1	ALDH4A1	ODC1	ENSG00000 114786	UBC	ASS1	NAGS
ONECUT1	NEUROG3	PDX1	FOXA2	HNF1B	HNF1A	NKX6-1	HNF4A	FOXM1	STAT5B	STAT5A
ORAOV1	YAE1D1	ABCE1	FGF19	FGF3	FGF4	PPFIA1	CTTN	ANO3	MSH3	MYEOV
OSMR	OSM	JAK1	IL6ST	JAK2	CTF1	CNTF	IL6	IL11	CLCF1	SOCS3
OSR1	VHL	SOX9	EGLN1	RUNX2	SHH	NOTCH1	NOTCH3	NOTCH2	IHH	DHH
P2RX4	CCL21	CCL2	IRF5	GATA2	IL2	PLA2G1B	P2RY12	P2RY1	P2RY13	P2RY2
PA2G4	ERBB3	NMD3	GTPBP4	AR	MRTO4	UBC	GNL3	RRP12	GNL3L	NIP7
PAIP1	PABPC1	EIF4A1	EIF4A2	EIF4E	EIF4B	CNOT1	EIF4G1	PARN	EIF4A3	CNOT7
PAPD7	EXOSC10	PAPOLA	SKIV2L2	ZCCHC7	TUT1	POLR3GL	POLI	POLM	ZCCHC9	URI1
PAX1	MEOX1	HOXA3	SHH	BMP4	NKX3-2	NOG	TCF15	WNT1	GLI1	HOXB3
PCDH10	BEX1	SRPX2	C3orf58	DLG4	RAD23A	PSMC2	KCNQ2	CMTM5	CMTM3	CSPG5
PCNA	FEN1	RFC3	RFC4	RFC5	RFC1	POLH	POLD1	LIG1	CDKN1A	MSH6
PCNP	UHRF2	UBC	NPDC1	ENSG00000 267022	ZNF223	PSPH	CA7	NPR2	NPR3	
PDCD1	CD274	PDCD1LG2	PTPN11	CD8A	HLA-DRB1	LCK	CSK	CD3D	CD3E	CD3G
PDCD4	EIF4A1	RPS6KB1	BTRC	UBC	EIF4A2	RPS6KA1	EIF4G1	FBXW11	RPS13	MYB
PDGFRA	PDGFC	PDGFA	PDGFB	PIK3R1	CRK	PDGFRB	CRKL	GRB2	PLCG1	CAV1
PDIA3	CALR	CANX	ERP27	PPIB	HLA-A	HSPA5	UBC	ERAP1	ERO1L	GANAB
PDK2	PDHX	PDHA1	AKT1	DLD	PRKCB	RPS6KB1	PRKCA	PDHB	DLAT	PDK1
PEBP1	RAF1	UBC	IKBKB	PRKCZ	CHUK	PRKCA	AURKB	BIRC5	CDCA8	INCENP
PECAM1	FLI1	CD34	XG	EWSR1	PRDM1	SLC4A3	ARSF	ARSH	SHOX	KRT7
PGR	SRC	NCOA1	NCOA3	ESR1	CDK2	SP1	CCND1	MAPK1	ERBB2	BRCA1
PHB	TP53	PHB2	UBC	RAF1	E2F1	SMARCA4	HDAC1	RB1	BRE	RNF2
PIGF	PIGG	PIGO	PIGB	PIGU	PIGK	PIGT	PIGS	PIGA	PIGC	PGF
PIGT	PIGK	PIGS	PIGU	GPAA1	PIGO	PIGA	PIGF	PGAP1	PLAUR	UBC
PLAGL1	CDKN1C	ADCYAP1R 1	IVL	NKX2-5	NPPA	HDAC1	RFX6	TP53	ABCC8	EP300
PLAT	SERPINE1	LRP1	PLG	F2	SERPINF2	SERPINI1	PDGFD	ITGAM	UBA1	HRG
PLOD2	GLT25D2	GLT25D1	WHSC1L 1	SETDB1	WHSC1	EHMT1	SETMAR	SUV39H2	NSD1	PLOD3
PLP2	PLP1	PDCL2	SH2D3C	PPP1CA	PDCL3	SPECC1	CCR1	BCAP31	SH2D3A	TUBB
PLSCR1	SHC1	BACE1	ABL1	SRC	KIAA1149	UBC	PLAC8	EGFR	ITPR1	PRTN3
POFUT1	NOTCH1	NOTCH3	NOTCH4	POFUT2	POGLUT1	NOTCH2	DLL1	JAG1	FUT11	FUT10
POP5	POP4	POP1	RPP25	RPP14	ENSG00000 255154	RPP30	RPP21	RNASEH1	TRIM39R	RPP38
POU2F3	LOR	DCAF11	ITGB3	CSNK2A2	HCFC2	CSNK2A1	LIPG	SPRR2A	PLA1A	ELF3
PPL	KRT8	EVPL	VIM	PPHLN1	AKT1	UBN1	COL17A1	CHL1	KAZN	DSG3
PPP1R3C	NHLRC1	EPM2A	UBC	GYS1	GYS2	GBE1	GYG1	GYG2	UBA52	UBB
PPP2R1B	PPP2R5A	PPP2CA								

Gene List No.2 to be continued…….

PRDM5	EHMT2	CCND1	HDAC1	KDM1A	MAOB	MAOA	KDM1B	EHMT1	PRDM8	GFI1
PRR11	UBC	PRR21	MZT2B							
PRSS2	APP	TCN1	GIF	DEFA5	F2RL1	VEGFA	F2R	SPINK1	ZMYND10	F2RL3
PRSS3	APP	TCN1	GIF	DEFA5	F2RL1	VEGFA	F2R	SPINK1	ZMYND10	F2RL3
PSMB9	PSMB2	PSMB4	PSMB1	PSMA3	PSMA5	PSMA4	PSMA1	PSMA2	PSMA7	HC3
PTEN	AKT1	TP53	PTK2	PDGFRB	SLC9A3R1	UBC	PIK3CA	JUN	PIK3CB	PIK3CD
PTGS2	TP53	PLA2G4A	JUN	PLA2G10	PLA2G2A	MAPK14	CAV1	MAPK1	PLA2G5	PLA2G1B
PTK6	STAP2	KHDRBS1	STAT3	SOCS3	KHDRBS2	PXN	BCAR1	ERBB2	HSP90AA1	TP53
PTPN6	EGFR	SIRPA	LCP2	GAB4	GAB1	SRC	SSTR2	BLNK	HOXA10	MLLT4
PTTG1	CDC20	FZR1	ANAPC11	ESPL1	UBC	CDC27	AURKA	TP53	DECR1	BUB1B
PWP2H	WDR36	RRP9	UTP18	HEATR1	UTP6	WDR75	BMS1	NOP14	NOC4L	NAT10
PYCARD	AIM2	NLRP3	CASP1	MEFV	CASP5	NLRC4	NLRP1	CARD8	PSTPIP1	TP53
RAF1	KRAS	HRAS	RAP1A	YWHAB	MAP2K1	HSP90AA1	YWHAZ	BRAF	YWHAQ	BAD
RALY	MAGOH	HNRNPD	EIF4A3	UBC	HNRNPA0	HNRNPH1	RBMX	HNRNPC	EIF2S2	RNPS1
RAPGEF3	ABCC8	RAP1A	RAP1B	RAP1GAP	AKAP6	RAF1	PDE4A	APBB1IP	GNAS	PDE3B
RARA	NCOR2	NCOA1	PML	NCOR1	NCOA3	NRIP1	NCOA2	MED1	RXRG	KAT2B
RARB	RXRG	NR4A2	RARA	NR2C1	NR0B2	NR2F1	NR4A1	HNF4G	RORC	ESR2
RARG	RXRG	NR2F6	MED1	NR0B2	NRBP1	NR4A1	HNF4G	NR4A2	RORB	VDR
RARRES3	MAVS	UBC	CRABP2	RARG	HRAS	TGM1	RARB	TNFAIP3	PTGDS	HSP90AA1
RASSF1	STK4	STK3	HRAS	CDC20	CNKSR1	MOAP1	MDM2	MLH1	E4F1	KRAS
RBL1	E2F4	TFDP1	TFDP2	CDK2	CDK4	E2F5	E2F1	MYBL2	E2F3	E2F2
RBL2	E2F4	CDK2	E2F5	HDAC1	LIN9	CDK4	LIN54	TFDP1	RBBP4	TFDP2
RBP1	STRA6	LRAT	TTR	RPE65	RBP4	APOB	APOA1	AKT1	APOE	SDC1
RCL1	BMS1	UTP20	BYSL	RRP9	DHX37	NOP14	NOC4L	IMP4	TSR1	WDR12
RGS1	GNAI1	GNAO1	GNAZ	GNAI3	GNAT3	CCR7	CXCR5	CXCL12	CCR1	GRM6
RHCG	ATP6V0A4	CD47	SLC9A3	SLC26A4	OPRM1	GNRH1	AQP1	WRB	AQP2	DNAH8
RNASEN	DGCR8	TP53	XPO5	DDX17	POLR2A	POLR2F	EIF2C1	POLR2C	EIF2C2	POLR2B
RNF114	UBC	XAF1	PARP11	TNFAIP3	UBE2D1	CPLX1	LAGE3	NUDCD2	HSPB1	DSTN
RPL37	RPL14	RPL18A	RPL13	RPL30	RPL4	RPL29	RPL27	RPL32	RPL18	RPL31
RPL39L	RPL31	RPL18A	UBA52	RPL15	RPL29	RPL27	RPL32	RPL38	RPL30	RPL36AL
RPRM	TP53	CCNB1	PGRMC1	TMEM79	HS3ST1	GADD45A	GTSE1	CNNM3	TP53AIP1	SCG2
RPS19	RPS14	RPS13	RPS8	RPS11	RPS16	RPS3A	RPS7	RPS4X	RPS23	RPS5
RPSA	RPS21	RPS16	RPS5	RPS3A	RPS8	RPS3	RPS13	RPS14	FAU	RPS20
RRAD	CAMK2G	PRKACA	PRKCA	NME1	CALM1	CSNK2A1	YWHAZ	MAP3K19	MAP3K3	MAP3K2
RRAS	RAF1	PIK3CA	PIK3CD	PIK3CG	EPHB2	RASA1	SOS1	PLXNB1	MLLT4	NF1
RRM2	RRM1	CDK1	TYMS	PCNA	TFDP1	TFDP2	POLA1	TXN	UBC	CDC6
RTN3	RTN4	KIAA1149	BACE1	UBC	BCL2	COL4A3BP	ZFYVE27	CRELD1	RTN1	REEP5
RUNX1	CBFB	SMAD3	CEBPA	RUNX1T1	KAT6A	SPI1	HDAC1	CTBP1	GATA1	TAL1
RUNX3	CBFB	SMAD3	UBC	SMAD4	EZH2	SP7	CDKN1A	CCND1	NTRK3	ZBTB7B
S100A10	ANXA2	TRPV6	UBC	PLG	AHNAK	DLC1	KCNK3	CFTR	TMEM65	JAK2
S100A7	FABP5	BRCA1	TBP	MYC	MAX	COPS5	S100A12	RNASE7	RASSF9	RASSF7
S100A8	S100A9	TLR4	LY96	MYD88	PTPN11	TIRAP	MIF	SRC	BAX	CDH1
S100A9	S100A8	LY96	TLR4	MYD88	TIRAP	SAP18	TOP1	MIF	MYB	PARP1
S100P	S100PBP	EZR	SNRPF	SUGT1	AGER	IQGAP1	S100A1	CACYBP	IGBP1	CA8
SAA1	LYZ	FPR1	APOA1	TTR	CXCR3	CXCL10	IL8	FPR2	CXCL9	APCS
SCAMP1										
SCD	UBC	SREBF1	PPARA	CYB5A	LEP	INS	PPARG	FABP4	ACACA	FASN
SCEL	SULF1	YIPF3	SYNPO	CYP26A1	TXNIP	ACTN4	NPHS1	POU4F2	RAD52	LHX9
SDC1	FGF2	HPSE	HGF	MMP1	GPC1	SDCBP	HSPG2	CASK	SDC2	SDC3
SDHA	SDHB	SDHC	SDHD	SDHAF2	SUCLA2	SUCLG2	UQCRFS1	SUCLG1	CYC1	FH
SEMG2	WFDC12	KLK3	KLK2	SPINT4	PRKAR2A	EPPIN	TGM1	SLPI	WFDC5	SERPINA5
SEPP1	TNF	LRP8	LRP2	KRT16	SEPT4	PVRL2	GOPC	GPX2	SMCP	GPX3
SERPINA3	DNAJC1	IL6	ELANE	KLK3	CTSG	KLK2	KLK4	CLEC1B	CTRL	CMA1
SERPINB5	TP53	ATF2	FBXO32	KHDRBS3	DNMT3A	CHUK	IL6	PLAU	IRF6	TP63
SFN	TP53	CDC25B	TSC2	UBC	RAF1	BAD	MAPK1	MST1R	AKT1	CDK1
SFRP1	WNT4	WNT5A	WNT3A	WNT1	WNT2	WNT8B	WNT8A	WNT7B	WNT10B	WNT7A
SFRP2	WNT3A	WNT4	WNT1	WNT3	WNT5A	PAX6	WNT7B	SFRP1	WNT7A	WNT11
SFRP4	WNT7A	WNT8B	WNT8A	WNT4	WNT5A	WNT3A	WNT10B	WNT7B	WNT1	WNT11
SFRP5	WNT8B	WNT5A	WNT11	WNT8A	WNT10B	WNT3A	WNT3	LHX5	WNT2B	WNT4
SIPA1	RRP1B	RAP1A	RAPGEF3	RASGRP1	RASGRP2	MLLT4	RAP1B	RAP1GAP2	RAPGEF4	YWHAB

Gene List No.2 to be continued…….

