International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 7 Issue 3, May-June 2023 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Esophageal Cancer & Associated Genes

Abdullah Haidary

Master of Science (Bioinformatics), Savitribai Phule Pune University (University of Pune), Pune, Maharashtra, India

ABSTRACT

Cancer is a genetically based illness that results from alterations in the genes that regulate how our cells operate. More genetic alterations, such as DNA mutations, are present in cancer cells than in healthy ones. Some of these alterations might not be related to the disease at all; they might be its effect rather than its cause. This initiative makes an effort to examine and research esophageal cancer. It can develop when a malignant tumor grows in the lining of the esophagus and has two variants, making it the eighth most common cancer in the world. Esophageal squamous cell carcinoma and esophageal adenocarcinoma. Risk factors, symptoms, and indicators are discussed.

The risk factors for esophageal cancer include smoking, drinking alcohol, and other behaviors. 575 altered genes that are connected to esophageal cancer, according to the cancer genome atlas. 180 examples of this malignancy were found in the cancer genome atlas. To further understand their relationships, these genes were examined. Additionally, using the k-means clustering approach, a network of all the genes connected to this malignancy has been developed in 3 and 5 groups.

KEYWORDS: *Esophageal cancer, genes, stages, network*

How to cite this paper: Abdullah Haidary "Esophageal Cancer & Associated Genes" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-7 | Issue-3, June 2023, pp.644-659, URL: www.ijtsrd.com/papers/ijtsrd45003.pdf

IJTSRD45003

Copyright © 2023 by author (s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (http://creativecommons.org/licenses/by/4.0)

INTRODUCTION

The term "cancer" refers to a collection of disorders. Every type of cancer causes more body cells to start dividing uncontrollably and spread to neighboring tissues. Cancer may develop practically anywhere in the billions of cells that make up the human body. When the body requires new cells, human cells divide regularly to create them. New cells are produced when the old or damaged cells die. But when cancer starts, this methodical procedure fails. As cells multiply, injured or new cells should survive when they should, and new cells form when old ones are no longer required. These additional cells have the ability to proliferate indefinitely and can become stems known as tissues.

Strong tissues, which are tissues, are formed by many malignancies. Leukemia and other blood malignancies frequently do not develop robust tissues. Cancerous tissues provide a threat because they have the potential to spread or attack surrounding tissues. Additionally, when these tissues expand, some cancer cells depart and go through the blood or lymph nodes to different regions of the body where they establish new tissues independent of the original tumor. Malignant tissue does not spread or assault neighboring tissues like deadly tissue does. However, benign tissue can occasionally be rather massive (Louis and Missouri, 2021). While the lethal tissue occasionally comes back after removal, they mostly do not. Contrary to many other toxic tissues in the body, damaged brain tissue can be fatal.

How cancer Arises?

Because genetic mutations regulate how our cells function, particularly how they grow and divide, cancer is a genetic condition. Cancer-causing genetic mutations can be passed down from parents. Genetic flaws or DNA damage brought on by genetic mutations may lead them to reoccur over the course of a person's lifespan. Smoke and radiation from the sun's UV rays are examples of chemicals and radiation that naturally cause cancer. Each cancer has a particular set of genetic alterations. More alterations will occur when the malignancy becomes worse. Distinct cells in the same plant may have distinct genetic alterations.

Typically, cancer cells have more genetic alterations than healthy cells do, including DNA mutations. Some of these alterations might not be connected to

cancer; rather than being the cause of cancer, they might be its result.

Esophageal cancer

It ranks as the eighth most prevalent cancer worldwide (Howlader et al., 2016). It can happen if a fatal tumor develops on the lining of the esophagus. There are currently few clinical trials available for the initial diagnosis and treatment of cancer, which has led to a life expectancy of 15–30%. It is one of the most lethal malignancies. Esophageal squamous cell carcinoma is the most common kind of histology cancer in China and the rest of the globe, whereas esophageal adenocarcinoma has been on the rise for decades in western nations. Esophageal cancer has a distinct national and ethnic origin.

Only a small portion of esophageal cancer and death rates are caused by risk factors including alcohol misuse and smoking worldwide, which is consistent with the fact that some of China's most prevalent regions have low drinking and smoking rates. Numerous genome connection studies have been carried out to define the cell base, to uncover genetic anomalies that cause cancer, and to direct the creation of efficient targeted medicines and diagnostic biomarkers. On the other hand, it is notable that this malignant disease affects more than 450,000 individuals globally each year and is the sixth most prevalent cancer in men and the ninth most common in women (Boyle and Levin, 2008). Squamous cell carcinoma (ESCC) accounts for around 90% of occurrences of cervical cancer, with the remaining cases being adenocarcinomas (EA). With an average incidence of 14.6 instances per 100,000 cases, Jaishan, China, had the highest recorded rate of ESCC mortality. The United Kingdom's other regions and Scotland both had the highest incidence rate for EA years (6.6 per 100,000) (Curado et al., 2007).

What causes esophagus cancer?

The diagnosis of cancer may be recorded due to a few circumstances. According to the expert, toxic chemicals like alcohol and cigarette usage can destroy the DNA cells in the stomach, which can lead to constipation cancer. DNA damage can also result from persistent gastrointestinal tract irritation, such as that brought on by reflux, Barrett's throat, achalasia, Plummer-Vinson syndrome, or scar-swallowing scar tissue. An alteration in the DNA of our cells is what leads to cancer. Our genes, which govern how our cells function, are made out of a substance called DNA in each of our unique cells. Some bacteria regulate how quickly cells divide to form new cells and how quickly they perish.

- Oncogenes are specialized genes that promote cell growth, division, and survival.
- Genes that stop tumors include those that help regulate cell division or lead cells to die when they should.

DNA mutations that activate oncogenes or genetic changes that shut down tumor-suppressing mutations can both lead to cancer. Cells begin to expand out of control as a result. The development of cancer frequently requires mutations in a variety of distinct genes.

Cancer cells that are being observed typically have several distinct gene alterations in their DNA. The existence of common genetic alterations in all esophageal malignancies is unclear, though.

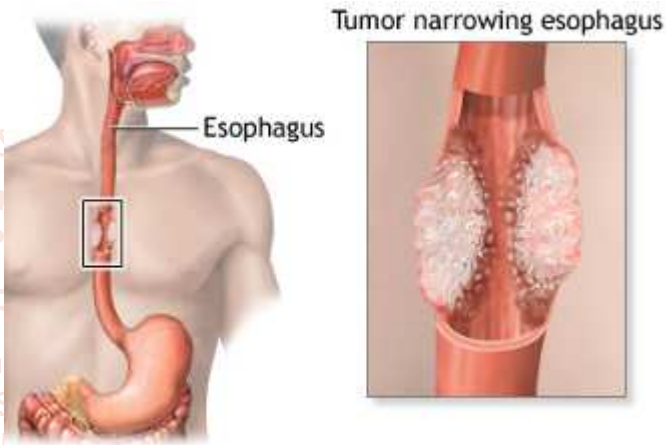

Figure 1: tumor narrowing the esophagus (source:https://medlineplus.gov/ency/imagepages/19928.htm)

Types of esophageal cancer

Squamous cell carcinoma of the esophagus and esophageal adenocarcinoma are the two primary types of esophageal cancer.

1. Squamous cell carcinoma of the esophagus

It happens when the cancerous cells that line the throat's flat, tiny cells. Although it can develop anywhere, this form of cancer often develops above or in the center of the neck. With 90% of the world's population represented by it, it is the most prevalent (EC) component outside of the United States. (Lepage et al., 2008).

China, Central Asia, and East and South Africa have the most expensive costs.

In the United States, the incidence of squamous cell carcinoma is to three years per 100,000 years per person. (Cook and Chow, 2009).

This condition occurs more frequently in black people and is comparable in gender between the ages of 60 and 70. Smoking, drinking, and achalasia are all significant esophageal squamous cell carcinoma risk factors. (Zendehdel et al., 2011).

2. Esophageal adenocarcinoma

The most prevalent kind of esophageal cancer in North America and Europe is esophageal adenocarcinoma (Freedman et al., 2007). According to National Cancer Institute data from 2013, the majority of cases affect people over the age of 50, and the incidence for people 65 and older ranges from 11.8 to 16.3 per 100,000 years per person. Men are eight times more likely to develop the disease than women are, and white people are five times more likely to do so than black people.

Smoking, obesity, and gastro esophageal reflux syndrome are the main risk factors for this malignancy (Rubenstein and Taylor, 2010). With a low conversion rate, Barrett esophagus is recognized as an esophageal cancer precursor condition. According to a study of 11,028 individuals with low- and high-grade dysplasia of the esophagus who had been monitored for more than five years, the annual incidence of esenapolis adenocarcinoma was 0.12%. Helicobacter pylori infection has been linked to a 41% lower incidence of esophageal cancer.

Because of the atrophy of the stomach mucosa brought on by Pylori, it is thought that the amount of gastric acid secreted, which leads to reflux disease and Barrett esophagus, is reduced (Xie et al., 2013). The company is still being looked at, and the American College of Gastroenterology's recommendations for treating Pylori infection are still given high marks.

Symptoms and signs

The following indications or symptoms might be experienced by someone with esophageal cancer. Congenital cancer patients occasionally do not experience these alterations. Or, a separate non-cancer health issue might be to blame for the symptoms. Difficulty and pain with swallowing, particularly when eating meat, bread, or raw vegetable. As the tumor grows, it can block the pathway to the stomach. Even liquid may be painful to swallow.

- Pressure or burning in the chest.
- Indigestion or heartburn.
- Vomiting.
- Frequent choking on food.
- Unexplained weight loss.
- Coughing or hoarseness.
- Pain behind the breastbone or in throat.

First-stage esophageal cancer may be fully asymptotic or present with some fictitious symptoms such dyspepsia, chest discomfort, or heartburn. Alternately, individuals may exhibit symptoms like bloody stools or anemia due to iron shortage. Patients may have odynophagia, mild to moderate dysphagia, or an external feeling. Esophagogastroduodenoscopy (EGD) and diagnostic mucosal biopsy can also be used to diagnose these symptoms in most cases. With timely high endoscopic Barrett's disease surveillance, some early malignancies are identified.

Esophageal cancer symptoms typically appear suddenly (ASCO foundation, 2019). Tumor growth is the primary symptom of esophageal cancer. After first having trouble swallowing solid food, people gradually have trouble swallowing ground food and liquids. 90% of patients report prolonged dysphagia as their primary symptom, with 50% or less light retention or when the luminal diameter is smaller than 13 mm. Patients with advanced illness are more likely to have weight loss and anorexia (caused by a poor diet for dysphagia or as a side effect of cancer), which puts them at risk for malnutrition. About half of patients experience odynophagia. Food cravings and pneumonia can result from esophageal blockage. Back or back discomfort, as well as tumor growth, can be caused by a pericardial or mediastinum tumour enlargement. The paralysis of the recurrent laryngeal nerves is frequently linked to hoarseness. Internal involvement or touching the abdomen might cause hiccups.

If the tumor is wounded and sensitive, hemorrhage or gastrointestinal bleeding may develop, as well as tiredness. If you have a persistent cough or pneumonia, you should tell your doctor in case a plant assault on surrounding airways resulted in a trachea or bronchiole-esophageal fistula.

What are the stages of esophageal cancer?

After a person is diagnosed with recording cancer, doctors will try to determine if it has spread, and if so, how far. This process is called stage. The cancer category describes how much cancer is in the body. It helps to find out how serious the cancer is and how to treat it. Doctors use the cancer category when talking about survival statistics (Posner et al., 2019).

The first stage cancer is called stage 0 (grade dysplasia). Then go from paragraph I (1) to IV (4). As a rule, the decline in the number, the cancer has spread slowly. A higher rate, such as stage IV, indicates that the cancer is more prevalent. And within the category, the previous book means lower category. Although each person's cancer experience is different, cancer with the same stages tends to have the same vision and is often treated the same way.