SIX1	EYA2	EYA1	DACH1	EYA4	MYOD1	MYOG	DACH2	PAX2	TLE1	CCNA1
SKP2	SKP1	CDKN1B	CUL1	CDK2	CDKN1A	CKS1B	RBX1	UBC	FZR1	MYC
SLC3A2	SLC7A5	SLC7A11	SLC7A8	UBC	SLC7A6	SLC7A7	SLC7A10	FAM57A	MYH14	ICAM1
SLC7A6OS	POLR2L	POLR2E	POLR2C	POLR2B	POLR2G	POLR2D	POLR2I	POLR2F	POLR2H	GTF2F2
SLURP1	LYPD6	LYNX1	GML	CHRFAM7A	CHRNA7	CAPN3	ITGA2	TTN	CHRNA4	ACADVL
SMAD2	FOXH1	TGFBR1	SKIL	ZFYVE9	SKI	SMAD4	UBC	CREBBP	ACVR1B	EP300
SMAD4	SKI	SMAD2	SMAD3	SMAD1	TRIM33	UBC	EP300	PIAS4	SKIL	BMP2
SOCS1	UCP1	DFFB	NRIP1	SLC7A1	SREBF1	XDH	PLIN1	CIDEC	CIDEB	PRKAB1
SOX1	NOG	CTNNB1	FOXA2	PAX6	NANOG	FOXA1	NES	NEUROG2	STAT3	NAT9
SP100	SUMO1	PML	CBX5	SUMO2	ETS1	UBE2I	TP53	CREBBP	OASL	IRF7
SPA17	ROPN1	AKAP3	ROPN1L	AKAP7	GAS2L2	ZPBP	RASL10A	GAS2L1	GAS2	PRSS50
SPARC	PLG	TGFB1	VEGFA	FN1	COL1A1	COL3A1	ALB	MMRN1	TIMP1	VWF
SPINK5	SFTPD	PPP2CA	CHGA	SST	ST14	KLK7	LAP3	KLK5	MEN1	LPA
SPP1	CD44	MMP7	ITGAV	ITGB1	ACP5	RUNX2	ITGA9	THBS1	BMP2	TP53
SPRR1A	IVL	LOR	S100A10	TCHH	S100A6	SPRR2A	ACTA1	SPRR2B	FLG	SPRR4
SPRR1B	FOSL1	JUN	MAPK7	SP1	MAP2K5	IVL	GLIS1	KLK7	LOR	KLK6
SPRR2A	SPRR2B	SND1	POU2F3	ELF3	SPRR1A	ITK	CTTN	HCK	ABL1	SRC
SPRR2B	SPRR2A	SPRR1A	SPRR3	SPRR1B	KIAA1614	SPRR4				
SPRR3	IVL	LOR	KRT16	SRSF2	S100A10	TCHH	EVPL	SPRR2A	FLG	SPRR2B
ST14	SPINT1	ENSG00000228325	ASIC2	TEK	HSPA4	F2RL1	ST13	MED16	CDCP1	LAP3
STARD10	STARD8	STARD6	STARD9	TMCO7	PI16	GRB2	STARD13	STAR	COL4A3BP	STARD3
STAT1	JAK2	PIAS1	CREBBP	JAK1	IFNGR1	STAT2	EGFR	EP300	TYK2	IRF1
STAT3	EGFR	SRC	JAK2	JAK1	EP300	JAK3	CCND1	MTOR	STAT1	CREBBP
STAU1	UPF1	CASC3	RPS6	PABPC1	DHX9	UBC	NCL	HNRNPU	PPP1CA	TUBA1A
STK11	CAB39	STRADA	STRADB	CDC37	CAB39L	ETV4	STK11IP	PRKAA1	HSP90AA1	PTEN
STK31	TDRKH	TDRD9	TDRD5	RNF17	PIWIL3	TDRD15	TDRD7	PIWIL1	FAM216A	PIWIL4
STX16	STX6	VTI1A	VAMP4	STX10	STX8	YKT6	VPS45	VAMP5	STX5	VAMP3
SULF1	ARSH	SLCO5A1	SHH	EGFR	BMP7	HNF1B	ARSG	SCEL	ARSF	ARSE
SULT2B1	CYP7A1	CYP3A7	CYP17A1	CYP3A5	CYP1A1	CYP1A2	CYP2E1	CYP3A4	CYP7B1	CYP11A1
SYCP2	SYCP3	SYCP1	TEX11	SYCE2	SMC1B	STAG3	TEX12	SYCE1	RAD21	STAG2
SYK	CBL	VAV1	BLNK	GRB2	TYROBP	LCP2	LAT	LCK	CBLB	HCLS1
TARS	UBC	SARS	HARS	FARS2	YARS	LARS	NARS	IARS	FARSB	KARS
TBC1D1	RAB10	RAB8A	RAB13	SLC2A4	RAB14	PRKAA2	PRKAB1	PRKAG3	AKT1	PRKAG1
TBX3	BMP4	SHH	GATA6	HDAC1	LEF1	IHH	DHH	MSX2	GATA3	CDKN2A
TCEA2	SUPT5H	SUPT4H1	RTF1	WHSC2	CDK9	LEO1	CTR9	CDC73	POLR2A	POLR2I
TERT	HSP90AA1	XRCC5	XRCC6	AKT1	SMG6	MTOR	PTGES3	MYC	PINX1	TEP1
TFF3	EGFR	ERBB2	STAT6	CTNNB1	TFF2	TFF1	HIF1A	ARNT	GHR	RELA
TFPI2	MMP9	MMP2	IL1B	TMPRSS4	PSAP	GCG	PLG	VCAN	F11	KLKB1
TFRC	TF	HFE	OPTN	MYC	HIF1A	UBC	B2M	SLC40A1	FTH1	STEAP3
TGFA	EGFR	ERBB2	ERBB3	TP53	ADAM17	ONECUT1	IDE	ERBB4	IGF1R	MET
TGFB1	TGFBR1	TGFBR2	DCN	SMAD3	SMAD2	SMAD7	ACVRL1	LTBP1	APP	LTBP4
TGFB2	TGFBR2	TGFBR1	TGFBR3	APP	ATF2	TGFB1	DCN	SMAD2	SMAD3	FMOD
TGFBR1	SMAD2	SMAD7	TGFB1	FKBP1A	TGFBR2	SMAD3	UBC	TGFB3	SMAD6	SMAD4
TGM3	CALML5	WWOX	TGFB1	SBSN	LOR	LGMN	ABCA12	USP32	USP53	CD79A
THBD	F2	PROC	PF4	F5	PROS1	F10	JUN	FOSL1	PF4V1	CLEC4C
THBS1	CD47	MMP9	CD36	MMP2	LRP1	ITGB3	PLG	COL18A1	FGF2	VEGFA
THBS2	MMP2	THBS3	MMP9	CD36	SPP1	ITGB1	PDGFB	FURIN	THBS1	ITGA6
TIMP1	MMP9	MMP3	MMP2	MMP1	MMP14	MMP15	TGFB1	STAT3	MMP10	SERPINE1
TIMP2	MMP2	MMP14	MMP16	MMP9	RECK	MMP1	MMP10	FURIN	MMP13	MMP8
TIMP3	MMP2	ADAM17	TP53	MMP9	EFEMP1	MMP14	MMP3	MMP23B	GSTP1	ADAMTS4
TKTL1	H6PD	TALDO1	G6PD	GMPS	PGK2	TEX28	PGK1	PGD	RPE	ACLY
TLR2	MYD88	HMGB1	TIRAP	CD36	NFKB1	TOLLIP	CD14	LY96	IRAK4	TRAF6
TLR4	LY96	MYD88	TIRAP	TRAF6	TICAM1	TICAM2	HMGB1	IRAK4	TRAF3	IRAK2
TLR8	MYD88	UNC93B1	TLR3	BTK	CTSK	CTSS	TLR9	CTSB	TLR7	LGMN
TMC6	SYNGR2	TK1	CANX	SLC30A9	UBC	WLS	SEC14L2	SEC14L4	SEC14L3	EVL
TMCO6	DNAJC18	WDR55	DNAJB14	UBQLN1	PCDHB1					
TMEM132A	UBC	HSPA5	NECAP1	SSR3	GSK3B	NECAP2	HSD17B2	NUDT3	RNF157	MCTS1
TMPRSS11	MAP2	MST1	CNPY1	SPINK6	TSPO2	TMEM4	THTPA	LRFN1	ERCC8	ZNF597

D										
TMPRSS11E	KRT16	SERPINE1	PVRL4	SPINT1	STEAP4	ANO1	DEFB107A	C14orf80	ANO2	CTAGE8
TNF	IFNG	ICAM1	ITGAM	FOXP3	MMP9	IKBKG	RELB	TNFRSF1A	MYC	CD83
TNFRSF10C	TNFSF10	TP53	FADD	BFAR	CFLAR	ZC3H12A	RAP1A	BMX	DCXR	CNTN3
TOB1	CNOT7	CPEB3	CNOT6L	CNOT1	RQCD1	CNOT10	TNKS1BP1	CNOT2	SMAD4	PABPC1
TOP2A	TOP1	TOP2B	SUMO2	SUMO1	UBC	PCNA	BLM	TOP1MT	RECQL	ESPL1
TP53	MDM2	CDKN1A	ATM	BRCA1	KAT2B	EP300	SIRT1	CDKN2A	MAPK8	CREBBP
TP73	MDM2	ABL1	YAP1	UBC	EP300	KAT2B	TP53BP2	TP53	MYC	E2F1
TPX2	AURKA	KIF15	KPNB1	PLK1	ZNF598	NUMA1	KPNA2	RAN	HMMR	PSAP
TRA2A	MAGOH	SRSF10	PRPF40A	SAP18	39326	U2SURP	IK	SRSF5	DHX15	HRH2
TREM2	TYROBP	PLXNA1	SYK	SEMA6D	PLCG2	LCP2	VAV2	VAV3	FYN	SOS1
TRIM29	HINT1	TP53	HDAC11	TRAT1	TRIM23	THY1	TRIM11	KAT5	GSK3B	PRKCA
TRIP13	RAD51C	USP32	USP10	HORMAD2	SPO11	HORMAD1	UBC	RAD1	TSEN34	TSEN2
TSPAN3	ITGB1	IGSF8	SCARB1	MMP2	MMP14	CD36	HMBS	MME	VAMP7	MMP9
TUBB2	TUBA4A	TUBA1A	TUBA1B	TUBB4B	UBC	TBCD	TBCA	TUBA1C	ACTB	SIRT2
TWIST1	TCF3	TP53	SNAI1	MYOD1	CDH2	CDH1	MSX2	CDKN2A	RELA	TCF4
TYK2	IFNAR1	STAT1	STAT5A	STAT5B	STAT2	STAT3	STAT6	SOCS1	PTPN1	SOCS3
TYMS	DHFR	SHMT1	PCNA	RRM2	DUT	DHFRL1	MTHFD1	DTYMK	DCTD	MTHFR
UFD1L	NPLOC4	VCP	UBL4A	UBC	UBE4B	YOD1	PLAA	SPATA5	USP13	UBE2N
UPK1A	UPK2	UPK3A	UPK3B	CD151	CD53	CD37	CD63	TMEM147		
VEGFA	KDR	FLT1	NRP1	HIF1A	PGF	VWF	TGFB1	SRC	IGF1	MMP9
VEGFB	FLT1	VEGFA	NRP1	FLT4	KDR	FIGF	VEGFC	IGF1	VWF	PGF
VEGFC	FLT4	KDR	NRP2	FLT1	IGF1R	IGF2	VWF	VEGFA	ALB	IGF1
VHL	HIF1A	EPAS1	CUL2	TCEB1	TCEB2	RBX1	UBC	HIF3A	EGLN1	FN1
VPS35	VPS29	VPS26A	VPS26B	SNX2	SORL1	MOGAT3	SNX1	SNX3	VPS36	VPS25
VSIG4	C3	C1R	CFB	BMP1						
VWF	GP1BA	ADAMTS13	VEGFA	ITGA2B	SERPINE1	ALB	F5	GP9	F9	ITGB3
WT1	TP53	PAX2	SOX9	NR0B1	WTAP	PAWR	FOXD1	NPHS1	PAX8	WNT4
XRCC1	PARP1	LIG3	PNKP	APTX	POLB	PARP2	APLF	OGG1	PCNA	PARG
YIF1A	YIPF5	VAPB	YIPF7	RABAC1	RAB1A	UBC	YIPF6	YIPF4	RAB1B	VAPA
ZBED2	AXIN1	AXIN2	SWI5	ZBED6	C7orf29	ZBED5	SYNRG	MTHFSD	THUMPD1	ZBTB41
ZNF217	HDAC2	CTBP1	KDM1A	RCOR1	CTBP2	HDAC1	SEPSECS	HDAC9	RBBP7	HDAC3

End of Gene List No.2

After finding the target genes with the help of STRING database, we go for validation of target genes through micro array data and check whether the genes we found, expressed in cervical cancer cells or not.