Most esophageal cancers start in the innermost part of the esophagus (epithelium) and grow into deeper layers over time.

The stages of constipation cancer are given numbers I to IV; as the number increases, the cancer has progressed significantly. The categories are:

Stage 0: Abnormal cells (not yet cancerous) are found only in a layer of esophagus cell cells.

Stage I: Cancer cells are found only in a layer of linear cells of the esophagus.

Stage II: Cancer has reached the lining of the muscles or the outer wall of the esophagus. In addition, the cancer may spread to nearby 1 to 2 lymph nodes (small part of the body system).

Stage III: Cancer has reached deep into the lining of the internal muscles or the wall of the connective tissue. It may spread across the abdomen into a nearby organ or spread to multiple lymph nodes near the abdomen.

Phase IV: This is the most advanced stage. The cancer has spread to other parts of the body or to lymph nodes far from the throat.

How the stage determined?

The most commonly used stage cancer screen system is the American Joint Committee on Cancer (AJCC) TNM system, which is based on three components:

Size (size) of the tumor (T): How advanced is the cancer of the stomach wall? Has the cancer spread to nearby organs or organs? To learn about the layers of the throat wall briefly below is explained.

Spread of nearby lymph nodes (N): cancer spread to nearby lymph nodes.

Distribution (metastasis) in remote areas (M): Has the cancer spread to distant lymph nodes or to distant organs such as the lungs or liver. The numbers or letters after T, N and M give more details about these things. Higher numbers mean that the cancer is more advanced.

Once the individual T, N, and M categories have been identified, this information is integrated into a process called stage aggregation to allocate to the whole category. For more information see Cancer Staging (Niederhuber et al., 2020)

Esophageal Cancer Risk Factors

Anything that raises your risk of contracting an illness like cancer is considered to be a risk. Risk factors for various malignancies vary. Smoking is one example of a dangerous drug that can be substituted. Others, like an individual's age or family history, are immutable.

Numerous variables that may impact your chance of developing cancer have been identified by scientists. Adenocarcinoma of the esophagus is more likely to develop in certain people, while squamous cell carcinoma of the esophagus is more likely in others.

However, having a risk or even greater does not guarantee that you will develop cancer. Additionally, not all individuals with the condition have identified risk factors.

There are various risk factors between the ESCC and the EAC. They have been shown to be major causes of harm to each histologic type.

Both smoking and drinking are known risk factors for ESCC, with heavy smokers having a risk that is 50% higher than that of non-smokers and non-smokers. It has recently been shown that a lack of the enzyme aldehyde dehydrogenase 2 (ALDH2), which is responsible for the so-called alcohol response, increases the risk of alcohol-related ESCC. In the East Asian population, there is a variant of ALDH2 caused by the substitution of lysine for glutamate at position 487, with the lysine allele rendering the protein it codes for inactive (Yoshida et al., 1984). The risk of ESCC can also rise while drinking hot beverages (Baan et al., 2007). Additionally, people with achalasia are more likely to acquire ESCC, and both ESCC and EAC can be side effects of post-traumatic stress disorder. The data is conflicting, however oncogenic papillomaviruses may raise the risk of ESCC.

Smoking, obesity, and gastro esophageal reflux disease (GERD) are risk factors for the EAC. When compared to people with less frequent episodes, individuals with at least weekly GERD signs had a five-fold increased probability of developing EAC, while daily signals have a seven-fold increased likelihood. In comparison to non-smokers, the risk of esophageal AC and stomach cardiac AC was 2.32 in current smokers and 1.62 in smokers, respectively. However, meta-analyses have clearly shown that there is no connection between alcohol use and the incidence of esophageal and gastric cardiac AC (Tramacere et al., 2011). The highest body mass index (BMI) that will correlate to a summary of the gastro esophageal AC estimation is 1.5, according to systematic reviews and meta-analyses. According to a recent prospective research of groups in the United States, those with a BMI below 35 had a 3.67 hazard ratio compared to people with a normal BMI. While adipokines and cytokines secreted from adipocytes and inflammatory cells are known to affect plant development, obesity may happen mechanically in reflux. Infection with Helicobacter pylori has been shown to lower the incidence of EAC by 41%

through gastritis atrophy and decreased acidity. The risk of ESCC and EAC is raised by radiotherapy for thoracic disorders such breast cancer and Hodgkin's lymphoma. Over time, both the ESCC and the EAC have increased. With up to eight males or one EAC woman and three men or one ESCC woman, there is a dominant masculine presence.

The majority of the fat in obese men is distributed around the abdomen, and a greater belly size is linked to a higher risk of EAC. However, the incidence of smoking and alcohol use among males can be used to explain why the ESCC is larger in men.

Although the effect of estrogen inhibition on the growth of esophageal cancer cells has been reported, there is no strong conclusion on the role of estrogen in human cancer etiology. The ESCC family type is uncommon, although family reunification has been reported in a high-risk area in China. In contrast, Barrett's esophagus family reunion with the EAC has been observed (Morton et al., 2014).

7% of Barrett's esophagus and EAC patients in a European cohort research were domestic cases. It is debatable whether endoscopic screening for high-risk people is beneficial. It has been claimed that lugol chromo endoscopy and cutting-edge imaging technologies such tiny imaging bands can both be useful for finding the initial ESCC (Muto et al., 2010).

In addition, it is advised that patients with newly discovered head and neck cancer undergo endoscopic esophageal screening. The efficiency of endoscopic screening or testing, however, among those most exposed to ESCC risk factors has not been examined in any trials. Contrarily, endoscopic screening is advised for individuals with numerous Barrett's esophagus risk factors, despite the fact that no randomized controlled research has demonstrated any benefit in reducing the risk of esophageal cancer-related mortality (Spechler, 2013). Endoscopic screening is advised for 3-5 years in people with Barrett's esophagus without dysplasia, and endoscopic therapy is the recommended course of action for those with high-grade dysplasia (HGD). However, due to a low yearly cancer rate of 0.1-0.3% among patients with non-dysplastic Barrett's esophagus (NDBE), there has recently been discussion about increasing recruiting of or terminating employment with these patients.

Table 1: risk factors of esophageal cancer

Squamous cell carcinoma	Adenocarcinoma
Cigarette smoking	Gastro-esophageal reflux disease
Alcohol drinking	Barrett's esophagus
ALDH2 deficiency	Reflux symptoms
Drinking very hot liquids	Obesity
Achalasia	Cigarette smoking
Caustic injury	Diet (high in processed meat, low in fruits, vegetables)
History of thoracic radiation	History of thoracic radiation
Tylosis	Anticholinergic agents
Human papilloma virus infection	Family history
N-nitrosamines	Helicobacter pylori infection (decreased risk)

Materials and methods

- For this present research article the TCGA (GDC, genomic data common) and PPI STRING database has been used.
- According to the TCGA there are some mutated genes related to esophageal cancer.
- By using the PPI string database a network of 575 genes mutated in esophageal cancer in 3 and 5 clusters has been created.
- The Cancer Genome Atlas, a landmark cancer genomics program, molecularly characterized over 20,000 primary cancer and matched normal samples 33 cancer types.
- This joint effort between NCI and the National human Genome Research Institute began in 2006.
- TCGA generated over 2.5 petabytes of genomic, epigenomic, transcriptomics and proteomics data.

- This data which already led to improve in our ability to diagnose, treat and prevent cancer.
- Gene regulatory network is a set of genes, or parts of genes, that interact with each other to control a specific cell function.
- These networks are important in development, differentiation and responding to environmental cues.
- They are made of a few highly connected nodes and many poorly connected nodes nested within a hierarchical regulatory regime.

Results and discussions

According to the TCGA, there are genetically altered genes linked to esophageal cancer. Below is a network of 575 genetically altered genes linked to constipation, and there is a genetic connection in three groups as well as five groups. A network of genes, or a network of genes, that work together to regulate a certain cell's function is known as a network that controls genes. Gene control networks play a crucial role in the creation, categorization, and application of environmental regulations. Several highly linked nodes and several unconnected nodes are often assumed to make up gene control networks, which are entrenched in states of high control. Therefore, genetic control networks gauge the pace of development of free-level networks. This structure is thought to be mutable due to the selective attachment of highly intertwined genes.

A renowned cancer program known as The Cancer Genome Atlas (TCGA) is distinguished by more than 20,000 primary cancer cells and general samples made up of 33 different forms of cancer. Researchers from other areas and universities have joined the NCI and National Human Genome Research Institute in this collaborative initiative, which started in 2006.

More than 2.5 petabytes of genomic, epigenomic, transcriptomic, and proteomic data were generated by TCGA over the course of the following twelve years. Everyone in the scientific community who can benefit from the knowledge, which has already contributed to the advancement of our capacity to identify, treat, and prevent cancer, will always have access to it.

In the TCGA category of GDC (general genomic data) there are 68 projects, 67 basic sites, 84609 cases, 23587 genes and 35587082 mutations available, which 2487 belongs to esophageal cancer. Among these programs, there are existing cancer projects. I did research on Esophagus cancer using the TCGA database.

I have found 180 cases of esophagus cancer, and 575 genes have been changed. The genes below are genetically modified genes.

Evaluation of gene regulatory network

There are primarily two ways that networks can evolve, first way by simultaneously. The first is that network topology can be changed by the addition or subtraction of nodes (genes) or parts of the network (modules) may be expressed in different contexts. The second way networks can evolve is by changing the strength of interactions between nodes, such as how strongly a transcription factor may bind to a cis- regulatory element.