6.1 GEO microarray data:

- GEO Accession Id: **GSE27469**

Number of Samples Taken = 82 patients with cervical cancer

In total, 82 patients with cervical cancer, stage 1b bulky through 4a, were included. Pretreatment biopsies were taken from the tumors and analyzed with Illumina gene expression BeadArrays (Human WG-6 v3).

Experiment Type: Expression profiling by array

82 Samples of GSE27469 (Drivers of gene expression in cervical cancer)

Treatment Protocol

The samples were immediately snap-frozen in liquid nitrogen and stored at -80oC until used for analysis. Biopsies with more than 50% tumor cells in HE stained sections from the central part of the specimen were used further. 1-4 tumor samples were collected from each patient at the time of diagnosis.

RNA was extracted with Trizol reagent, followed by double precipitation with isopropanol and final precipitation with 5M lithium chloride. RNA from different biopsies of the same tumor was pooled. Quality control was performed with the Agilent Bioanalyzer.

The samples were hybridized to the arrays at 58oC overnight, using the standard Illumina hybridization protocol.

GSM678702

Cohort (gse36562 study): DCE-MRI cohort
hypoxia score: High
tissue: cervical tumor
figo stage: 2B
histology: Squamous cell carcinoma
cohort: Validation cohort
cohort (gse38964 study): Integrative cohort
3p status: Loss

GSM678703

Cohort (gse36562 study): Validation cohort
hypoxia score: High
tissue: cervical tumor
figo stage: 3B
histology: Squamous cell carcinoma
cohort: Basic cohort

lymph node status (gse38433 study): 1
cohort (gse38964 study): Integrative cohort
3p status: Loss

GSM678704

Cohort (gse36562 study): DCE-MRI cohort
hypoxia score: Low
tissue: cervical tumor
figo stage: 3B
histology: Squamous cell carcinoma
cohort: Basic cohort**lymph node status (gse38433 study):** 1
cohort (gse38964 study): Integrative cohort
3p status: Loss

GSM678705

Cohort (gse36562 study): Validation cohort
hypoxia score: High
tissue: cervical tumor
figo stage: 3B
histology: Squamous cell carcinoma
cohort: Basic cohort

lymph node status (gse38433 study): 0
cohort (gse38964 study): Integrative cohort
3p status: Loss

GSM678706

Cohort (gse36562 study): DCE-MRI cohort
hypoxia score: Low
tissue: cervical tumor
figo stage: 3B
histology: Squamous cell carcinoma
cohort: Basic cohort.

lymph node status (gse38433 study): 1
cohort (gse38964 study): Integrative cohort
3p status: Loss

The total number of sample is 82, which had been used during micro array experiment and here different conditions which have been applied during micro array experiment on 5 cell samples are shown.

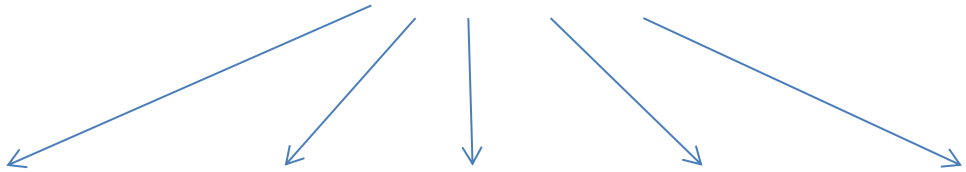
- GEO Accession Id: **GSE30758**

Number of Samples Taken = 152

Out of the 152 women, 75 developed a cervical intraepithelial neoplasia of grade 2 or higher (CIN2+) after 3 years of sample collection. The rest of women (77) remained disease free (CIN2-).

Overall Design: Bisulphite converted DNA from the 152 samples were hybridized to the Illumina Infinium 27k Human Methylation Beadchip v1.2.

152 Samples of GSE30758 (genomic DNA from normal cells of the uterine cervix)



GSM763063

Sample type: bisulphite converted genomic DNA from a prospective CIN2- **outcome:** CIN2+(=1)/CIN2-(=0) 1
Age: 25
Hpv status: 0
pcr: Round2Cyt 1
sentrix id: 1
sentrix position: A
bisulphite 1: BS conversion C1 Grn 2431
bisulphite 2: BS conversion C2 Grn 2403

GSM763064

Sample type: bisulphite converted genomic DNA from a prospective CIN2- **outcome:** CIN2+(=1)/CIN2-(=0) 1
Age: 19
Hpv status: 0
pcr: Round2Cyt 1
sentrix id: 1
sentrix position: B
bisulphite 1: BS conversion C1 Grn 2690
bisulphite 2: BS conversion C2 Grn 2111

GSM763065

Sample type: bisulphite converted genomic DNA from a prospective CIN2- **outcome:** CIN2+(=1)/CIN2-(=0) 1
Age: 19
Hpv status: 0
pcr: Round2Cyt 1
sentrix id: 1
sentrix position: C
bisulphite 1: BS conversion C1 Grn 2414
bisulphite 2: BS conversion C2 Grn 1939

GSM763066

Sample type: bisulphite converted genomic DNA from a prospective CIN2- **outcome:** CIN2+(=1)/CIN2-(=0) 1
Age: 40
Hpv status: 0
pcr: Round2Cyt 1
sentrix id: 1
sentrix position: D
bisulphite 1: BS conversion C1 Grn 2176
bisulphite 2: BS conversion C2 Grn 1612

GSM763067

Sample type: bisulphite converted genomic DNA from a prospective CIN2- **outcome:** CIN2+(=1)/CIN2-(=0) 1
Age: 25
Hpv status: 1
pcr: Round2Cyt 1
sentrix id: 1
sentrix position: E
bisulphite 1: BS conversion C1 Grn 2036
bisulphite 2: BS conversion C2 Grn1826

The total number of normal cell samples is 182, which had been taken during micro array experiments. Here, randomly 5 samples were selected from 182 samples and shown their different experimental conditions.

Downloaded the gene expression microarray data of normal cervix cell and cervical cancer cell having GEO id GSE30758 and GSE27469 respectively and then find how many genes (present in Gene List No.1 & 2) are common in these two microarray data. After comparison, we found that 3000 genes are common in normal cervix gene micro array data and the list no.2 (prepared with the help of STRING database) and 2998 genes are common in Gene List No.2 and cervical cancer gene microarray data. From this comparison, we conclude that the genes present in Gene List No.1 &2 are associated with cervical cancer progression because the genes are expressed in cervical cancer cells with different expression profiles as compared to normal cervix cells. We also compared the expression values (log FC) of common genes present in both micro array data and found that there is a change in expression value of each common gene in normal and cancer cell conditions.

The list of common genes and comparison of their fold change (log FC) is given below:

				GEO ID=GSE30758				GEO ID=GSE27469				
				micro array data of normal cervix cell				cervical cancer micro array data				
cc network gene		common genes		ID	genes	logFC		ID	Gene.symbol	logFC		common genes
UBC		UBC		cg17792192	CAMSAP1	-0.119		ILMN_1736007	A1BG	-0.01118049		UBC
TP53		TP53		cg00113020	LILRB4	-0.106		ILMN_2055271	A1BG	-0.00032064		TP53
EGFR		EGFR		cg08418978	CLDN10	-0.114		ILMN_1883785	A1CF	-0.041859		EGFR
MYC		MYC		cg20383064	BFSP2	-0.123		ILMN_1893846	A1CF	0.01724954		MYC
CCND1		CCND1		cg14502651	PRH2	-0.122		ILMN_2383229	A1CF	0.02119657		CCND1
GRB2		GRB2		cg03552688	BPESC1	-0.129		ILMN_1806310	A1CF	-0.01597895		GRB2
AKT1		AKT1		cg02623400	SERPIND1	-0.118		ILMN_1779670	A1CF	-0.00405414		AKT1
SRC		SRC		cg07873128	OSBPL5	-0.0508		ILMN_1745607	A2M	0.12964894		SRC
CDK2		CDK2		cg18389810	C14orf8	-0.124		ILMN_2136495	A2ML1	0.13227821		CDK2
JUN		JUN		cg03622431	DNASE2B	-0.1		ILMN_2295659	A3GALT2	-0.02501706		JUN
CTNNB1		CTNNB1		cg03455024	SLC12A3	-0.102		ILMN_1735045	A4GALT	-0.02728861		CTNNB1
VEGFA		TGFB1		cg17474651	MAGEC3	-0.11		ILMN_1680754	A4GNT	0.03968825		VEGFA
TGFB1		PCNA		cg14272706	ALS2CR12	-0.107		ILMN_1851496	AA06	-0.00603029		TGFB1
PCNA		MMP9		cg17108383	PCDHGC5	-0.093		ILMN_1755321	AAAS	0.01772988		PCNA
MMP9		HSP90AA1		cg04592706	ABCA4	-0.108		ILMN_1698554	AACS	-0.05684599		MMP9
HSP90AA1		NTN4		cg03752087	CASP14	-0.108		ILMN_1760414	AADAC	-0.11699257		HSP90AA1
NTN4		CDH1		cg00037763	SETBP1	-0.1		ILMN_1752884	AADACL2	-0.00747106		NTN4
CDH1		STAT3		cg23773532	GPRC5D	-0.064		ILMN_1660703	AADACL3	-0.03187472		CDH1
STAT3		IL6		cg24198678	DNASE1	-0.106		ILMN_1668851	AADACL4	0.02112043		STAT3
IL6		ERBB2		cg26787239	IL4	-0.0603		ILMN_1726986	AADAT	0.29761671		IL6
ERBB2		MMP2		cg23749046	GPR61	-0.109		ILMN_2270015	AADAT	0.08684805		ERBB2
MMP2		FGF2		cg02005755	SLC13A4	-0.112		ILMN_1809959	AADAT	-0.03506827		MMP2

Genes that are associated with cervical cancer (present in Gene List No.2)

Gene Expression microarray data of normal cervix cell **(GEO id: GSE 27469)**

Gene Expression microarray data of cervical cancer cell **(GEO id: GSE 30758)**

Common genes found in Gene List No. 2 and micro array data

Table 1. Showing common genes present in different columns

Table 2. Represents difference between Fold Change (logFC) of genes present in normal and cancer cell condition

		log FC	log FC
2	common genes	normal cell	cancer cell
3	UBC	0.0263	0.05400587
4	TP53	0.0171	-0.08339098
5	EGFR	0.00184	-0.10306953
6	MYC	0.00553	-0.32775617
7	CCND1	0.0276	-0.25182692
8	GRB2	0.00209	-0.03298683
9	AKT1	-0.00518	-0.13560081
10	SRC	-0.114	0.07349216
11	CDK2	0.0186	-0.08665518
12	JUN	-0.00189	0.0745052
13	CTNNB1	0.0101	-0.0451723
14	VEGFA	#N/A	-0.05183348
15	TGFB1	-0.0116	-0.0202905
16	PCNA	0.0293	-0.16351064
17	MMP9	-0.0364	0.07143603
18	HSP90AA1	0.00599	-0.06423078
19	NTN4	-0.0611	-0.03597337
20	CDH1	-0.0994	-0.02005108
21	STAT3	0.0309	-0.15269625
22	IL6	0.00551	0.31800967
23	ERBB2	0.0225	-0.04825521
24	MMP2	-0.0395	0.02102479
25	FGF2	0.0387	-0.02556014

After the validation process, Cytoscape tool was used for network construction. The validation process concluded that when a normal cervix cell changed into a cervical cancer cell, the fold change (log FC) of 3000 genes are changed accordingly.

6.2 Cytoscape:

- Prepared a table of two columns in excel: the first column contained Source genes(genes in Gene List No.1) and second column filled with target genes.(searched by online tool STRING)
- Import the excel table in Cytoscape tool and constructed the network genes that are associated with cervical cancer.