Table 2: List of 575 mutated genes in esophagus cancer

No	Gene ID	Symbol	Name	SSM Affected Cases in Cohort	Mutations
1	ENSG00000141510	TP53	tumor protein p53	157 / 180 (87.22%)	123
2	ENSG00000167548	KMT2D	lysine (K)-specific methyltransferase 2D	29 / 180 (16.11%)	34
3	ENSG00000116044	NFE2L2	nuclear factor, erythroid 2-like 2	22 / 180 (12.22%)	22
4	ENSG00000196159	FAT4	FAT atypical cadherin 4	21 / 180 (11.67%)	25
5	ENSG00000178568	ERBB4	erb-b2 receptor tyrosine kinase 4	19 / 180 (10.56%)	20
6	ENSG00000055609	KMT2C	lysine (K)-specific methyltransferase 2C	19 / 180 (10.56%)	20
7	ENSG00000117713	ARID1A	AT rich interactive domain 1A (SWI-like)	19 / 180 (10.56%)	19
8	ENSG00000148400	NOTCH1	notch 1	18 / 180 (10.00%)	22
9	ENSG00000147889	CDKN2A	cyclin-dependent kinase inhibitor 2A	17 / 180 (9.44%)	16
10	ENSG00000140937	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)	17 / 180 (9.44%)	20
11	ENSG00000121879	PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha	16 / 180 (8.89%)	10
12	ENSG00000157168	NRG1	neuregulin 1	16 / 180 (8.89%)	17
13	ENSG00000087460	GNAS	GNAS complex locus	16 / 180 (8.89%)	16
14	ENSG00000198795	ZNF521	zinc finger protein 521	14 / 180 (7.78%)	15
15	ENSG00000196367	TRRAP	transformation/transcription domain-associated protein	14 / 180 (7.78%)	16
16	ENSG00000141646	SMAD4	SMAD family member 4	13 / 180 (7.22%)	13
17	ENSG00000127616	SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	13 / 180 (7.22%)	12
18	ENSG00000196712	NF1	neurofibromin 1	13 / 180 (7.22%)	11
19	ENSG00000005339	CREBBP	CREB binding protein	12 / 180 (6.67%)	13
20	ENSG00000152217	SETBP1	SET binding protein 1	12 / 180 (6.67%)	12
21	ENSG00000177084	POLE	polymerase (DNA directed), epsilon, catalytic subunit	12 / 180 (6.67%)	13
22	ENSG00000145675	PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)	12 / 180 (6.67%)	12
23	ENSG00000079102	RUNX1T1	runt-related transcription factor 1; translocated to, 1 (cyclin D-related)	12 / 180 (6.67%)	13
24	ENSG00000109670	FBXW7	F-box and WD repeat domain containing 7, E3 ubiquitin protein ligase	12 / 180 (6.67%)	13
25	ENSG00000100393	EP300	E1A binding protein p300	12 / 180 (6.67%)	12
26	ENSG00000181690	PLAG1	pleiomorphic adenoma gene 1	11 / 180 (6.11%)	14
27	ENSG00000118971	CCND2	cyclin D2	11 / 180 (6.11%)	9
28	ENSG00000163513	TGFBR2	transforming growth factor, beta receptor II (70/80kDa)	11 / 180 (6.11%)	11
29	ENSG00000134982	APC	adenomatous polyposis coli	11 / 180 (6.11%)	10
30	ENSG00000049618	ARID1B	AT rich interactive domain 1B (SWI1-like)	11 / 180 (6.11%)	12
31	ENSG00000171862	PTEN	phosphatase and tensin homolog	11 / 180 (6.11%)	15
32	ENSG00000144218	AFF3	AF4/FMR2 family, member 3	11 / 180 (6.11%)	12
33	ENSG00000073614	KDM5A	lysine (K)-specific demethylase 5A	11 / 180 (6.11%)	10
34	ENSG00000173821	RNF213	ring finger protein 213	11 / 180 (6.11%)	14
35	ENSG00000171456	ASXL1	additional sex combs like transcriptional regulator 1	10 / 180 (5.56%)	12
36	ENSG00000127152	BCL11B	B-cell CLL/lymphoma 11B (zinc finger protein)	10 / 180 (5.56%)	15
37	ENSG00000083857	FAT1	FAT atypical cadherin 1	10 / 180 (5.56%)	10
38	ENSG00000119866	BCL11A	B-cell CLL/lymphoma 11A (zinc finger protein)	10 / 180 (5.56%)	10
39	ENSG00000141736	ERBB2	erb-b2 receptor tyrosine kinase 2	10 / 180 (5.56%)	9
40	ENSG00000140836	ZFHX3	zinc finger homeobox 3	10 / 180 (5.56%)	15
41	ENSG00000140538	NTRK3	neurotrophic tyrosine kinase, receptor, type 3	10 / 180 (5.56%)	10
42	ENSG00000127914	AKAP9	A kinase (PRKA) anchor protein 9	10 / 180 (5.56%)	14
43	ENSG00000143924	EML4	echinoderm microtubule associated protein like 4	9 / 180 (5.00%)	11
44	ENSG00000164754	RAD21	RAD21 homolog (S. pombe)	9 / 180 (5.00%)	8
45	ENSG00000165671	NSD1	nuclear receptor binding SET domain protein 1	9 / 180 (5.00%)	12
46	ENSG00000177565	TBL1XR1	transducin (beta)-like 1 X-linked receptor 1	9 / 180 (5.00%)	13
47	ENSG00000118058	KMT2A	lysine (K)-specific methyltransferase 2A	9 / 180 (5.00%)	10
48	ENSG00000178573	MAF	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog	9 / 180 (5.00%)	10
49	ENSG00000103197	TSC2	tuberous sclerosis 2	9 / 180 (5.00%)	9
50	ENSG00000184937	WT1	Wilms tumor 1	9 / 180 (5.00%)	9
51	ENSG00000196498	NCOR2	nuclear receptor corepressor 2	9 / 180 (5.00%)	11
52	ENSG00000185920	PTCH1	patched 1	9 / 180 (5.00%)	11
53	ENSG00000119508	NR4A3	nuclear receptor subfamily 4, group A, member 3	9 / 180 (5.00%)	8
54	ENSG00000196220	SRGAP3	SLIT-ROBO Rho GTPase activating protein 3	9 / 180 (5.00%)	9
55	ENSG00000108375	RNF43	ring finger protein 43	9 / 180 (5.00%)	9
56	ENSG00000165731	RET	ret proto-oncogene	9 / 180 (5.00%)	10
57	ENSG00000168769	TET2	tet methylcytosine dioxygenase 2	9 / 180 (5.00%)	15
58	ENSG00000175054	ATR	ATR serine/threonine kinase	9 / 180 (5.00%)	10
59	ENSG00000095015	MAP3K1	mitogen-activated protein kinase kinase kinase 1, E3 ubiquitin protein ligase	9 / 180 (5.00%)	9
60	ENSG00000141867	BRD4	bromodomain containing 4	8 / 180 (4.44%)	8
61	ENSG00000160613	PCSK7	proprotein convertase subtilisin/kexin type 7	8 / 180 (4.44%)	8
62	ENSG00000116062	MSH6	mutS homolog 6	8 / 180 (4.44%)	8
63	ENSG00000119335	SET	SET nuclear proto-oncogene	8 / 180 (4.44%)	8
64	ENSG00000100503	NIN	ninein (GSK3B interacting protein)	8 / 180 (4.44%)	8
65	ENSG00000129204	USP6	ubiquitin specific peptidase 6	8 / 180 (4.44%)	8
66	ENSG00000114861	FOXP1	forkhead box P1	8 / 180 (4.44%)	9
67	ENSG00000099949	LZTR1	leucine-zipper-like transcription regulator 1	8 / 180 (4.44%)	8
68	ENSG00000057657	PRDM1	PR domain containing 1, with ZNF domain	8 / 180 (4.44%)	8
69	ENSG00000122025	FLT3	fms-related tyrosine kinase 3	8 / 180 (4.44%)	8
70	ENSG00000184384	MAML2	mastermind-like 2 (Drosophila)	8 / 180 (4.44%)	8
71	ENSG00000116128	BCL9	B-cell CLL/lymphoma 9	8 / 180 (4.44%)	9
72	ENSG00000147050	KDM6A	lysine (K)-specific demethylase 6A	8 / 180 (4.44%)	8
73	ENSG00000171735	CAMTA1	calmodulin binding transcription activator 1	8 / 180 (4.44%)	10
74	ENSG00000149948	HMGA2	high mobility group AT-hook 2	8 / 180 (4.44%)	9
75	ENSG00000122566	HNRNPA2B1	heterogeneous nuclear ribonucleoprotein A2/B1	8 / 180 (4.44%)	8
76	ENSG00000007237	GAS7	growth arrest-specific 7	8 / 180 (4.44%)	8
77	ENSG00000171094	ALK	anaplastic lymphoma receptor tyrosine kinase	7 / 180 (3.89%)	7
78	ENSG00000204764	RANBP17	RAN binding protein 17	7 / 180 (3.89%)	7
79	ENSG00000163629	PTPN13	protein tyrosine phosphatase, non-receptor type 13 (APO-1/CD95 (Fas)-associated phosphatase)	7 / 180 (3.89%)	8
80	ENSG00000126777	KTN1	kinectin 1 (kinesin receptor)	7 / 180 (3.89%)	7

No	Gene ID	Symbol	Name	SSM Affected Cases in Cohort	Mutations
161	ENSG00000151532	VTI1A	vesicle transport through interaction with t-SNAREs 1A	5 / 180 (2.78%)	6
162	ENSG00000136352	NKX2-1	NK2 homeobox 1	5 / 180 (2.78%)	7
163	ENSG00000100644	HIF1A	hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)	5 / 180 (2.78%)	7
164	ENSG00000065526	SPEN	spen family transcriptional repressor	5 / 180 (2.78%)	5
165	ENSG00000118689	FOXO3	forkhead box O3	5 / 180 (2.78%)	5
166	ENSG00000145012	LPP	LIM domain containing preferred translocation partner in lipoma	5 / 180 (2.78%)	5
167	ENSG00000177606	JUN	jun proto-oncogene	5 / 180 (2.78%)	5
168	ENSG00000196531	NACA	nascent polypeptide-associated complex alpha subunit	5 / 180 (2.78%)	5
169	ENSG00000110395	CBL	Cbl proto-oncogene, E3 ubiquitin protein ligase	5 / 180 (2.78%)	5
170	ENSG00000037280	FLT4	fms-related tyrosine kinase 4	5 / 180 (2.78%)	7
171	ENSG00000109906	ZBTB16	zinc finger and BTB domain containing 16	5 / 180 (2.78%)	6
172	ENSG00000170759	KIF5B	kinesin family member 5B	5 / 180 (2.78%)	5
173	ENSG00000198793	MTOR	mechanistic target of rapamycin (serine/threonine kinase)	5 / 180 (2.78%)	5
174	ENSG00000136238	RAC1	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	5 / 180 (2.78%)	5
175	ENSG00000187741	FANCA	Fanconi anemia, complementation group A	5 / 180 (2.78%)	5
176	ENSG00000165392	WRN	Werner syndrome, RecQ helicase-like	5 / 180 (2.78%)	6
177	ENSG00000156531	PHF6	PHD finger protein 6	5 / 180 (2.78%)	5
178	ENSG00000105568	PPP2R1A	protein phosphatase 2, regulatory subunit A, alpha	5 / 180 (2.78%)	5
179	ENSG00000105173	CCNE1	cyclin E1	5 / 180 (2.78%)	5
180	ENSG00000180644	PRF1	perforin 1 (pore forming protein)	4 / 180 (2.22%)	4
181	ENSG00000107779	BMPR1A	bone morphogenetic protein receptor, type IA	4 / 180 (2.22%)	4
182	ENSG00000198400	NTRK1	neurotrophic tyrosine kinase, receptor, type 1	4 / 180 (2.22%)	4
183	ENSG00000196914	ARHGEF12	Rho guanine nucleotide exchange factor (GEF) 12	4 / 180 (2.22%)	4
184	ENSG00000163399	ATP1A1	ATPase, Na+/K+ transporting, alpha 1 polypeptide	4 / 180 (2.22%)	5
185	ENSG00000167460	TPM4	tropomyosin 4	4 / 180 (2.22%)	5
186	ENSG00000078674	PCM1	pericentriolar material 1	4 / 180 (2.22%)	4
187	ENSG00000181449	SOX2	SRY (sex determining region Y)-box 2	4 / 180 (2.22%)	3
188	ENSG00000182197	EXT1	exostosin glycosyltransferase 1	4 / 180 (2.22%)	4
189	ENSG00000121966	CXCR4	chemokine (C-X-C motif) receptor 4	4 / 180 (2.22%)	4
190	ENSG00000166949	SMAD3	SMAD family member 3	4 / 180 (2.22%)	4
191	ENSG00000136492	BRIP1	BRCA1 interacting protein C-terminal helicase 1	4 / 180 (2.22%)	4
192	ENSG00000178053	MLF1	myeloid leukemia factor 1	4 / 180 (2.22%)	4
193	ENSG00000133392	MYH11	myosin, heavy chain 11, smooth muscle	4 / 180 (2.22%)	4
194	ENSG00000078403	MLLT10	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 10	4 / 180 (2.22%)	4
195	ENSG00000182162	P2RY8	purinergic receptor P2Y, G-protein coupled, 8	4 / 180 (2.22%)	4
196	ENSG00000175832	ETV4	ets variant 4	4 / 180 (2.22%)	4
197	ENSG00000181555	SETD2	SET domain containing 2	4 / 180 (2.22%)	4
198	ENSG00000111252	SH2B3	SH2B adaptor protein 3	4 / 180 (2.22%)	4
199	ENSG00000182872	RBM10	RNA binding motif protein 10	4 / 180 (2.22%)	4
200	ENSG00000183508	FAM46C	family with sequence similarity 46, member C	4 / 180 (2.22%)	4
201	ENSG00000168610	STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	4 / 180 (2.22%)	4
202	ENSG00000198900	TOP1	topoisomerase (DNA) I	4 / 180 (2.22%)	5
203	ENSG00000119772	DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha	4 / 180 (2.22%)	4
204	ENSG00000105221	AKT2	v-akt murine thymoma viral oncogene homolog 2	4 / 180 (2.22%)	4
205	ENSG00000067842	ATP2B3	ATPase, Ca++ transporting, plasma membrane 3	4 / 180 (2.22%)	5
206	ENSG00000171791	BCL2	B-cell CLL/lymphoma 2	4 / 180 (2.22%)	4
207	ENSG00000127329	PTPRB	protein tyrosine phosphatase, receptor type, B	4 / 180 (2.22%)	4
208	ENSG00000179583	CIITA	class II, major histocompatibility complex, transactivator	4 / 180 (2.22%)	4
209	ENSG00000197157	SND1	staphylococcal nuclease and tudor domain containing 1	4 / 180 (2.22%)	4
210	ENSG00000085224	ATRX	alpha thalassemia/mental retardation syndrome X-linked	4 / 180 (2.22%)	4
211	ENSG00000134352	IL6ST	interleukin 6 signal transducer	4 / 180 (2.22%)	4
212	ENSG00000172493	AFF1	AF4/FMR2 family, member 1	4 / 180 (2.22%)	4
213	ENSG00000128513	POT1	protection of telomeres 1	4 / 180 (2.22%)	4
214	ENSG00000160271	RALGDS	ral guanine nucleotide dissociation stimulator	4 / 180 (2.22%)	3
215	ENSG00000139219	COL2A1	collagen, type II, alpha 1	4 / 180 (2.22%)	4
216	ENSG00000110619	CARS	cysteinyl-tRNA synthetase	4 / 180 (2.22%)	4
217	ENSG00000138698	RAP1GDS1	RAP1, GTP-GDP dissociation stimulator 1	4 / 180 (2.22%)	4
218	ENSG00000127946	HIP1	huntingtin interacting protein 1	4 / 180 (2.22%)	4
219	ENSG00000245848	CEBPA	CCAAT/enhancer binding protein (C/EBP), alpha	4 / 180 (2.22%)	4
220	ENSG00000206503	HLA-A	major histocompatibility complex, class I, A	4 / 180 (2.22%)	4
221	ENSG00000110987	BCL7A	B-cell CLL/lymphoma 7A	4 / 180 (2.22%)	5
222	ENSG00000147065	MSN	moesin	4 / 180 (2.22%)	4
223	ENSG00000104320	NBN	nibrin	4 / 180 (2.22%)	4
224	ENSG00000204843	DCTN1	dynactin 1	4 / 180 (2.22%)	4
225	ENSG00000071564	TCF3	transcription factor 3	4 / 180 (2.22%)	4
226	ENSG00000185630	PBX1	pre-B-cell leukemia homeobox 1	4 / 180 (2.22%)	4
227	ENSG00000147548	WHSC1L1	Wolf-Hirschhorn syndrome candidate 1-like 1	4 / 180 (2.22%)	8
228	ENSG00000187098	MITF	microphthalmia-associated transcription factor	4 / 180 (2.22%)	4
229	ENSG00000081237	PTPRC	protein tyrosine phosphatase, receptor type, C	4 / 180 (2.22%)	4
230	ENSG00000106031	HOXA13	homeobox A13	4 / 180 (2.22%)	2
231	ENSG00000133639	BTG1	B-cell translocation gene 1, anti-proliferative	4 / 180 (2.22%)	4
232	ENSG00000085832	EPS15	epidermal growth factor receptor pathway substrate 15	4 / 180 (2.22%)	4
233	ENSG00000163518	FCRL4	Fc receptor-like 4	4 / 180 (2.22%)	4
234	ENSG00000089280	FUS	FUS RNA binding protein	4 / 180 (2.22%)	4
235	ENSG00000157388	CACNA1D	calcium channel, voltage-dependent, L type, alpha 1D subunit	4 / 180 (2.22%)	4
236	ENSG00000183765	CHEK2	checkpoint kinase 2	4 / 180 (2.22%)	3
237	ENSG00000047410	TPR	translocated promoter region, nuclear basket protein	4 / 180 (2.22%)	4
238	ENSG00000066117	SMARCD1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1	4 / 180 (2.22%)	4
239	ENSG00000132475	H3F3B	H3 histone, family 3B (H3.3B)	4 / 180 (2.22%)	4
240	ENSG00000184402	SS18L1	synovial sarcoma translocation gene on chromosome 18-like 1	3 / 180 (1.67%)	3

No	Gene ID	Symbol	Name	SSM Affected Cases in Cohort	Mutations
241	ENSG00000097007	ABL1	ABL proto-oncogene 1, non-receptor tyrosine kinase	3 / 180 (1.67%)	3
242	ENSG00000122779	TRIM24	tripartite motif containing 24	3 / 180 (1.67%)	3
243	ENSG00000166886	NAB2	NGFI-A binding protein 2 (EGR1 binding protein 2)	3 / 180 (1.67%)	3
244	ENSG00000156970	BUB1B	BUB1 mitotic checkpoint serine/threonine kinase B	3 / 180 (1.67%)	3
245	ENSG00000116251	RPL22	ribosomal protein L22	3 / 180 (1.67%)	1
246	ENSG00000064933	PMS1	PMS1 postmeiotic segregation increased 1 (S. cerevisiae)	3 / 180 (1.67%)	3
247	ENSG00000112576	CCND3	cyclin D3	3 / 180 (1.67%)	3
248	ENSG00000204103	MAFB	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog B	3 / 180 (1.67%)	3
249	ENSG00000136997	MYC	v-myc avian myelocytomatosis viral oncogene homolog	3 / 180 (1.67%)	4
250	ENSG00000169032	MAP2K1	mitogen-activated protein kinase kinase 1	3 / 180 (1.67%)	3
251	ENSG00000049540	ELN	elastin	3 / 180 (1.67%)	3
252	ENSG00000084676	NCOA1	nuclear receptor coactivator 1	3 / 180 (1.67%)	3
253	ENSG00000184507	NUTM1	NUT midline carcinoma, family member 1	3 / 180 (1.67%)	3
254	ENSG00000160789	LMNA	lamin A/C	3 / 180 (1.67%)	3
255	ENSG00000159216	RUNX1	runt-related transcription factor 1	3 / 180 (1.67%)	3
256	ENSG00000171723	GPHN	gephyrin	3 / 180 (1.67%)	3
257	ENSG00000064012	CASP8	caspase 8, apoptosis-related cysteine peptidase	3 / 180 (1.67%)	4
258	ENSG00000110092	CCND1	cyclin D1	3 / 180 (1.67%)	3
259	ENSG00000080824	HSP90AA1	heat shock protein 90kDa alpha (cytosolic), class A member 1	3 / 180 (1.67%)	3
260	ENSG00000065361	ERBB3	erb-b2 receptor tyrosine kinase 3	3 / 180 (1.67%)	3
261	ENSG00000115524	SF3B1	splicing factor 3b, subunit 1, 155kDa	3 / 180 (1.67%)	4
262	ENSG00000132170	PPARG	peroxisome proliferator-activated receptor gamma	3 / 180 (1.67%)	3
263	ENSG00000168646	AXIN2	axin 2	3 / 180 (1.67%)	3
264	ENSG00000116560	SFPQ	splicing factor proline/glutamine-rich	3 / 180 (1.67%)	3
265	ENSG00000170234	PWWP2A	PWWP domain containing 2A	3 / 180 (1.67%)	3
266	ENSG00000167751	KLK2	kallikrein-related peptidase 2	3 / 180 (1.67%)	3
267	ENSG00000066455	GOLGA5	golgin A5	3 / 180 (1.67%)	3
268	ENSG00000123268	ATF1	activating transcription factor 1	3 / 180 (1.67%)	3
269	ENSG00000108654	DDX5	DEAD (Asp-Glu-Ala-Asp) box helicase 5	3 / 180 (1.67%)	3
270	ENSG00000125952	MAX	MYC associated factor X	3 / 180 (1.67%)	3
271	ENSG00000141380	SS18	synovial sarcoma translocation, chromosome 18	3 / 180 (1.67%)	3
272	ENSG00000130675	MNX1	motor neuron and pancreas homeobox 1	3 / 180 (1.67%)	3
273	ENSG00000100311	PDGFB	platelet-derived growth factor beta polypeptide	3 / 180 (1.67%)	5
274	ENSG00000157873	TNFRSF14	tumor necrosis factor receptor superfamily, member 14	3 / 180 (1.67%)	3
275	ENSG00000169184	MN1	meningioma (disrupted in balanced translocation) 1	3 / 180 (1.67%)	3
276	ENSG00000128052	KDR	kinase insert domain receptor	3 / 180 (1.67%)	3
277	ENSG00000133703	KRAS	Kirsten rat sarcoma viral oncogene homolog	3 / 180 (1.67%)	2
278	ENSG00000159784	FAM131B	family with sequence similarity 131, member B	3 / 180 (1.67%)	3
279	ENSG00000135903	PAX3	paired box 3	3 / 180 (1.67%)	3
280	ENSG00000077782	FGFR1	fibroblast growth factor receptor 1	3 / 180 (1.67%)	3
281	ENSG00000072274	TFRC	transferrin receptor	3 / 180 (1.67%)	3
282	ENSG00000157613	CREB3L1	cAMP responsive element binding protein 3-like 1	3 / 180 (1.67%)	3
283	ENSG00000141367	CLTC	clathrin, heavy chain (Hc)	3 / 180 (1.67%)	3
284	ENSG00000143437	ARNT	aryl hydrocarbon receptor nuclear translocator	3 / 180 (1.67%)	3
285	ENSG00000135679	MDM2	MDM2 proto-oncogene, E3 ubiquitin protein ligase	3 / 180 (1.67%)	3
286	ENSG00000135100	HNF1A	HNF1 homeobox A	3 / 180 (1.67%)	4
287	ENSG00000136167	LCP1	lymphocyte cytosolic protein 1 (L-plastin)	3 / 180 (1.67%)	3
288	ENSG00000160957	RECQL4	RecQ protein-like 4	3 / 180 (1.67%)	3
289	ENSG00000129152	MYOD1	myogenic differentiation 1	3 / 180 (1.67%)	3
290	ENSG00000129514	FOXA1	forkhead box A1	3 / 180 (1.67%)	3
291	ENSG00000108821	COL1A1	collagen, type I, alpha 1	3 / 180 (1.67%)	3
292	ENSG00000171843	MLLT3	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 3	3 / 180 (1.67%)	5
293	ENSG00000130382	MLLT1	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 1	3 / 180 (1.67%)	3
294	ENSG00000006468	ETV1	ets variant 1	3 / 180 (1.67%)	3
295	ENSG00000169083	AR	androgen receptor	3 / 180 (1.67%)	3
296	ENSG00000169696	ASPSCR1	alveolar soft part sarcoma chromosome region, candidate 1	3 / 180 (1.67%)	3
297	ENSG00000104408	EIF3E	eukaryotic translation initiation factor 3, subunit E	3 / 180 (1.67%)	3
298	ENSG00000105976	MET	MET proto-oncogene, receptor tyrosine kinase	3 / 180 (1.67%)	3
299	ENSG00000164438	TLX3	T-cell leukemia homeobox 3	3 / 180 (1.67%)	3
300	ENSG00000082898	XPO1	exportin 1	3 / 180 (1.67%)	4
301	ENSG00000072062	PRKACA	protein kinase, cAMP-dependent, catalytic, alpha	3 / 180 (1.67%)	3
302	ENSG00000162434	JAK1	Janus kinase 1	3 / 180 (1.67%)	3
303	ENSG00000163655	GMPS	guanine monphosphate synthase	3 / 180 (1.67%)	3
304	ENSG00000112561	TFEB	transcription factor EB	3 / 180 (1.67%)	3
305	ENSG00000168685	IL7R	interleukin 7 receptor	3 / 180 (1.67%)	3
306	ENSG00000070371	CLTCL1	clathrin, heavy chain-like 1	3 / 180 (1.67%)	4
307	ENSG00000116990	MYCL	v-myc avian myelocytomatosis viral oncogene lung carcinoma derived homolog	3 / 180 (1.67%)	3
308	ENSG00000072364	AFF4	AF4/FMR2 family, member 4	3 / 180 (1.67%)	3
309	ENSG00000205927	OLIG2	oligodendrocyte lineage transcription factor 2	3 / 180 (1.67%)	3
310	ENSG00000105639	JAK3	Janus kinase 3	3 / 180 (1.67%)	3
311	ENSG00000197299	BLM	Bloom syndrome, RecQ helicase-like	3 / 180 (1.67%)	3
312	ENSG00000130779	CLIP1	CAP-GLY domain containing linker protein 1	3 / 180 (1.67%)	3
313	ENSG00000088038	CNOT3	CCR4-NOT transcription complex, subunit 3	3 / 180 (1.67%)	3
314	ENSG00000151702	FLI1	Fli-1 proto-oncogene, ETS transcription factor	3 / 180 (1.67%)	3
315	ENSG00000179348	GATA2	GATA binding protein 2	3 / 180 (1.67%)	3
316	ENSG00000164362	TERT	telomerase reverse transcriptase	3 / 180 (1.67%)	3
317	ENSG00000096968	JAK2	Janus kinase 2	3 / 180 (1.67%)	3
318	ENSG00000113721	PDGFRB	platelet-derived growth factor receptor, beta polypeptide	3 / 180 (1.67%)	4
319	ENSG00000165409	TSHR	thyroid stimulating hormone receptor	3 / 180 (1.67%)	3
320	ENSG00000105810	CDK6	cyclin-dependent kinase 6	3 / 180 (1.67%)	3