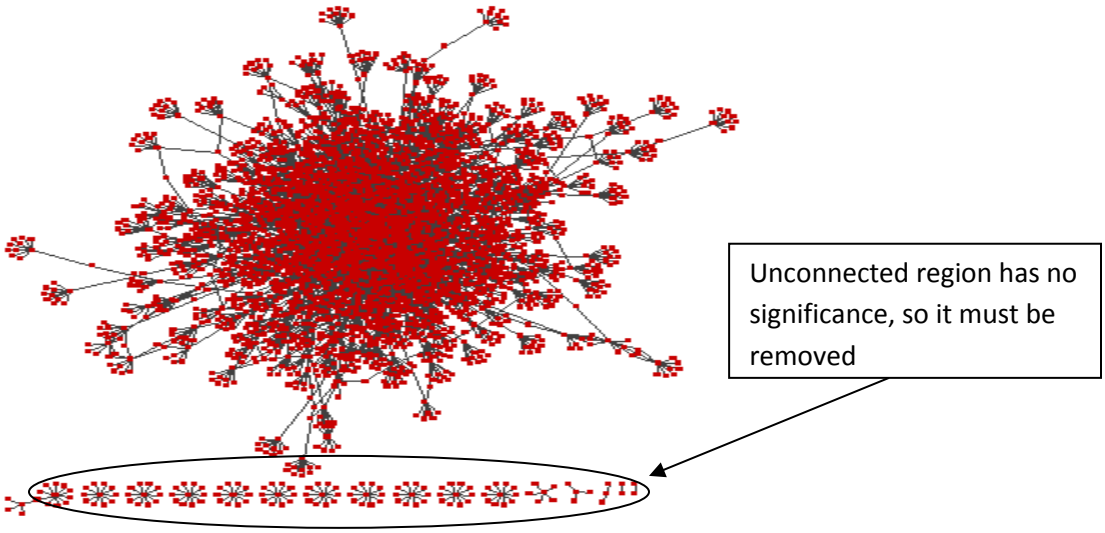
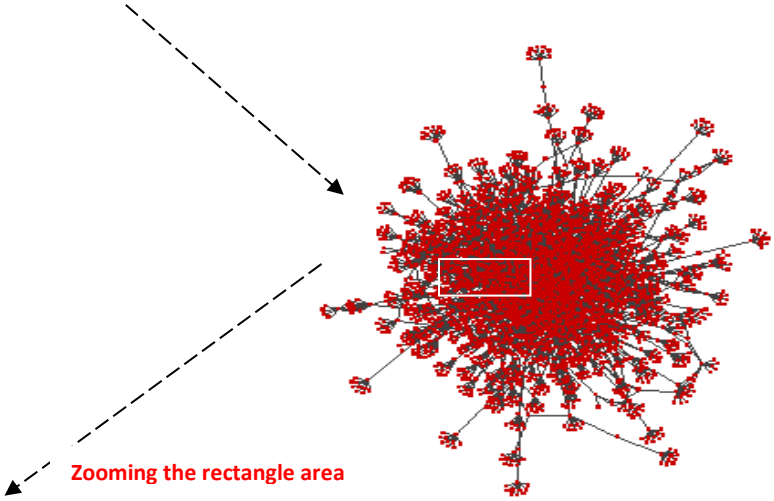


Figure 12. Cervical cancer gene network that having 3241 nodes and 5211 edges.

- After removing non-interacted regions, we removed all self-loops and duplicated edges, in order to make gene network more significant and relevant. Now the total number of nodes is 3099 and 4932 edges.



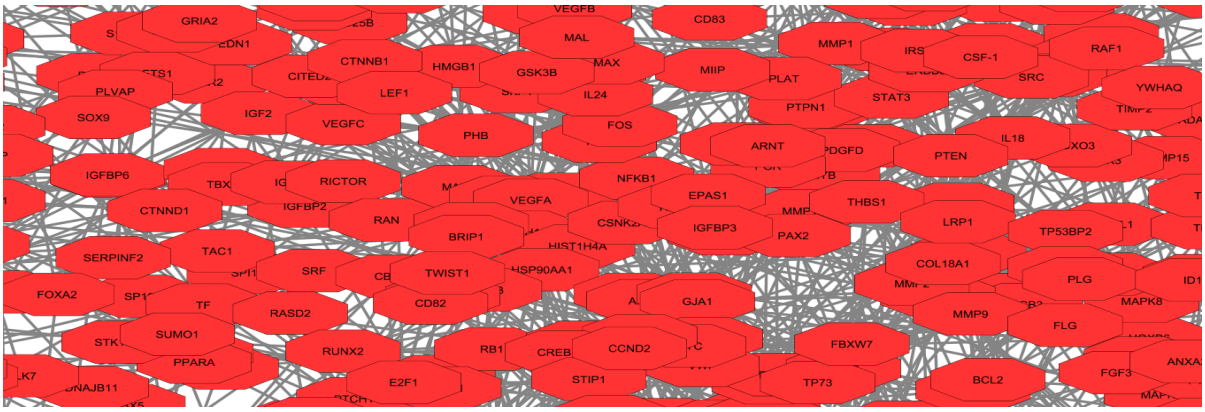


Figure 13: nodes (vertices) of the gene network, labeled with gene name.

Once the network was obtained, NetworkAnalyzer (plug-in of Cytoscape) used for the topological studies of the network.

6.2 Network Analyzer:

- Analyzed the topological properties of gene network (degree distribution, average clustering coefficient, closeness centrality, betweenness centrality, neighborhood connectivity) by using NetworkAnalyzer plug-in of Cytoscape.
- All biological networks are scale-free, the characteristic of scale free network is that its topological properties should follow a power law. According to network theory, if a network is evolved in nature, then its topological properties must follow a power law and network must be scaled-free.

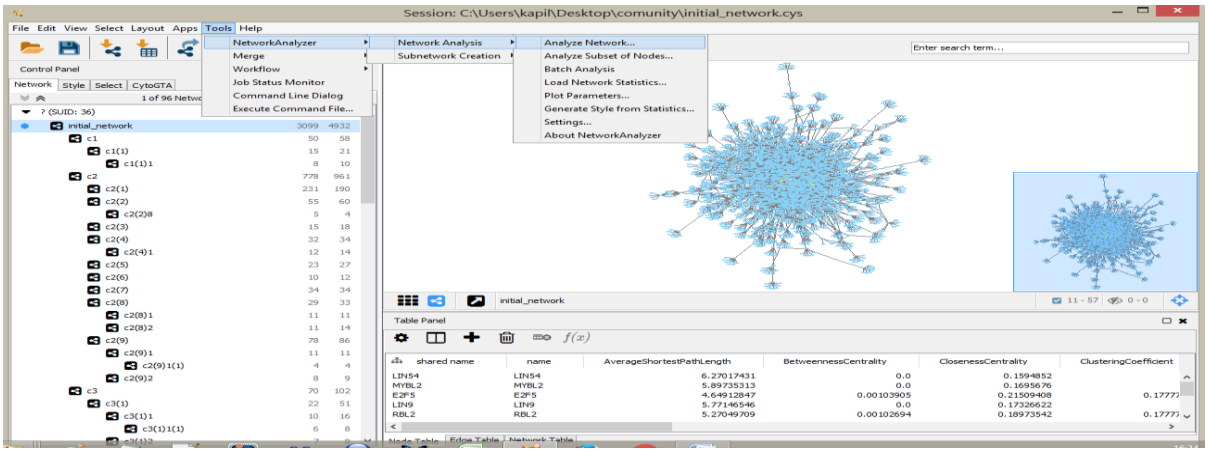


Figure 14. To run NetworkAnalyzer, select **Tools → NetworkAnalyzer → Network Analysis → Analyze Network**.

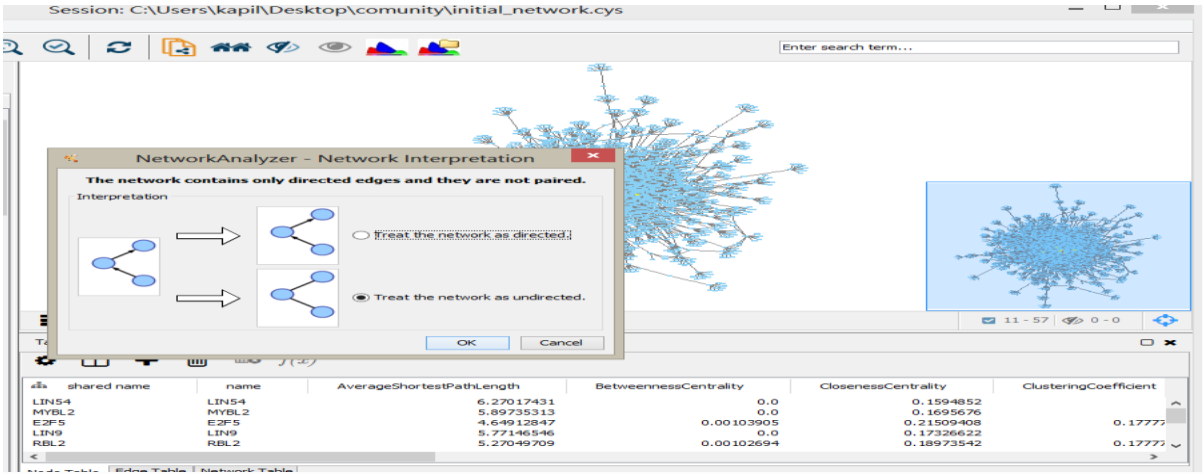


Figure 15. For scale free network, choose option" Treat the network as undirected" which means in the case of scale free network, information can be transferred between any two nodes in both directions.

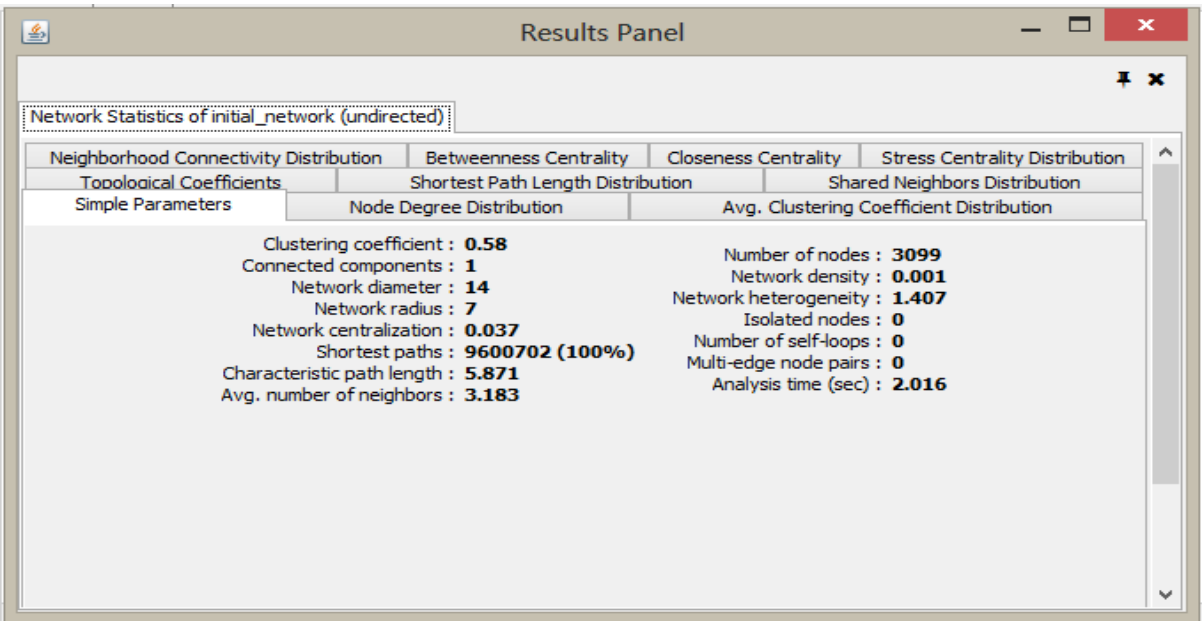


Figure 16. The result obtained by NetworkAnalyzer that defines simple parameters of gene network like clustering coefficient, network diameter, radius, shortest path, network density etc.

6.3 Power Law:

Power law also is known as scaling law is a functional relationship between two quantities, such that change in one quantity results in a proportional change in second quantities and does not depend on the initial size of both quantities [51].

In other words, one quantity varies as a power of the second.

A power law distribution can be represented by the mathematical function $\mathbf{Y = k X^{\alpha}}$, where X and Y are variables of our interest, k is a constant and $\boldsymbol{\alpha}$ is the law's exponent.

6.4 Study of Topological properties of cervical cancer gene network:

If a network is evolved in nature as well as scale-free, then some topological properties should follow a power law [58]. The topological properties of the network are given below:

- **Degree Distribution:**
  The degree distribution of a scale free network is that where the probability of a node having a given degree has a scale-invariant decay as degree grows [58]. It means it follows a power-law of the form
  P (k) ~ k^ (-ɣ)

Where ɣ > 1 is a constant and k=1, 2, 3,.…N.

The degree distribution of a particular node can be calculated by the following formula

$$P(k) = \frac{n_k}{N}$$

Here, *p(k)* is probability degree distribution of a node having degree K

N_k is a number of nodes having degree k

N represents total number of nodes in a network

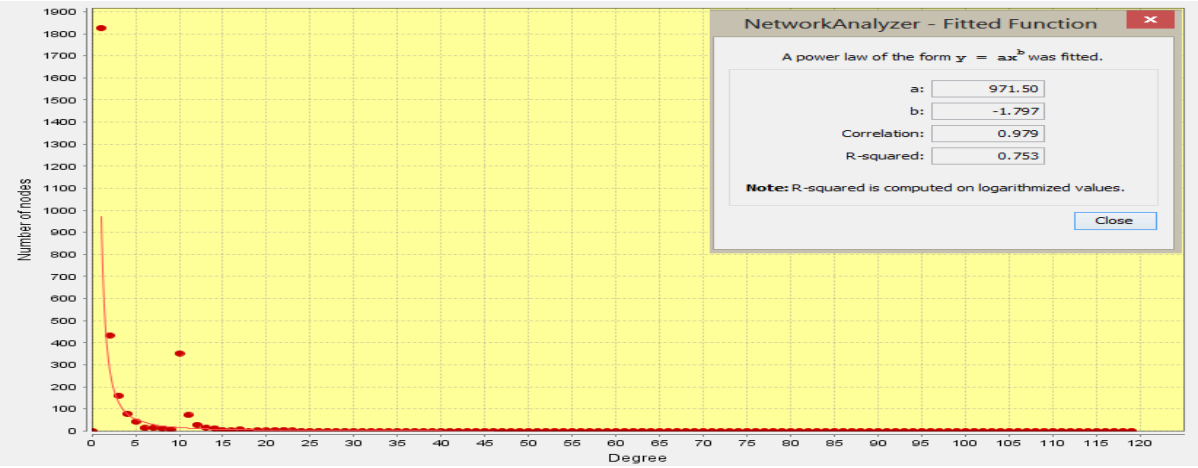


Figure 17: The node degree distribution of cervical cancer associated gene network follow the power law and the value of ɣ is 1.8.