No	Gene ID	Symbol	Name	SSM Affected Cases in Cohort	Mutations
321	ENSG00000197323	TRIM33	tripartite motif containing 33	3 / 180 (1.67%)	4
322	ENSG00000161547	SRSF2	serine/arginine-rich splicing factor 2	3 / 180 (1.67%)	4
323	ENSG00000187239	FNBP1	formin binding protein 1	3 / 180 (1.67%)	3
324	ENSG00000124145	SDC4	syndecan 4	3 / 180 (1.67%)	3
325	ENSG00000033030	ZCCHC8	zinc finger, CCHC domain containing 8	3 / 180 (1.67%)	3
326	ENSG00000108946	PRKAR1A	protein kinase, cAMP-dependent, regulatory, type I, alpha	3 / 180 (1.67%)	4
327	ENSG00000067560	RHOA	ras homolog family member A	3 / 180 (1.67%)	3
328	ENSG00000139263	LRIG3	leucine-rich repeats and immunoglobulin-like domains 3	3 / 180 (1.67%)	3
329	ENSG00000204209	DAXX	death-domain associated protein	2 / 180 (1.11%)	2
330	ENSG00000182054	IDH2	isocitrate dehydrogenase 2 (NADP+), mitochondrial	2 / 180 (1.11%)	3
331	ENSG00000147655	RSPO2	R-spondin 2	2 / 180 (1.11%)	2
332	ENSG00000168421	RHOH	ras homolog family member H	2 / 180 (1.11%)	2
333	ENSG00000029725	RABEP1	rabaptin, RAB GTPase binding effector protein 1	2 / 180 (1.11%)	2
334	ENSG00000138081	FBXO11	F-box protein 11	2 / 180 (1.11%)	2
335	ENSG00000169714	CNBP	CCHC-type zinc finger, nucleic acid binding protein	2 / 180 (1.11%)	2
336	ENSG00000187735	TCEA1	transcription elongation factor A (SII), 1	2 / 180 (1.11%)	2
337	ENSG00000184634	MED12	mediator complex subunit 12	2 / 180 (1.11%)	2
338	ENSG00000187621	TCL6	T-cell leukemia/lymphoma 6 (non-protein coding)	2 / 180 (1.11%)	2
339	ENSG00000018408	WWTR1	WW domain containing transcription regulator 1	2 / 180 (1.11%)	2
340	ENSG00000157764	BRAF	B-Raf proto-oncogene, serine/threonine kinase	2 / 180 (1.11%)	2
341	ENSG00000167985	SDHAF2	succinate dehydrogenase complex assembly factor 2	2 / 180 (1.11%)	2
342	ENSG00000143549	TPM3	tropomyosin 3	2 / 180 (1.11%)	2
343	ENSG00000105662	CRTC1	CREB regulated transcription coactivator 1	2 / 180 (1.11%)	2
344	ENSG00000181163	NPM1	nucleophosmin (nucleolar phosphoprotein B23, numatrin)	2 / 180 (1.11%)	2
345	ENSG00000113916	BCL6	B-cell CLL/lymphoma 6	2 / 180 (1.11%)	2
346	ENSG00000073803	MAP3K13	mitogen-activated protein kinase kinase kinase 13	2 / 180 (1.11%)	2
347	ENSG00000138363	ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	2 / 180 (1.11%)	2
348	ENSG00000128487	SPECC1	sperm antigen with calponin homology and coiled-coil domains 1	2 / 180 (1.11%)	1
349	ENSG00000120457	KCNJ5	potassium channel, inwardly rectifying subfamily J, member 5	2 / 180 (1.11%)	2
350	ENSG00000172409	CLP1	cleavage and polyadenylation factor I subunit 1	2 / 180 (1.11%)	2
351	ENSG00000105369	CD79A	CD79a molecule, immunoglobulin-associated alpha	2 / 180 (1.11%)	2
352	ENSG00000109685	WHSC1	Wolf-Hirschhorn syndrome candidate 1	2 / 180 (1.11%)	2
353	ENSG00000182158	CREB3L2	cAMP responsive element binding protein 3-like 2	2 / 180 (1.11%)	3
354	ENSG00000114999	TTL	tubulin tyrosine ligase	2 / 180 (1.11%)	2
355	ENSG00000134323	MYCN	v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog	2 / 180 (1.11%)	2
356	ENSG00000009709	PAX7	paired box 7	2 / 180 (1.11%)	2
357	ENSG00000175387	SMAD2	SMAD family member 2	2 / 180 (1.11%)	2
358	ENSG00000072694	FCGR2B	Fc fragment of IgG, low affinity IIb, receptor (CD32)	2 / 180 (1.11%)	2
359	ENSG00000146374	RSPO3	R-spondin 3	2 / 180 (1.11%)	2
360	ENSG00000073578	SDHA	succinate dehydrogenase complex, subunit A, flavoprotein (Fp)	2 / 180 (1.11%)	2
361	ENSG00000158169	FANCC	Fanconi anemia, complementation group C	2 / 180 (1.11%)	2
362	ENSG00000132002	DNAJB1	DnaJ (Hsp40) homolog, subfamily B, member 1	2 / 180 (1.11%)	2
363	ENSG00000183161	FANCF	Fanconi anemia, complementation group F	2 / 180 (1.11%)	2
364	ENSG00000069399	BCL3	B-cell CLL/lymphoma 3	2 / 180 (1.11%)	2
365	ENSG00000039068	CDH1	cadherin 1, type 1, E-cadherin (epithelial)	2 / 180 (1.11%)	2
366	ENSG00000123473	STIL	SCL/TAL1 interrupting locus	2 / 180 (1.11%)	2
367	ENSG00000110841	PPFIBP1	PTPRF interacting protein, binding protein 1 (liprin beta 1)	2 / 180 (1.11%)	2
368	ENSG00000102974	CTCF	CCCTC-binding factor (zinc finger protein)	2 / 180 (1.11%)	2
369	ENSG00000126752	SSX1	synovial sarcoma, X breakpoint 1	2 / 180 (1.11%)	2
370	ENSG00000066468	FGFR2	fibroblast growth factor receptor 2	2 / 180 (1.11%)	2
371	ENSG00000135111	TBX3	T-box 3	2 / 180 (1.11%)	2
372	ENSG00000175595	ERCC4	excision repair cross-complementation group 4	2 / 180 (1.11%)	2
373	ENSG00000204370	SDHD	succinate dehydrogenase complex, subunit D, integral membrane protein	2 / 180 (1.11%)	2
374	ENSG00000147140	NONO	non-POU domain containing, octamer-binding	2 / 180 (1.11%)	2
375	ENSG00000153814	JAZF1	JAZF zinc finger 1	2 / 180 (1.11%)	2
376	ENSG00000165556	CDX2	caudal type homeobox 2	2 / 180 (1.11%)	2
377	ENSG00000162924	REL	v-rel avian reticuloendotheliosis viral oncogene homolog	2 / 180 (1.11%)	2
378	ENSG00000128602	SMO	smoothened, frizzled class receptor	2 / 180 (1.11%)	2
379	ENSG00000204531	POU5F1	POU class 5 homeobox 1	2 / 180 (1.11%)	2
380	ENSG00000203734	ECT2L	epithelial cell transforming 2 like	2 / 180 (1.11%)	2
381	ENSG00000126012	KDM5C	lysine (K)-specific demethylase 5C	2 / 180 (1.11%)	2
382	ENSG00000134371	CDC73	cell division cycle 73	2 / 180 (1.11%)	2
383	ENSG00000196092	PAX5	paired box 5	2 / 180 (1.11%)	2
384	ENSG00000102145	GATA1	GATA binding protein 1 (globin transcription factor 1)	2 / 180 (1.11%)	2
385	ENSG00000132155	RAF1	Raf-1 proto-oncogene, serine/threonine kinase	2 / 180 (1.11%)	2
386	ENSG00000096384	HSP90AB1	heat shock protein 90kDa alpha (cytosolic), class B member 1	2 / 180 (1.11%)	2
387	ENSG00000148737	TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)	2 / 180 (1.11%)	2
388	ENSG00000160867	FGFR4	fibroblast growth factor receptor 4	2 / 180 (1.11%)	2
389	ENSG00000142208	AKT1	v-akt murine thymoma viral oncogene homolog 1	2 / 180 (1.11%)	2
390	ENSG00000184702	05-Sep	septin 5	2 / 180 (1.11%)	2
391	ENSG00000118503	TNFAIP3	tumor necrosis factor, alpha-induced protein 3	2 / 180 (1.11%)	2
392	ENSG00000117400	MPL	MPL proto-oncogene, thrombopoietin receptor	2 / 180 (1.11%)	2
393	ENSG00000163041	H3F3A	H3 histone, family 3A	2 / 180 (1.11%)	2
394	ENSG00000172936	MYD88	myeloid differentiation primary response 88	2 / 180 (1.11%)	2
395	ENSG00000104365	IKBKB	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta	2 / 180 (1.11%)	2
396	ENSG00000189283	FHIT	fragile histidine triad	2 / 180 (1.11%)	2
397	ENSG00000151348	EXT2	exostosin glycosyltransferase 2	2 / 180 (1.11%)	2
398	ENSG00000134899	ERCC5	excision repair cross-complementation group 5	2 / 180 (1.11%)	2
399	ENSG00000114423	CBLB	Cbl proto-oncogene B, E3 ubiquitin protein ligase	2 / 180 (1.11%)	2
400	ENSG00000184640	09-Sep	septin 9	2 / 180 (1.11%)	2

No	Gene ID	Symbol	Name	SSM Affected Cases in Cohort	Mutations
401	ENSG00000109220	CHIC2	cysteine-rich hydrophobic domain 2	2 / 180 (1.11%)	2
402	ENSG00000137812	CASC5	cancer susceptibility candidate 5	2 / 180 (1.11%)	2
403	ENSG00000143322	ABL2	ABL proto-oncogene 2, non-receptor tyrosine kinase	2 / 180 (1.11%)	2
404	ENSG00000125354	06-Sep	septin 6	2 / 180 (1.11%)	2
405	ENSG00000103126	AXIN1	axin 1	2 / 180 (1.11%)	2
406	ENSG00000185338	SOCS1	suppressor of cytokine signaling 1	2 / 180 (1.11%)	2
407	ENSG00000078399	HOXA9	homeobox A9	2 / 180 (1.11%)	3
408	ENSG00000163161	ERCC3	excision repair cross-complementation group 3	2 / 180 (1.11%)	3
409	ENSG00000198286	CARD11	caspase recruitment domain family, member 11	2 / 180 (1.11%)	2
410	ENSG00000162613	FUBP1	far upstream element (FUSE) binding protein 1	2 / 180 (1.11%)	2
411	ENSG00000123983	ACSL3	acyl-CoA synthetase long-chain family member 3	2 / 180 (1.11%)	2
412	ENSG00000128656	CHN1	chimerin 1	2 / 180 (1.11%)	2
413	ENSG00000184675	AMER1	APC membrane recruitment protein 1	2 / 180 (1.11%)	2
414	ENSG00000107882	SUFU	suppressor of fused homolog (Drosophila)	2 / 180 (1.11%)	2
415	ENSG00000197535	MYO5A	myosin VA (heavy chain 12, myoxin)	2 / 180 (1.11%)	2
416	ENSG00000002834	LASP1	LIM and SH3 protein 1	2 / 180 (1.11%)	2
417	ENSG00000131759	RARA	retinoic acid receptor, alpha	2 / 180 (1.11%)	2
418	ENSG00000079805	DNM2	dynamin 2	2 / 180 (1.11%)	2
419	ENSG00000197646	PDCD1LG2	programmed cell death 1 ligand 2	2 / 180 (1.11%)	2
420	ENSG00000173757	STAT5B	signal transducer and activator of transcription 5B	2 / 180 (1.11%)	2
421	ENSG00000147403	RPL10	ribosomal protein L10	2 / 180 (1.11%)	3
422	ENSG00000166888	STAT6	signal transducer and activator of transcription 6, interleukin-4 induced	2 / 180 (1.11%)	3
423	ENSG00000178105	DDX10	DEAD (Asp-Glu-Ala-Asp) box polypeptide 10	1 / 180 (0.56%)	1
424	ENSG00000136936	XPA	xeroderma pigmentosum, complementation group A	1 / 180 (0.56%)	1
425	ENSG00000130844	ZNF331	zinc finger protein 331	1 / 180 (0.56%)	1
426	ENSG00000213190	MLLT11	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 11	1 / 180 (0.56%)	1
427	ENSG00000104419	NDRG1	N-myc downstream regulated 1	1 / 180 (0.56%)	1
428	ENSG00000171302	CANT1	calcium activated nucleotidase 1	1 / 180 (0.56%)	1
429	ENSG00000178691	SUZ12	SUZ12 polycomb repressive complex 2 subunit	1 / 180 (0.56%)	1
430	ENSG00000051108	HERPUD1	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1	1 / 180 (0.56%)	1
431	ENSG00000172175	MALT1	MALT1 paracaspase	1 / 180 (0.56%)	1
432	ENSG00000138413	IDH1	isocitrate dehydrogenase 1 (NADP+), soluble	1 / 180 (0.56%)	1
433	ENSG00000077150	NFKB2	nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	1 / 180 (0.56%)	1
434	ENSG00000068078	FGFR3	fibroblast growth factor receptor 3	1 / 180 (0.56%)	1
435	ENSG00000141985	SH3GL1	SH3-domain GRB2-like 1	1 / 180 (0.56%)	1
436	ENSG00000121289	CEP89	centrosomal protein 89kDa	1 / 180 (0.56%)	1
437	ENSG00000132781	MUTYH	mutY homolog	1 / 180 (0.56%)	1
438	ENSG00000136826	KLF4	Kruppel-like factor 4 (gut)	1 / 180 (0.56%)	1
439	ENSG00000019582	CD74	CD74 molecule, major histocompatibility complex, class II invariant chain	1 / 180 (0.56%)	1
440	ENSG00000111275	ALDH2	aldehyde dehydrogenase 2 family (mitochondrial)	1 / 180 (0.56%)	1
441	ENSG00000168092	PAFAH1B2	platelet-activating factor acetylhydrolase 1b, catalytic subunit 2 (30kDa)	1 / 180 (0.56%)	1
442	ENSG00000105619	TFPT	TCF3 (E2A) fusion partner (in childhood Leukemia)	1 / 180 (0.56%)	1
443	ENSG00000163930	BAP1	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)	1 / 180 (0.56%)	1
444	ENSG00000104903	LYL1	lymphoblastic leukemia associated hematopoiesis regulator 1	1 / 180 (0.56%)	1
445	ENSG00000106462	EZH2	enhancer of zeste 2 polycomb repressive complex 2 subunit	1 / 180 (0.56%)	1
446	ENSG00000154803	FLCN	folliculin	1 / 180 (0.56%)	1
447	ENSG00000123080	CDKN2C	cyclin-dependent kinase inhibitor 2C (p18, inhibits CDK4)	1 / 180 (0.56%)	1
448	ENSG00000154767	XPC	xeroderma pigmentosum, complementation group C	1 / 180 (0.56%)	1
449	ENSG00000196588	MKL1	megakaryoblastic leukemia (translocation) 1	1 / 180 (0.56%)	1
450	ENSG00000183770	FOXL2	forkhead box L2	1 / 180 (0.56%)	1
451	ENSG00000145819	ARHGAP26	Rho GTPase activating protein 26	1 / 180 (0.56%)	1
452	ENSG00000241476	SSX2	synovial sarcoma, X breakpoint 2	1 / 180 (0.56%)	1
453	ENSG00000145216	FIP1L1	factor interacting with PAPOLA and CPSF1	1 / 180 (0.56%)	1
454	ENSG00000139163	ETNK1	ethanolamine kinase 1	1 / 180 (0.56%)	1
455	ENSG00000110777	POU2AF1	POU class 2 associating factor 1	1 / 180 (0.56%)	1
456	ENSG00000135363	LMO2	LIM domain only 2 (rhombotin-like 1)	1 / 180 (0.56%)	1
457	ENSG00000213281	NRAS	neuroblastoma RAS viral (v-ras) oncogene homolog	1 / 180 (0.56%)	1
458	ENSG00000166407	LMO1	LIM domain only 1 (rhombotin 1)	1 / 180 (0.56%)	1
459	ENSG00000102034	ELF4	E74-like factor 4 (ets domain transcription factor)	1 / 180 (0.56%)	1
460	ENSG00000182866	LCK	LCK proto-oncogene, Src family tyrosine kinase	1 / 180 (0.56%)	1
461	ENSG00000047932	GOPC	golgi-associated PDZ and coiled-coil motif containing	1 / 180 (0.56%)	1
462	ENSG00000131653	TRAF7	TNF receptor-associated factor 7, E3 ubiquitin protein ligase	1 / 180 (0.56%)	1
463	ENSG00000123388	HOXC11	homeobox C11	1 / 180 (0.56%)	1
464	ENSG00000070404	FSTL3	follistatin-like 3 (secreted glycoprotein)	1 / 180 (0.56%)	1
465	ENSG00000099956	SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	1 / 180 (0.56%)	1
466	ENSG00000179218	CALR	calreticulin	1 / 180 (0.56%)	1
467	ENSG00000128714	HOXD13	homeobox D13	1 / 180 (0.56%)	1
468	ENSG00000134086	VHL	von Hippel-Lindau tumor suppressor, E3 ubiquitin protein ligase	1 / 180 (0.56%)	1
469	ENSG00000095002	MSH2	mutS homolog 2	1 / 180 (0.56%)	1
470	ENSG00000124795	DEK	DEK proto-oncogene	1 / 180 (0.56%)	1
471	ENSG00000112081	SRSF3	serine/arginine-rich splicing factor 3	1 / 180 (0.56%)	1
472	ENSG00000073584	SMARCE1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	1 / 180 (0.56%)	1
473	ENSG00000068323	TFE3	transcription factor binding to IGHM enhancer 3	1 / 180 (0.56%)	1
474	ENSG00000164919	COX6C	cytochrome c oxidase subunit VIc	1 / 180 (0.56%)	1
475	ENSG00000156052	GNAQ	guanine nucleotide binding protein (G protein), q polypeptide	1 / 180 (0.56%)	1
476	ENSG00000136754	ABI1	abl-interactor 1	1 / 180 (0.56%)	1
477	ENSG00000164398	ACSL6	acyl-CoA synthetase long-chain family member 6	1 / 180 (0.56%)	1
478	ENSG00000147257	GPC3	glypican 3	1 / 180 (0.56%)	1
479	ENSG00000126524	SBDS	Shwachman-Bodian-Diamond syndrome	1 / 180 (0.56%)	1
480	ENSG00000115170	ACVR1	activin A receptor, type I	1 / 180 (0.56%)	1