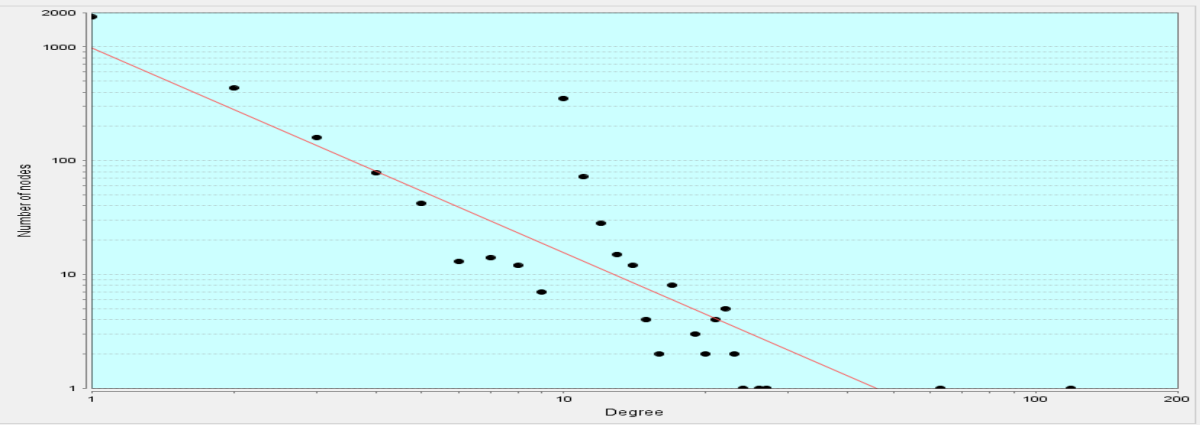


Figure 18: Log-Log plot of degree distribution p(K) and degree (k), which is a straight line that shows it follows a power law. The value of ɣ in this graph is 1.78. the graph is plotted with the help of NetworkAnalyzer, a plug-in of Cytoscape.

- **Betweenness centrality:**

  Betweenness centrality is calculated for networks those do not contain multiple edges. which do This property measures the extent to which a node presents on shortest paths between other nodes, in a network, that means Nodes with high betweenness may have great influence on network by virtue of their control over information passing between other nodes, and removal of these nodes will disrupt communications between other nodes [52].The value of betweenness centrality for each node lies between 0 and 1.

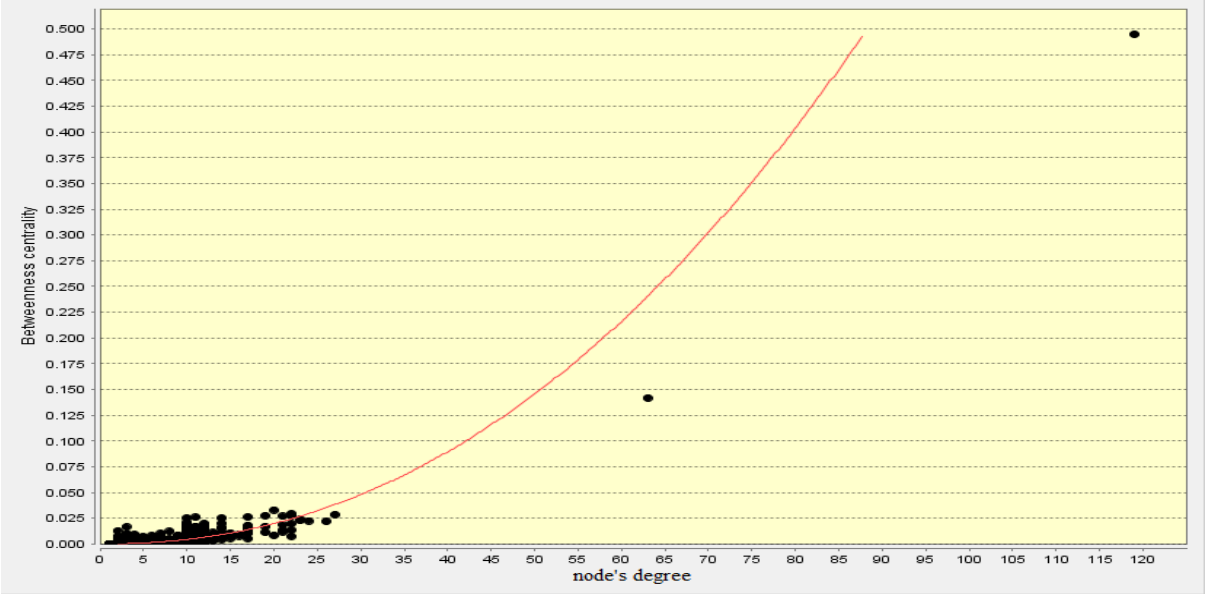


Figure 19: Graph plotted between betweenness centrality of each node and their respective degree. There are only a few nodes (hardly 5-10) that have high betweenness centrality, that confirmed their influence or dominance on the network.

- **Closeness centrality:**

  Closeness centrality of a node is defined as the reciprocal of the sum of the length of the shortest paths between the node and all other nodes in the network [53][55].

  In simple words, we can say that closeness is reciprocal of the farness. The closeness centrality of a node can be any number between 0 and 1. (0 for isolated node and 1 for highly centered node in a network)

$$C(x) = \frac{1}{\sum_y d(y,x)}.$$

Here d(y,x) is the distance between node x and y. in order to generate its normalized from, multiplied above formula by N-1, where N is the total number of nodes present in a network. For a large network, N-1 reduced to N.

$$C(x) = \frac{N}{\sum_y d(y,x)}.$$

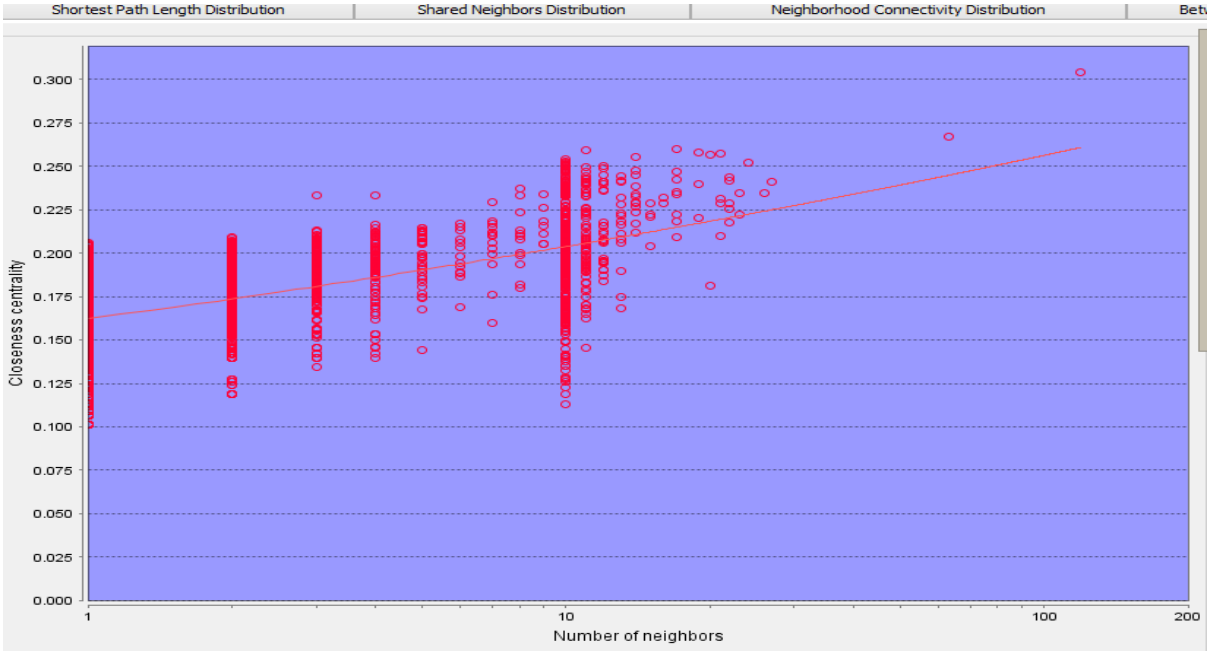


Figure 20: Graph of closeness centrality vs a number of neighbors. There are few nodes that have closeness centrality value greater than average, shown in the graph, hence confirmed their dominance in the network.

- **Average clustering coefficient:**

  Clustering coefficient:

  It is a measure of the ability of nodes, to exist in a cluster together (according to graph theory) in a network or graph.

  The mathematical representation of clustering coefficient **C_n**, for undirected networks, is given below

  $$C_n = 2e_n /(k_n(k_n-1))$$

  k_n is a number of neighbors (degree) of node n

  e_n is a number of connected pairs between all neighbors of node n

For directed networks, clustering coefficient is defined as $C_n = e_n /(k_n(k_n-1))$

The clustering coefficient of each node lies between 0 and 1. The average clustering coefficient is the average of clustering coefficient for all nodes(that have degree >1) present in a network [55].

We can also define clustering coefficient of a node n by the following formula:

$$C = \frac{3 \times \text{number of triangles}}{\text{number of connected triplets of vertices}} = \frac{\text{number of closed triplets}}{\text{number of connected triplets of vertices}}.$$

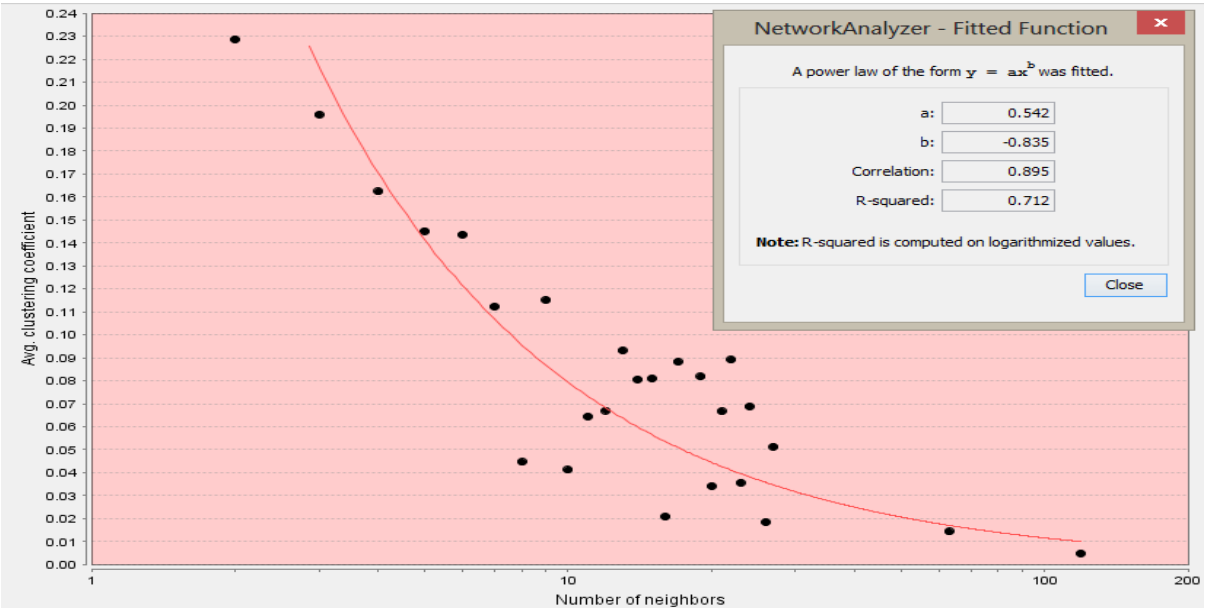


Figure 21: Graph-plot of Avg. clustering coefficient and a number of neighbors (obtained by NetworkAnalyzer, a plug-in of Cytoscape). The curved line represents that graph follows power law.

- **Topological coefficient:**

The topological coefficient $\mathbf{T_n}$ gives the information about how many neighbors are shared by a node n with other nodes of the network [57][58].

NetworkAnalyzer(plug-in of Cytoscape) calculates the topological coefficient for all nodes that have neighbors greater than 1. The value of $\mathbf{T_n}$ is zero for those nodes that have one or no neighbors.

$$\mathbf{T_n = avg(j(n,m))}$$

Where j(n,m) is defined as the total number of other nodes(m) that share at least one node with n.

The chart of the topological coefficients can be used to estimate the tendency of the nodes in the network to have shared neighbors.

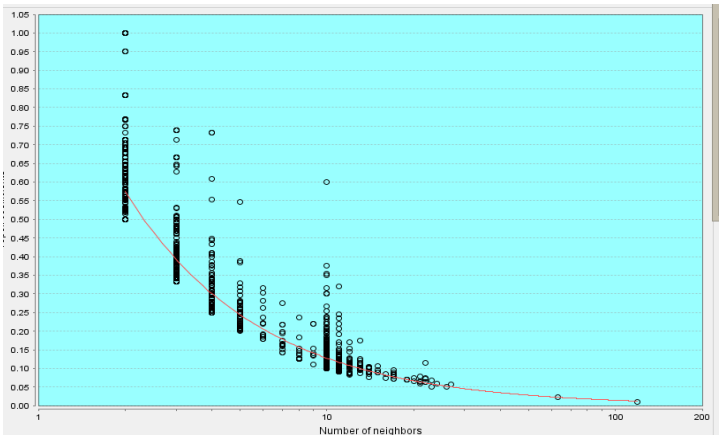


Figure:22
The graph of topological coefficients vs number of neighbours. It can be used to examine the ability of the node to have shared neighbours, in the network.

Stress distribution:

Stress distribution is a property of undirected network (a network without multiple edges) and is defined for a node n as the total number of shortest paths that pass through node n. Higher the number of shortest paths passed through a node, higher the stress for that node.

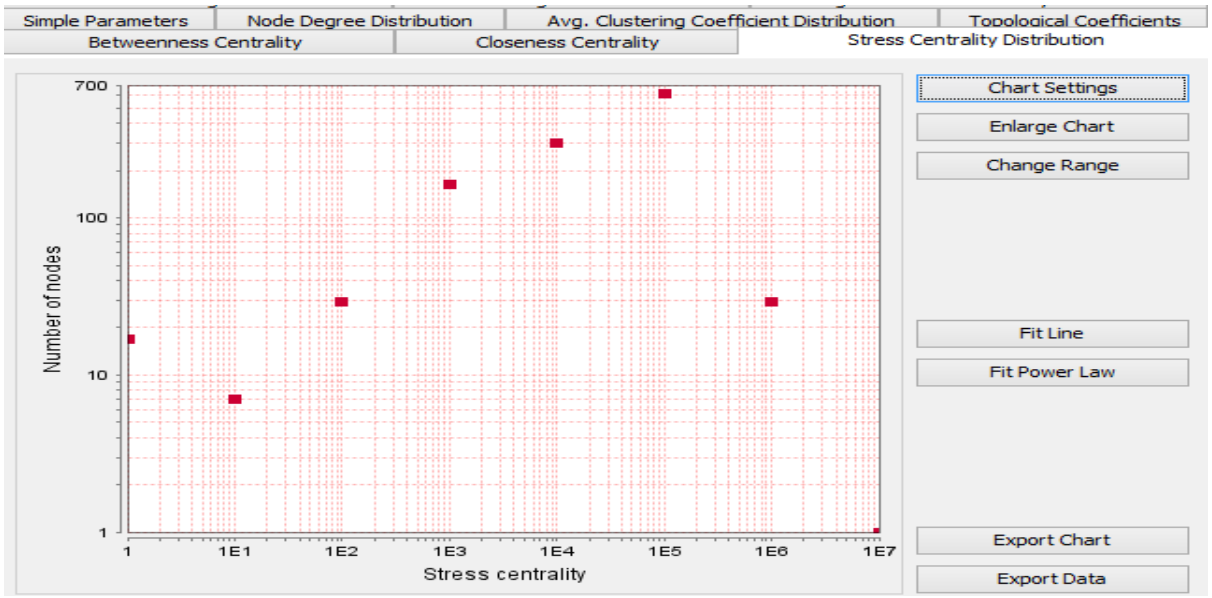


Figure 23: Graph for stress centrality distribution. If a node has high-stress centrality value, it has more influence in the network. This figure clearly shown that there are some nodes present, in our network have high-stress centrality.

After analyzing the topological properties of cervical cancer associated gene network we conclude that

- **The network made by us is a scale free network and it is evolved in nature because it fulfilled all the conditions that are necessary for being a scale free biological network.**
- **The network contains a few nodes that are dominant or showing huge influence in the network.**
- **The next task is to find out the most significant nodes, in order to pursue our goal; we took the help of R statistical programming language.**

The next step was to find communities and sub-communities present in the network. R Studio (an IDE based on R language) used for this, and scripts used is shown below:

Finding the communities by using R:

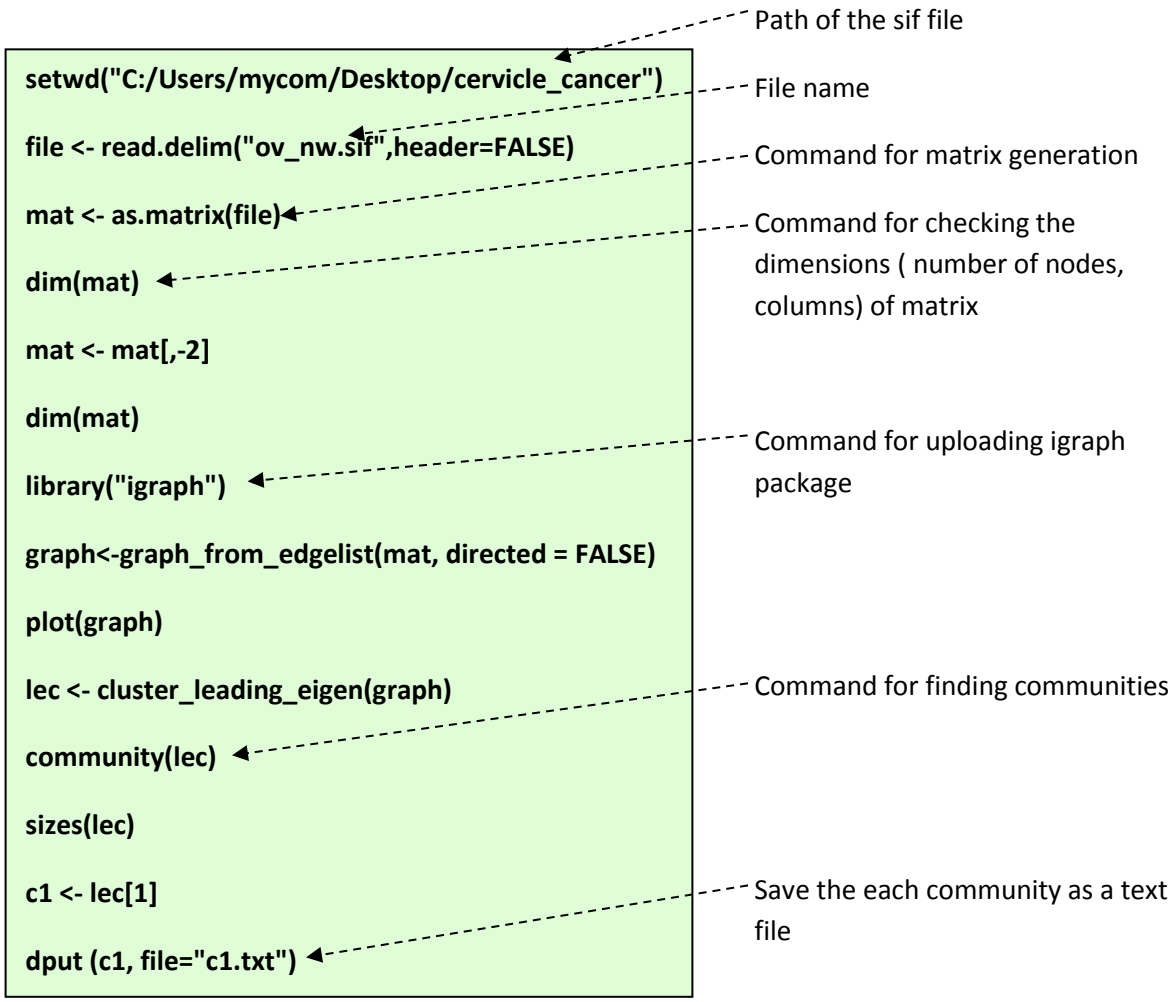


Figure 24: The above figure represents the script that is used for finding community, in R studio.

Followed steps in order to find community

1. Initially, export the network and save as .sif file
2. Set the path of the folder, in which .sif file saved
3. Provide the name of .sif file, so that it can be read
4. Make the matrix
5. Check the dimensions of matrix
6. Upload graph package in R
7. Find communities and save each as txt file for further use

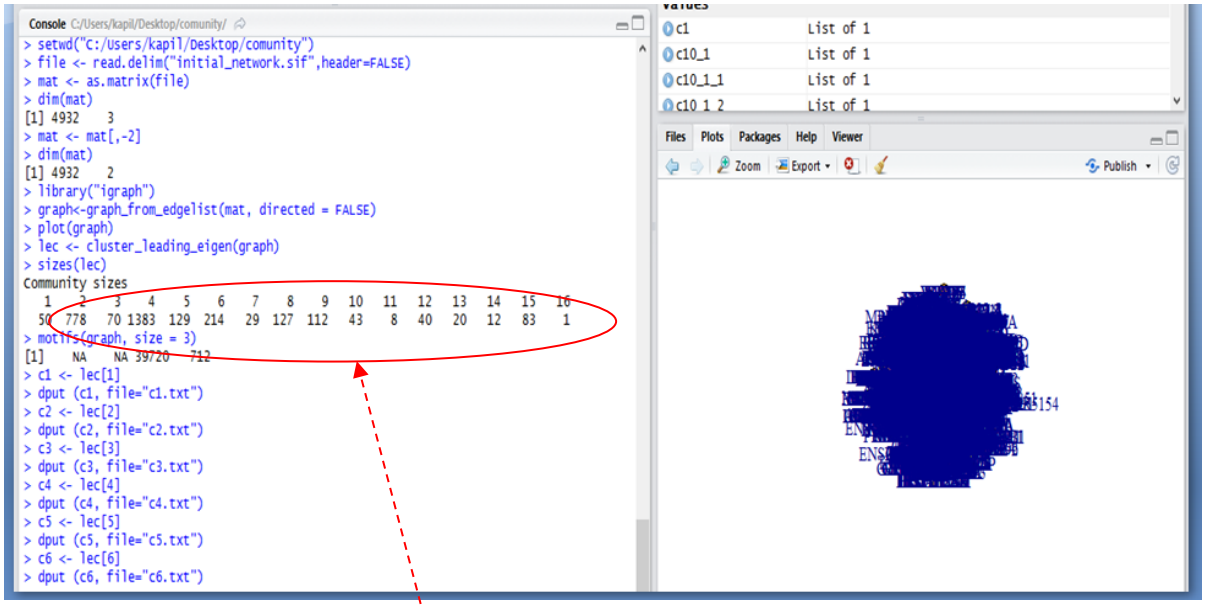


Figure 25: Initially we found that there are 16 communities in the network (at level 1)

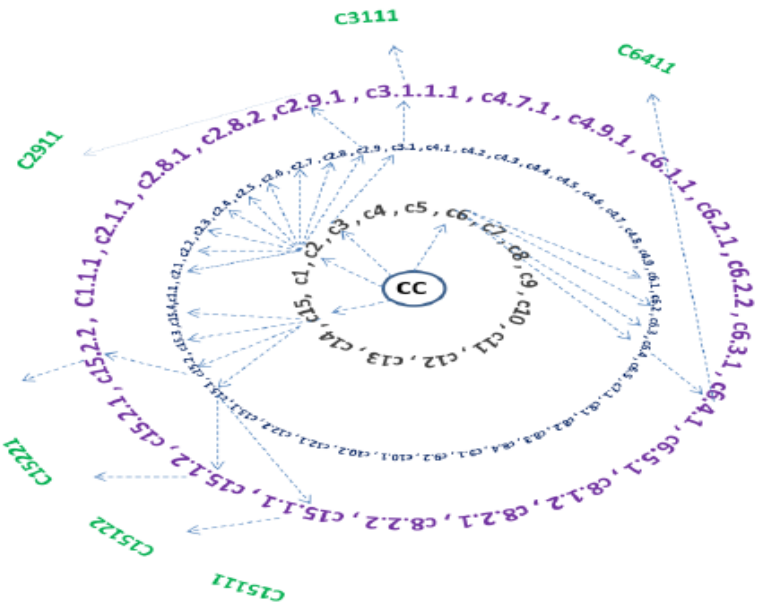


Figure 26. The system level organization of communities and sub-communities at a different level (Indicated by various concentric circles) and arrows show sub-communities constructed from previous communities/modules.