No	Gene ID	Symbol	Name	SSM Affected Cases in Cohort	Mutations
481	ENSG00000143252	SDHC	succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa	1 / 180 (0.56%)	1
482	ENSG00000127083	OMD	osteomodulin	1 / 180 (0.56%)	1
483	ENSG00000142273	CBLC	Cbl proto-oncogene C, E3 ubiquitin protein ligase	1 / 180 (0.56%)	1
484	ENSG00000213672	NCKIPSD	NCK interacting protein with SH3 domain	1 / 180 (0.56%)	1
485	ENSG00000184481	FOXO4	forkhead box O4	1 / 180 (0.56%)	1
486	ENSG00000108953	YWHAE	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon	1 / 180 (0.56%)	1
487	ENSG00000139083	ETV6	ets variant 6	1 / 180 (0.56%)	1
488	ENSG00000162367	TAL1	T-cell acute lymphocytic leukemia 1	1 / 180 (0.56%)	1
489	ENSG00000088256	GNA11	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)	1 / 180 (0.56%)	1
490	ENSG00000156076	WIF1	WNT inhibitory factor 1	1 / 180 (0.56%)	1
491	ENSG00000104884	ERCC2	excision repair cross-complementation group 2	1 / 180 (0.56%)	1
492	ENSG00000073921	PICALM	phosphatidylinositol binding clathrin assembly protein	1 / 180 (0.56%)	1
493	ENSG00000158711	ELK4	ELK4, ETS-domain protein (SRF accessory protein 1)	1 / 180 (0.56%)	1
494	ENSG00000213066	FGFR1OP	FGFR1 oncogene partner	1 / 180 (0.56%)	1
495	ENSG00000163902	RPN1	ribophorin I	1 / 180 (0.56%)	1
496	ENSG00000168172	HOOK3	hook microtubule-tethering protein 3	1 / 180 (0.56%)	1
497	ENSG00000126934	MAP2K2	mitogen-activated protein kinase kinase 2	1 / 180 (0.56%)	1
498	ENSG00000107807	TLX1	T-cell leukemia homeobox 1	1 / 180 (0.56%)	1
499	ENSG00000140577	CRTC3	CREB regulated transcription coactivator 3	1 / 180 (0.56%)	1
500	ENSG00000082805	ERC1	ELKS/RAB6-interacting/CAST family member 1	1 / 180 (0.56%)	1
501	ENSG00000067955	CBFB	core-binding factor, beta subunit	1 / 180 (0.56%)	1
502	ENSG00000108091	CCDC6	coiled-coil domain containing 6	1 / 180 (0.56%)	1
503	ENSG00000179295	PTPN11	protein tyrosine phosphatase, non-receptor type 11	1 / 180 (0.56%)	1
504	ENSG00000186051	TAL2	T-cell acute lymphocytic leukemia 2	1 / 180 (0.56%)	1
505	ENSG00000112039	FANCE	Fanconi anemia, complementation group E	1 / 180 (0.56%)	1
506	ENSG00000137309	HMGA1	high mobility group AT-hook 1	1 / 180 (0.56%)	2
507	ENSG00000105656	ELL	elongation factor RNA polymerase II	1 / 180 (0.56%)	1
508	ENSG00000128713	HOXD11	homeobox D11	1 / 180 (0.56%)	2
509	ENSG00000205755	CRLF2	cytokine receptor-like factor 2	1 / 180 (0.56%)	2
510	ENSG00000123364	HOXC13	homeobox C13	1 / 180 (0.56%)	1
511	ENSG00000126883	NUP214	nucleoporin 214kDa	1 / 180 (0.56%)	1
512	ENSG00000255292	AP002884.2	AP002884.2	0 / 180 (0.00%)	0
513	ENSG00000120217	CD274	CD274 molecule	0 / 180 (0.00%)	0
514	ENSG00000076685	NT5C2	5'-nucleotidase, cytosolic II	0 / 180 (0.00%)	0
515	ENSG00000160201	U2AF1	U2 small nuclear RNA auxiliary factor 1	0 / 180 (0.00%)	0
516	ENSG00000135446	CDK4	cyclin-dependent kinase 4	0 / 180 (0.00%)	0
517	ENSG00000146232	NFKBIE	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon	0 / 180 (0.00%)	0
518	ENSG00000175197	DDIT3	DNA-damage-inducible transcript 3	0 / 180 (0.00%)	0
519	ENSG00000111276	CDKN1B	cyclin-dependent kinase inhibitor 1B (p27, Kip1)	0 / 180 (0.00%)	0
520	ENSG00000137193	PIM1	Pim-1 proto-oncogene, serine/threonine kinase	0 / 180 (0.00%)	0
521	ENSG00000153944	MSI2	musashi RNA-binding protein 2	0 / 180 (0.00%)	0
522	ENSG00000048462	TNFRSF17	tumor necrosis factor receptor superfamily, member 17	0 / 180 (0.00%)	0
523	ENSG00000174775	HRAS	Harvey rat sarcoma viral oncogene homolog	0 / 180 (0.00%)	0
524	ENSG00000005073	HOXA11	homeobox A11	0 / 180 (0.00%)	0
525	ENSG00000115808	STRN	striatin, calmodulin binding protein	0 / 180 (0.00%)	0
526	ENSG00000122406	RPL5	ribosomal protein L5	0 / 180 (0.00%)	0
527	ENSG00000214827	MTCP1	mature T-cell proliferation 1	0 / 180 (0.00%)	0
528	ENSG00000185499	MUC1	mucin 1, cell surface associated	0 / 180 (0.00%)	0
529	ENSG00000157765	SLC34A2	solute carrier family 34 (type II sodium/phosphate cotransporter), member 2	0 / 180 (0.00%)	0
530	ENSG00000142867	BCL10	B-cell CLL/lymphoma 10	0 / 180 (0.00%)	0
531	ENSG00000007312	CD79B	CD79b molecule, immunoglobulin-associated beta	0 / 180 (0.00%)	0
532	ENSG00000067082	KLF6	Kruppel-like factor 6	0 / 180 (0.00%)	0
533	ENSG00000163497	FEV	FEV (ETS oncogene family)	0 / 180 (0.00%)	0
534	ENSG00000184923	NUTM2A	NUT family member 2A	0 / 180 (0.00%)	0
535	ENSG00000118260	CREB1	cAMP responsive element binding protein 1	0 / 180 (0.00%)	0
536	ENSG00000198625	MDM4	MDM4, p53 regulator	0 / 180 (0.00%)	0
537	ENSG00000109471	IL2	interleukin 2	0 / 180 (0.00%)	0
538	ENSG00000164683	HEY1	hes-related family bHLH transcription factor with YRPW motif 1	0 / 180 (0.00%)	0
539	ENSG00000133895	MEN1	multiple endocrine neoplasia I	0 / 180 (0.00%)	0
540	ENSG00000165025	SYK	spleen tyrosine kinase	0 / 180 (0.00%)	0
541	ENSG00000175643	RMI2	RecQ mediated genome instability 2	0 / 180 (0.00%)	0
542	ENSG00000188199	NUTM2B	NUT family member 2B	0 / 180 (0.00%)	0
543	ENSG00000023445	BIRC3	baculoviral IAP repeat containing 3	0 / 180 (0.00%)	0
544	ENSG00000134574	DDB2	damage-specific DNA binding protein 2, 48kDa	0 / 180 (0.00%)	0
545	ENSG00000026103	FAS	Fas cell surface death receptor	0 / 180 (0.00%)	0
546	ENSG00000261652	C15orf65	chromosome 15 open reading frame 65	0 / 180 (0.00%)	0
547	ENSG00000083093	PALB2	partner and localizer of BRCA2	0 / 180 (0.00%)	0
548	ENSG00000119537	KDSR	3-ketodihydrosphingosine reductase	0 / 180 (0.00%)	0
549	ENSG00000186575	NF2	neurofibromin 2 (merlin)	0 / 180 (0.00%)	0
550	ENSG00000091483	FH	fumarate hydratase	0 / 180 (0.00%)	0
551	ENSG00000170791	CHCHD7	coiled-coil-helix-coiled-coil-helix domain containing 7	0 / 180 (0.00%)	0
552	ENSG00000182185	RAD51B	RAD51 paralog B	0 / 180 (0.00%)	0
553	ENSG00000163026	C2orf44	chromosome 2 open reading frame 44	0 / 180 (0.00%)	0
554	ENSG00000054118	THRAP3	thyroid hormone receptor associated protein 3	0 / 180 (0.00%)	0
555	ENSG00000092820	EZR	ezrin	0 / 180 (0.00%)	0
556	ENSG00000119414	PPP6C	protein phosphatase 6, catalytic subunit	0 / 180 (0.00%)	0
557	ENSG00000101213	PTK6	protein tyrosine kinase 6	0 / 180 (0.00%)	0
558	ENSG00000143294	PRCC	papillary renal cell carcinoma (translocation-associated)	0 / 180 (0.00%)	0
559	ENSG00000015285	WAS	Wiskott-Aldrich syndrome	0 / 180 (0.00%)	0
560	ENSG00000144476	ACKR3	atypical chemokine receptor 3	0 / 180 (0.00%)	0

No	Gene ID	Symbol	Name	SSM Affected Cases in Cohort	Mutations
561	ENSG00000100721	TCL1A	T-cell leukemia/lymphoma 1A	0 / 180 (0.00%)	0
562	ENSG00000100814	CCNB1IP1	cyclin B1 interacting protein 1, E3 ubiquitin protein ligase	0 / 180 (0.00%)	0
563	ENSG00000221829	FANCG	Fanconi anemia, complementation group G	0 / 180 (0.00%)	0
564	ENSG00000158715	SLC45A3	solute carrier family 45, member 3	0 / 180 (0.00%)	0
565	ENSG00000138592	USP8	ubiquitin specific peptidase 8	0 / 180 (0.00%)	0
566	ENSG00000138795	LEF1	lymphoid enhancer-binding factor 1	0 / 180 (0.00%)	0
567	ENSG00000169249	ZRSR2	zinc finger (CCCH type), RNA-binding motif and serine/arginine rich 2	0 / 180 (0.00%)	0
568	ENSG00000114354	TFG	TRK-fused gene	0 / 180 (0.00%)	0
569	ENSG00000125618	PAX8	paired box 8	0 / 180 (0.00%)	0
570	ENSG00000184012	TMPRSS2	transmembrane protease, serine 2	0 / 180 (0.00%)	0
571	ENSG00000197880	MDS2	myelodysplastic syndrome 2 translocation associated	0 / 180 (0.00%)	0
572	ENSG00000103522	IL21R	interleukin 21 receptor	0 / 180 (0.00%)	0
573	ENSG00000164985	PSIP1	PC4 and SFRS1 interacting protein 1	0 / 180 (0.00%)	0
574	ENSG00000117118	SDHB	succinate dehydrogenase complex, subunit B, iron sulfur (Ip)	0 / 180 (0.00%)	0
575	ENSG00000076242	MLH1	mutL homolog 1	0 / 180 (0.00%)	0

Result of network analysis

Number of nodes: 570
Number of edges: 11511
Average node degree: 40.4
Avg. local clustering coefficient: 0.432
Expected number of edges: 4603
PPI enrichment p-value: < 1.0e-16

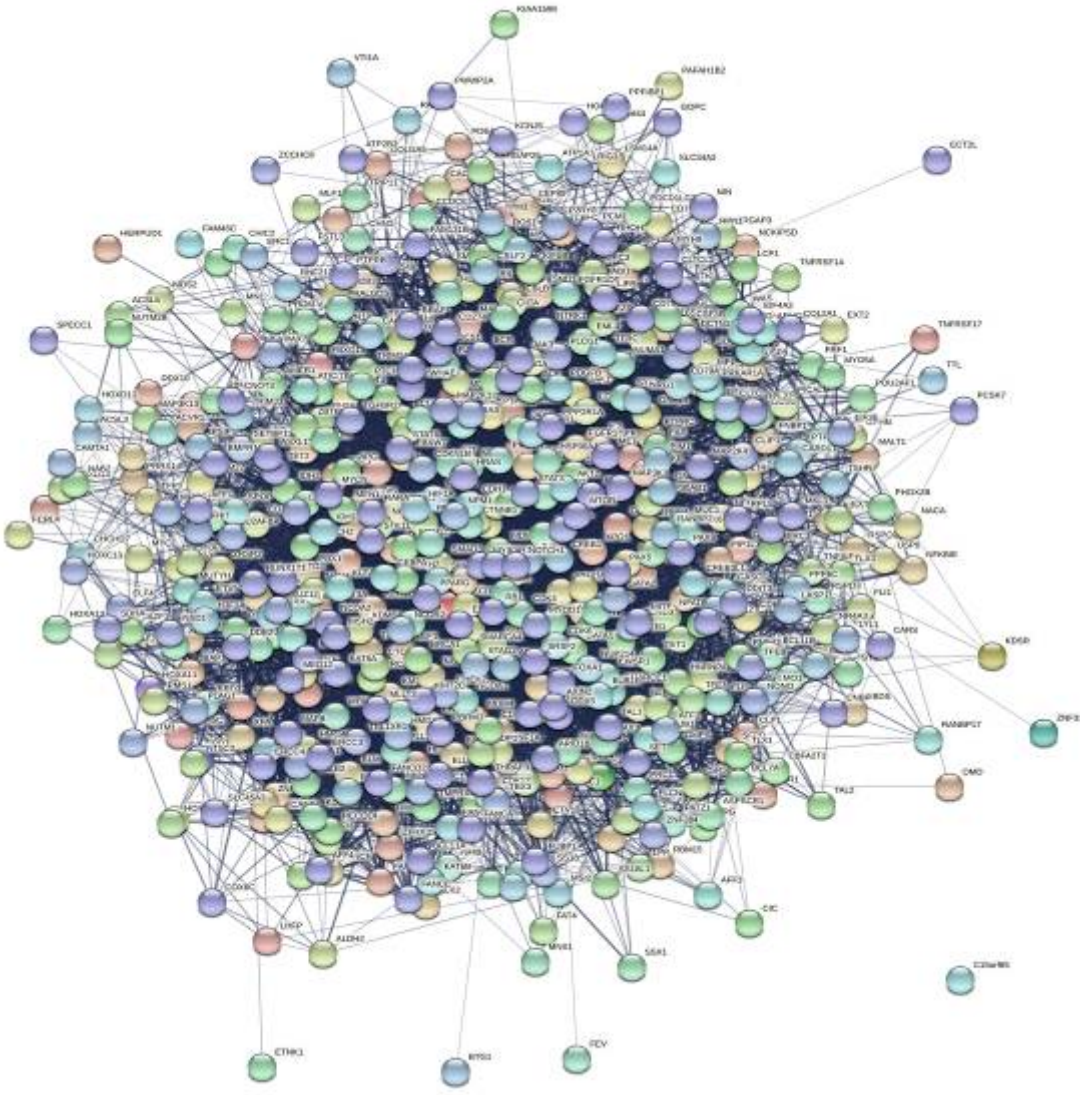

Figure 2: Network of 575 genes of esophageal cancer

The analysis section provides short statistics for the installed network, such as the number of nodes and terminals. The average node degree is the number of connections (in the point space) of the average protein in a network.

The coupling coefficient is a measure of how nodes are connected to a network.

Highly connected networks with high values.

A small PPI enrichment p indicates that the nodes are not configured and that the visible number of edges is important. Note that some of the richness situations are expected and that there are numbers that need to be translated with some caution.

- Three clusters of network using k-means clustering
- 1. Red (356 genes) 2. Green (140 genes) 3. Blue (74 genes)
- TP53 belongs to 1 cluster (Red).