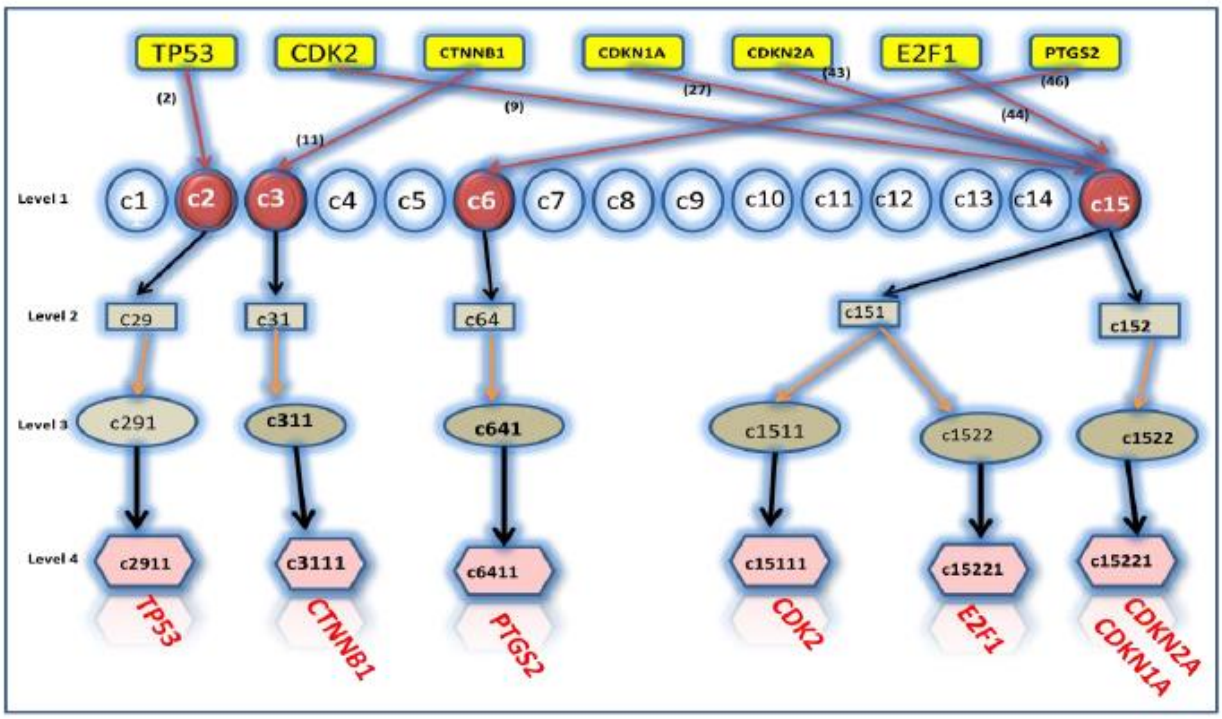


Figure 27. Paths of fundamental key regulators from complete network to motif through various modules/sub-modules at various level of organization.

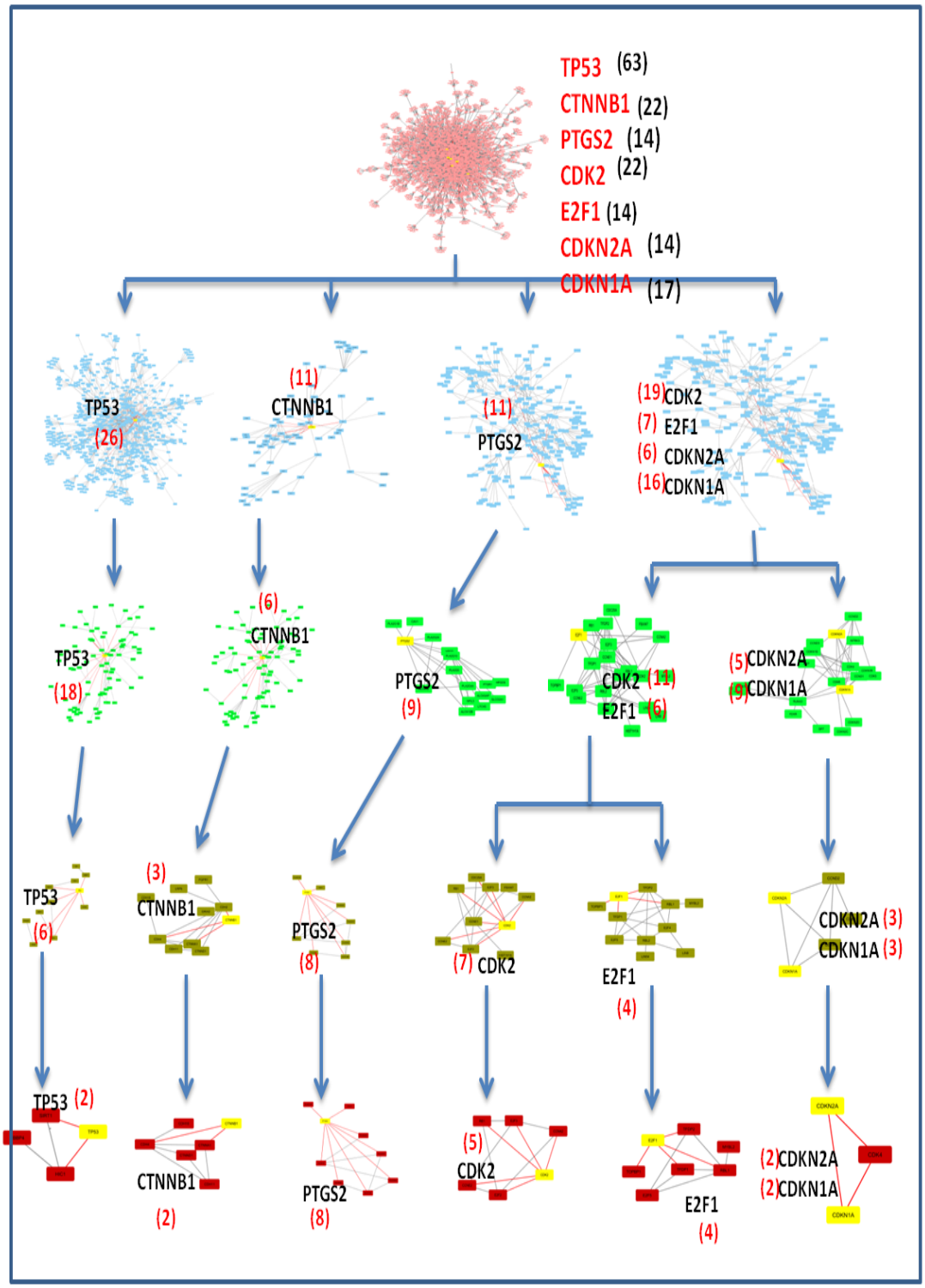


Figure 28: Communities at different level. The nodes represented by yellow colour are fundamental key regulators and highly dominating nodes in whole network (because they have highest degree in each level and also exist in closed triangular loop) . The whole concept is based on barabasi-albert model, which is a bottom-up approach.

Figure 27 shows that initially, motif combined to each other and from sub-communities. Sub-communities forms communities and communities form meta-communities and finally, the network is obtained.

With the help of R scripts, we found 16 communities at level 1 and found sub communities in further levels. We continued the process of finding communities until or unless we get motif. The motif is the closed triangular shaped tread of three nodes. As we known that biological network is not constructed, they evolved with time and follow growth and preferential attachment concepts. So, we can say that cervical cancer network is also evolved and formation of network start from a motif that is shown in figures no 25,27. Now the most important node that is also known as a key regulator of a particular network is that have a maximum degree at each level which means such nodes have great dominancy at each level of network and play a crucial role in network construction because real network follows preferential attachment which means richer gets richer. Barabasi-Albert model is a bottom up approach. Here, there are 7 important key regulators (CDK2, E2F1, CDKN1A, CDKN2A, TP53, PTGS2 and CTNNB1) that are the highest degree at each level, so that's why they are important for network and transferring of biological information through out the network. If these key regulators are destroyed the whole network would collapse. Hence these important key regulators can be used as early detection biomarkers and helpful in network medicine and design new therapeutics strategies for cervical cancer.

7. CONCLUSION:

As we know that cervical cancer is responsible for a large number of death in females, worldwide. More than 3000 genes are associated with cervical cancer progression. Here, we proved that the network made by us is a scale free network, which is the property of network which is real and evolved in nature by the passage of time. Scale free networks have some specific features and one of them is they contain hub nodes. By using R language we obtained communities, sub communities at different levels and also identify the most crucial nodes of cervical cancer PPI network called fundamental key regulators, the process of retrieving key regulators of the network is based on Barabasi-Albert model (generative model of PPI network, a backward approach). Because our network is scale free, it is robust against random node failure but vulnerable to deliberate attacks, which means if we target these fundamental key regulators (deliberate attacks), the whole network will break into tiny noncommunicating islands that will be helpful in inhibition of cervical cancer progression. Hence we can design early diagnostic markers as well as therapeutics strategies for cervical cancer, by identifying new markers.

8. REFERENCES:

1. Appleby P, Beral V, Berrington de González A, Colin D, Franceschi S, Goodhill A, Green J, Peto J, Plummer M, Sweetland S. Cervical cancer and hormonal contraceptives: collaborative reanalysis of individual data for 16,573 women with cervical cancer and 35,509 women without cervical cancer from 24 epidemiological studies. Lancet. 2007;370:1609–1621.
2. Arbyn M, Castellsagué X, de Sanjosé S, Bruni L, Saraiya M, Bray F, Ferlay J. Worldwide burden of cervical cancer in 2008. Ann Oncol. 2011;22:2675–2686.
3. Saxena U, Sauvaget C, Sankaranarayanan R. Evidence-based screening, early diagnosis and treatment strategy of cervical cancer for national policy in low- resource countries: example of India. Asian Pac J Cancer Prev. 2012;13:1699–1703.
4. Garland SM. Can cervical cancer be eradicated by prophylactic HPV vaccination? Challenges to vaccine implementation. Indian J Med Res. 2009;130:311–321.
5. Woodman CB, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer. 2007;7:11–22.
6. Choi YH, Chapman R, Gay N, Jit M. Potential overestimation of HPV vaccine impact due to unmasking of non-vaccine types: quantification using a multi-type mathematical model. Vaccine. 2012;30:3383–3388.
7. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Muñoz N. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol. 1999;189:12–19.
8. Clifford GM, Rana RK, Franceschi S, Smith JS, Gough G, Pimenta JM. Human papillomavirus genotype distribution in low-grade cervical lesions: comparison by geographic region and with cervical cancer. Cancer Epidemiol Biomarkers Prev. 2005;14:1157–1164.
9. Sukasem C, Pairoj W, Saekang N, Pombubpha H, Srichunrasami C, Pongtippan A, Junyangdikul P, Chantratita W. Molecular epidemiology of human papillomavirus genotype in women with high-grade squamous intraepithelial lesion and cervical cancer: will a quadrivalent vaccine be necessary in Thailand? J Med Virol. 2011;83:119–126.
10. Odida M, de Sanjosé S, Quint W, Bosch XF, Klaustermeier J, Weiderpass E. Human Papillomavirus type distribution in invasive cervical cancer in Uganda. BMC Infect Dis. 2008;8:85.
11. Sahasrabuddhe VV, Mwanahamuntu MH, Vermund SH, Huh WK, Lyon MD, Stringer JS, Parham GP. Prevalence and distribution of HPV genotypes among HIV-infected women in Zambia. Br J Cancer. 2007;96:1480–1483.
12. Desruisseau AJ, Schmidt-Grimminger D, Welty E. Epidemiology of HPV in HIV-positive and HIV-negative fertile women in Cameroon, West Africa. Infect Dis Obstet Gynecol. 2009;2009:810596.

13. Bosch FX, Burchell AN, Schiffman M, Giuliano AR, de Sanjose S, Bruni L, Tortolero-Luna G, Kjaer SK, Muñoz N. Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia. Vaccine. 2008;26 Suppl 10:K1–16.
14. Trottier H, Mahmud S, Costa MC, Sobrinho JP, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, Franco EL. Human papillomavirus infections with multiple types and risk of cervical neoplasia. Cancer Epidemiol Biomarkers Prev. 2006;15:1274–1280.
15. Rodríguez AC, Schiffman M, Herrero R, Wacholder S, Hildesheim A, Castle PE, Solomon D, Burk R. Rapid clearance of human papillomavirus and implications for clinical focus on persistent infections. J Natl Cancer Inst. 2008;100:513–517.
16. Naucler P, Ryd W, Törnberg S, Strand A, Wadell G, Elfgren K, Rådberg T, Strander B, Johansson B, Forslund O, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. N Engl J Med. 2007;357:1589–1597.
17. Winer RL, Hughes JP, Feng Q, O'Reilly S, Kiviat NB, Holmes KK, Koutsky LA. Condom use and the risk of genital human papillomavirus infection in young women. N Engl J Med. 2006;354:2645–2654.
18. Wen LM, Estcourt CS, Simpson JM, Mindel A. Risk factors for the acquisition of genital warts: are condoms protective? Sex Transm Infect. 1999;75:312–316.
19. Harper DM, Demars LR. Primary Strategies for HPV Infection and Cervical Cancer Prevention. Clin ObstetGynecol. 2014;57:256–278.
20. FUTURE II Study Group. Quadrivalent vaccine against human papillomavirus to prevent high-grade cervical lesions. N Engl J Med. 2007;356:1915–1927.
21. ACIP Adult Immunization Work Group, Bridges CB, Woods L, Coyne-Beasley T. Advisory Committee on Immunization Practices (ACIP) recommended immunization schedules for adults aged 19 years and older- United States, 2013. MMWR Surveill Summ. 2013;62 Suppl 1:9–19.
22. Brankovic I, Verdonk P, Klinge I. Applying a gender lens on human papillomavirus infection: cervical cancer screening, HPV DNA testing, and HPV vaccination. Int J Equity Health. 2013;12:14.
23. Calloway C, Jorgensen CM, Saraiya M, Tsui J. A content analysis of news coverage of the HPV vaccine by U.S. newspapers, January 2002-June 2005. J Womens Health (Larchmt) 2006;15:803–809.
24. Saslow D, Castle PE, Cox JT, Davey DD, Einstein MH, Ferris DG, Harper DM, Kinney W, Moscicki AB, Noller KL, Wheeler CM, Ades T, Andrews KS, Doroshenk MK, Kahn KG, Schmidt C, Shafey O, Smith RA, Partridge EE; Gyenocologic Cancer Advisory Group, Garcia F. American Cancer Society Guideline for human papillomavirus (HPV) vaccine use to prevent cervical cancer and its precursors. CA Cancer J Clin. 2007;57:7–28.
25. Casper MJ, Carpenter LM. Sex, drugs, and politics: the HPV vaccine for cervical cancer. Sociol Health Illn. 2008;30:886–899.

26. Leader AE, Weiner JL, Kelly BJ, Hornik RC, Cappella JN. Effects of information framing on human papillomavirus vaccination. J Womens Health (Larchmt) 2009;18:225–233.
27. Hull SC, Caplan AL. The case for vaccinating boys against human papillomavirus. Public Health Genomics. 2009;12:362–367.
28. Taira AV, Neukermans CP, Sanders GD. Evaluating human papillomavirus vaccination programs. Emerg Infect Dis. 2004;10:1915–1923.
29. Schiffman M, Castle PE. The promise of global cervical-cancer prevention. N Engl J Med. 2005;353:2101–2104.
30. Ladner J, Besson MH, Hampshire R, Tapert L, Chirenje M, Saba J. Assessment of eight HPV vaccination programs implemented in lowest income countries. BMC Public Health. 2012;12:370.
31. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *CA Cancer J Clin* ; 62: 147-172.
32. Fisher JW, Brundage SI. The challenge of eliminating cervical cancer in the United States: a story of politics, prudishness, and prevention. Women Health. 2009;49:246–261.
33. Peterson NB, Murff HJ, Cui Y, Hargreaves M, Fowke JH. Papanicolaou testing among women in the southern United States. J Womens Health (Larchmt) 2008;17:939–946.
34. Arbyn M, Bergeron C, Klinkhamer P, Martin-Hirsch P, Siebers AG, Bulten J. Liquid compared with conventional cervical cytology: a systematic review and meta-analysis. Obstet Gynecol. 2008;111:167–177.
35. Moyer VA. Screening for cervical cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2012;156:880–891, W312.
36. Rijkaart DC, Berkhof J, van Kemenade FJ, Coupe VM, Rozendaal L, Heideman DA, Verheijen RH, Bulk S, Verweij W, Snijders PJ, et al. HPV DNA testing in population based cervical screening (VUSA-Screen study): results and implications. Br J Cancer. 2012;106:975–981.
37. Stoler MH, Castle PE, Solomon D, Schiffman M. The expanded use of HPV testing in gynecologic practice per ASCCP-guided management requires the use of well-validated assays. Am J Clin Pathol. 2007;127:335–337.
38. Ideker, T. et al. (2008) Protein networks in disease. Genome Res., 18, 644652.
39. Palla, G. et al. (2005) Uncovering the overlapping community structure of complex networks in nature and society. Nature,435, 814818.
40. Santo, F. et al. (2010) Community detection in graphs. *Phys Rep.*, 486, 75174.

41. Pieroni, E., et al. (2008) Protein networking. insights into global functional organization of proteomes. Proteomics. 8,799816.
42. Barzel, B. et al. (2013) Universality in network dynamics. Nat Phys., 9, 673681.

43. Ravasz, E. et al. (2002) Hierarchical organization of modularity in metabolic networks. Science, 297, 15511555.

44. Girvan, M. et al. (2002) Community structure in social and biological networks. Proc Natl Acad Sci USA, 99, 78217826.

45. Maslov, S. et al. (2002) Specificity and stability in topology of protein networks. Science, 296, 910913.

46. Ravasz, E. et al.(2003) Hierarchical organization in complex networks. Phys. Rev., E 67, 026112.

47. Barabasi, A.L. et al. (2004) Network biology: understanding the cells functional organization. Nat Rev Genet., 5, 10113.

48. Romualdo, P.S. et al. (2001) Dynamical and Correlation Properties of the Internet. Phys. Rev. Lett., 87, 258701.

49. Nafis, S. et al. (2015) Apoptosis regulatory protein-protein interaction demonstrates hierarchical scale-free fractal network.Brif. Bioinf., 16, 675-699 .

50. Nafis, S. et al. (2016) Identification of key regulators and their controlling mechanism in a combinatorial apoptosis network: a systems biology approach. Mol. BioSyst., DOI: 10.1039/c6mb00526h.

51. Clauset, A. et al. (2009) Power-law distributions in empirical data. *SIAM Rev. Soc. Ind. Appl. Math.*, 51, 661703.

52. Brandes, U. (2001) A faster algorithm for betweenness centrality. *J. Math. Sociol.*, 25, 163-177.

53. Bonacich, P. (1987) Power and centrality: a family of measures. *Am. J. Sociol.*, 92, 1170-1182.

54. Ahmad, F.K. et al. (2012) The inference of cervical cancer metastasis through gene regulatory networks. Journal of biomedical informatics, 45, 350-362.

55. Borgatti, S.P. et al. (2006) A Graph-theoretic perspective on centrality. Social Networks, 28, 466484.

56. Mason, O. et al. (2007) Graph theory and networks in biology. *IET Syst. Biol.*, 1, 89-119.

57. Canright, G. et al. (2004) Roles in networks. *Sci. Comput. Program.*, 53, 195-214.

58. Canright, G. et al. (2006) Spreading on networks: a topographic view. *Complexus,* 3, 131-146.

59. Peng, X., et al. (2015) Rechecking the Centrality-Lethality Rule in the Scope of Prote... Subcellular Localization Interaction Networks. PLoS One., 10, e0130743.

60. Jeong, H., et al. (2001) Lethality and centrality in protein networks. Nature 411, 4142 (2001).

61. Kauffman, S.A. (1993) *The Origins of Order: Self-organization and Selection in Evolution*. Oxford University Press.

62. Whirl-Carrillo, M., et al. (2012) Pharmacogenomics Knowledge for Personalized Medicine. Clinical Pharmacology and Therapeutics. 92,414-417.

63. Cannistraci, C.V., et al. (2013) From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. Sci. Rep., 3, 1613.

64. Mandelbrot B., et al. (1997) A multifractal model of asset returns.Cowles Foundation Discussion, Paper No. 1164.

65. Prigogine, I., et al. (1984) Order out of Chaos.*Bantam Books, New York*.

66. Heylighen F., et al. (2001) The science of self-organization and adaptivity.*The encyclopedia of life support systems*, 5,253-280.

67. Holland, J.H., (1996) Hidden Order: How adaptation builds complexity.*Basic book*.

68. Holland, J.H., (1992) Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology.

69. *Control and Artificial Intelligence*, *MIT Press, Cambridge MA, USA*.

70. Hernandez-Toro J., et al. (2007) APID2NET: unified interactome graphic analyzer. Bioinformatics, 23, 24952497.

71. Barab´asi, A.L., et al. (1999) Emergence of scaling in random networks. *Science*, 286, 509-512.

72. Albert, R. and Barabasi A.L. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47 (2002).

73. Maslov, S., et al. (2002) Specificity and stability in topology of protein networks. *Science*, 296, 910-913 (2002).

74. Barrat, A., et al. (2004) The architecture of complex weighted networks. Proc. Natl. Acad. Sc., 101, 3747-3752.

75. Traag V.A., et al. (2011) Narrow scope for resolution-limit-free community detection. arXiv, 1104.

76. Traag V.A., et al. (2015) Significant Scales in Community Structure. Sci. Rep., 3, 2930.

77. Gakidou, E., Nordhages, S., Obermeyer, Z. **Coverage of cervical cancer screening in 57 countries: low average levels and large inequalities.** *PLoS Medicine*. 2008;5:863–868.

78. Facione, N.C., Giancarlo, C.A. **Narratives of Cervical symptom discovery and cancer diagnosis: psychologic risk for advanced cancer at diagnosis.** *Cancer Nursing*. 1998;21:430–440.

79. Heystek, M.J., De Jonge, E.T., Meyer, H.P., Lindeque, B.G. **Screening for cervical neoplasia in Mamelodi: lessons from an unscreened population.** *South African Medical Journal*. 1995;85:1180–1182.

80. LK Borysiewicz, A Fiander, M Nimako , etal: A recombinant vaccinia virus encoding human papillomavirus types 16 and 18, E6 and E7 proteins as immunotherapy for cervical cancerLancet 347: 1523– 1527,1996 Crossref, Medline

81. Phase II trial of ifosfamide and mesna in patients with advanced or recurrent squamous carcinoma of the cervix who had never received chemotherapy: A Gynecologic Oncology Group study Am J Obstet Gynecol 168: 805– 807,1993 Crossref, Medline

9. APPENDIX

Functions of important key regulators associated with Cervical Cancer

CDK2	Cyclins Dependant Kinase 2	This gene encodes a member of a family of serine/threonine protein kinases that participate in cell cycle regulation. The encoded protein is the catalytic subunit of the cyclin-dependent protein kinase complex, which regulates progression through the cell cycle. Activity of this protein is especially critical during the G1 to S phase transition
E2F1	E2F transcription factor 1	The protein encoded by this gene is a member of the E2F family of transcription factors. The E2F family plays a crucial role in the control of cell cycle and action of tumor suppressor proteins and is also a target of the transforming proteins of small DNA tumor viruses
TP53	Tumor Protein P53	This gene encodes a tumor suppressor protein containing transcriptional activation, DNA binding, and oligomerization domains. The encoded protein responds to diverse cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. Mutations in this gene are associated with a variety of human cancers.
CDKN1A	Cyclin Dependent Kinase Inhibitor 1A	This gene encodes a potent cyclin-dependent kinase inhibitor. The encoded protein binds to and inhibits the activity of cyclin-cyclin-dependent kinase2 or -cyclin-dependent kinase4 complexes, and thus functions as a regulator of cell cycle progression at G1. The expression of this gene is tightly controlled by the tumor suppressor protein p53, through which this protein mediates the p53-dependent cell cycle G1 phase arrest in response to a variety of stress stimuli. This protein can interact with proliferating cell nuclear antigen, a DNA polymerase accessory factor, and plays a regulatory role in S phase DNA replication and DNA damage repair.
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A	This gene generates several transcript variants which differ in their first exons. At least three alternatively spliced variants encoding distinct proteins have been reported, two of which encode structurally related isoforms known to function as inhibitors of CDK4 kinase. The remaining transcript includes an alternate first exon located 20 Kb upstream of the remainder of the gene; this transcript contains an alternate open reading frame (ARF) that specifies a protein which is structurally unrelated to the products of the other variants. This ARF product functions as a stabilizer of the tumor suppressor protein p53 as it can interact with, and sequester, the E3 ubiquitin-protein ligase MDM2, a protein responsible for the degradation of p53. In spite of the structural and functional differences, the CDK inhibitor isoforms and the ARF product encoded by this gene, through the regulatory roles of CDK4 and p53 in cell cycle G1 progression, share a common functionality in cell cycle G1 control. This gene is frequently mutated or deleted in a wide variety of tumors, and is known to be an important tumor suppressor gene.

PTGS2	Prostaglandin-Endoperoxide Synthase 2	This gene encodes an enzyme that is a member of the prostaglandin G/H synthase family. The encoded protein converts arachidonic acid to prostaglandin endoperoxide H2 which is a key enzymatic step in prostaglandin biosynthesis. This gene is the inducible prostaglandin G/H synthase family member that is upregulated during inflammation. Aberrant regulation of this gene is associated with cancer progression in several tissues and an increased risk of cardiovascular events.
CTNNB1	Catenin Beta-1	Key downstream component of the canonical Wnt signaling pathway. In the absence of Wnt, forms a complex with AXIN1, AXIN2, APC, CSNK1A1 and GSK3B that promotes phosphorylation on N-terminal Ser and Thr residues and ubiquitination of CTNNB1 via BTRC and its subsequent degradation by the proteasome. In the presence of Wnt ligand, CTNNB1 is not ubiquitinated and accumulates in the nucleus, where it acts as a coactivator for transcription factors of the TCF/LEF family, leading to activate Wnt responsive genes. Involved in the regulation of cell adhesion, as component of an E-cadherin:catenin adhesion complex. Acts as a negative regulator of centrosome cohesion. Involved in the CDK2/PTPN6/CTNNB1/CEACAM1 pathway of insulin internalization. Blocks anoikis of malignant kidney and intestinal epithelial cells and promotes their anchorage-independent growth by down-regulating DAPK2.