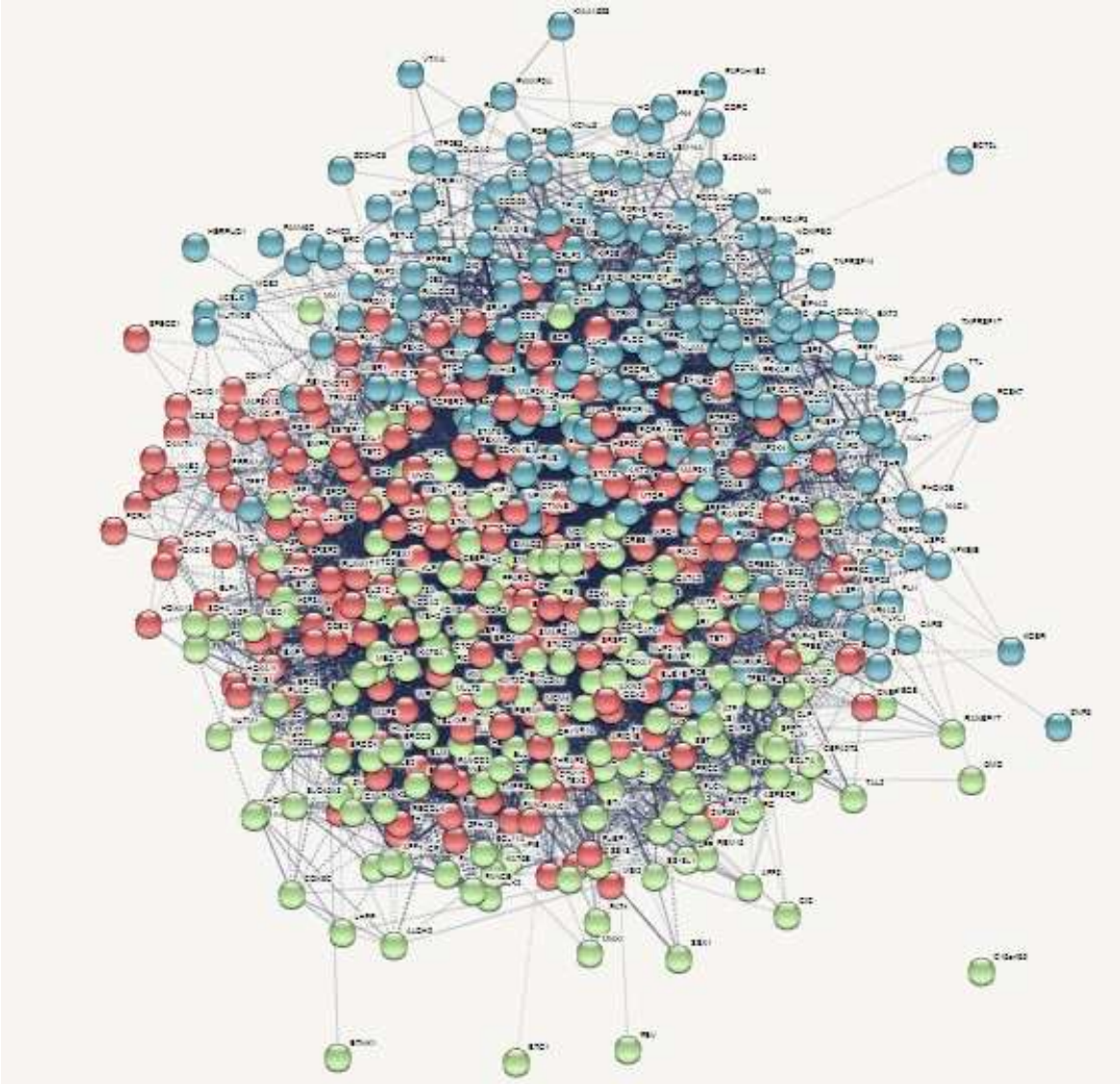

Figure 3: Three clusters of network k-means cluster

- Cluster: the organization of unlabeled data into similarity groups called clusters.
- On the other hand it is a collection of data items which are similar between them and dissimilar to data items in other clusters.
- K-means is a partitional clustering algorithm.
- The k-means algorithm partitions the given data into k clusters.
- k- is specified by the user.
- K-cluster randomly choose k data points to be the initial centroids, cluster centers.
- It assign each data point to the closest centroid
- Five clusters of network using k-means clustering
- 1. Red (181 genes)
- 2. Yellow (74 genes)
- 3. Green (70 genes)
- 4. Cyan (74 genes)
- 5.Blue (171 genes)

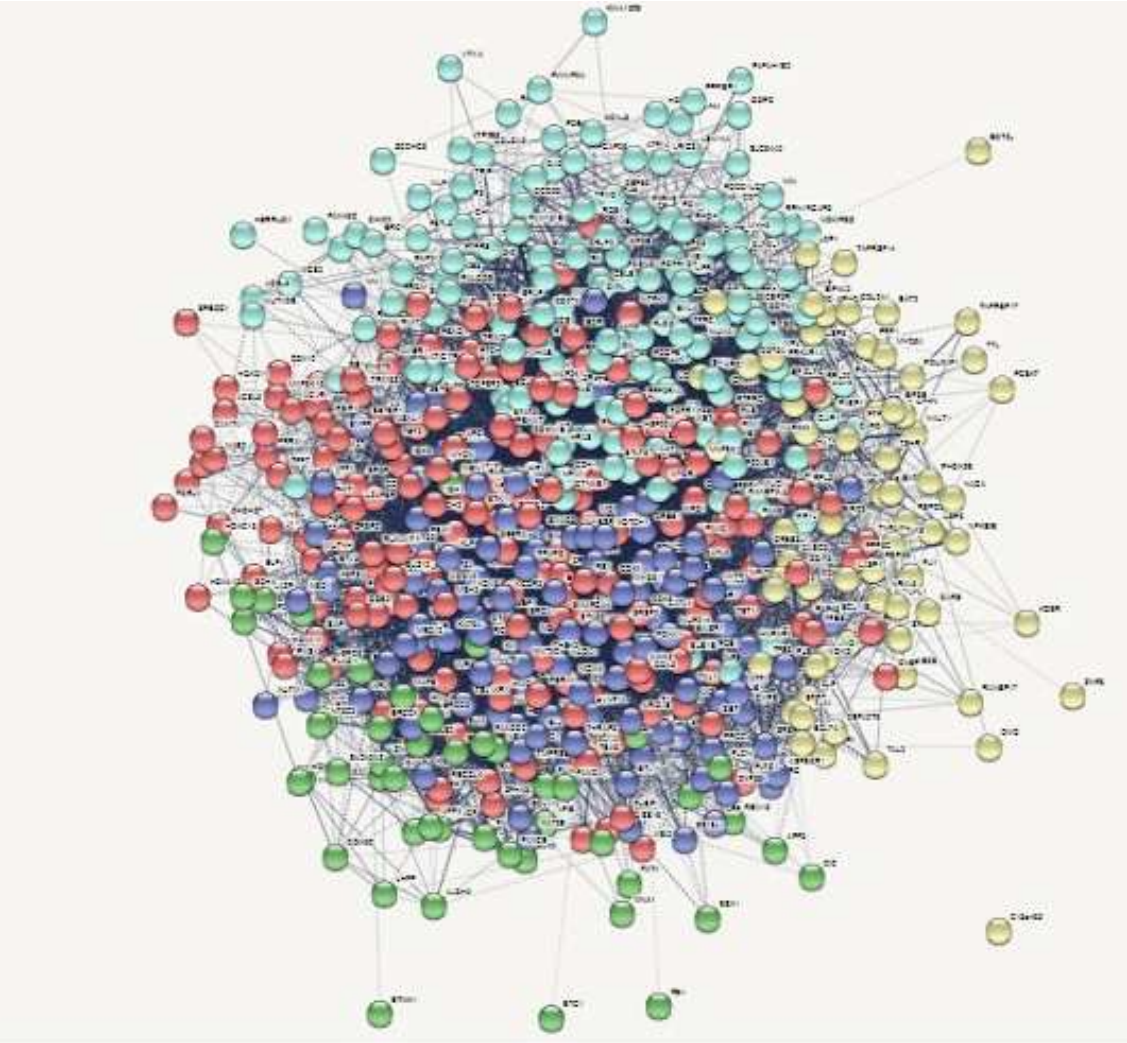

Figure 4: Five clusters of genes network k-means cluster

Conclusion

In this this research article I did some analysis related to esophageal cancer mutated genes, I used PPI STRING database and made the network of 575 genes from 180 reported studies related to esophageal cancer. It has been cleared that the esophageal cancer is the most 8th cancer in world and many people is dying every year. On the other hand men involved more than women, still there is no any evidence that why men involve than women; but as the investigation is going the most important factor is the hormone of estrogen, which it is available with female and now it's the most significant factor. Also to say that the risk factors that people involve to this cancer are: smoking, drinking alcohol and some others.

Reference

[1] Boyle, P Levin B (2008). Cancer, international Agency for Research on. World cancer report, 2008. IARC Press.

[2] Cook MB, Chow WH, Devesa SS. Oesophageal cancer incidence in the United States by race, sex, and histologic type, 1977-2005. Br J Cancer. 2009;101(5): 855-859

[3] Curado M, Edwards B, Shin H, Storm H, Ferlay J, Heanue M, et al. (2007). Cancer Incidence in Five continents, Vol. IX. IARC Scientific Publications No. 160, Lyon, IARC.

[4] Howlader N, Noone AM, Krapcho M, et al., eds.; National Cancer Institute. SEER Cancer Statistics Review, 1975-2013. Based on November 2015 SEER data submission, posted to the SEER website, April 2016. http://seer.cancer.gov/csr/1975_2013/.

[5] Lepage C, Rachet B, Jooste V, Faivre J, Coleman MP. Continuing rapid increase in esophageal adenocarcinoma in England and Wales. Am J Gastroenterol. 2008; 103(11):2694-2699.

[6] Morton LM, Gilbert ES, Stovall M, van Leeuwen FE, Dores GM, Lynch CF, Hall P, Smith SA, Weathers RE, Storm HH, Hodgson DC, Kleinerman RA, Joensuu H, Johannesen TB, Andersson M, Holowaty EJ, Kaijsar M, Pukkala E, Vaalavirta L, Fossa SD, Landmark F, Travis LB, Lamart S, Simon SL, Fraumeni JF Jr, Aleman BM, Curtis RE. Risk of esophageal cancer following radiotherapy for

Hodgkin lymphoma. Haematologica 2014; 99:e193-6.

[7] Zendehdel K, Nyrén O, Edberg A, Ye W. Risk of esophageal adenocarcinoma in achalasia patients, a retrospective cohort study in Sweden. Am J Gastroenterol. 2011; 106 (1): 57-61.

[8] Rubenstein JH, Taylor JB. Meta-analysis: the association of esophageal adenocarcinoma with symptoms of gastro-esophageal reflux. Aliment Pharmacol Ther. 2010; 32(10): 1222-1227.

[9] Xie FJ, Zhang YP, Zheng QQ, et al. Helicobacter pylori infection and esophageal cancer risk: an updated meta-analysis. World J Gastroenterology. 2013; 19(36): 6098-6107.

[10] Posner MC, Goodman KA, and Ilson DH. Ch 52 - Cancer of the Esophagus. Hellman S, Rosenberg SA, eds., Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology. 11th ed. Philadelphia, Pa: Lippincott-Williams &Wilkins; 2019.

[11] Tramacere I, La Vecchia C, Negri E. Tobacco smoking and esophageal and gastric cardiac adenocarcinoma: a meta-analysis. Epidemiology. 2011; 22(3):344-349.

[12] https://www.cancer.gov

[13] https://acsjournals.onlinelibrary.wiley.com

[14] https://www.uicc.org/news/globocan-2020-new-global-cancer-data

[15] https://www.who.int/news-room/fact-sheets/detail/cancer

[16] https://www.nih.gov/research-training

[17] https://www.cancer.org/cancer/esophagus-cancer.html