MethSRD, SNiPSRD & InteractSRD: A computational database for genes associated with dermatological disorders

A <u>Major Project</u> dissertation submitted in partial fulfilment of the requirement for the degree of

Master of Technology (Bioinformatics)

Submitted by

Mottadi Shiva
(2K15/BIO/09)
Delhi Technological University, Delhi, India

Under the supervision of

Dr. Yasha Hasija

Assistant Professor

Department of Biotechnology

Delhi Technological University, Delhi, India

CERTIFICATE

This is to certify that the M. Tech. dissertation entitled **"*MethSRD, SNiPSRD & InteractSRD: A computational database for genes associated with dermatological disorders*"**, submitted by **MOTTADI SHIVA (2K15/BIO/09)** in partial fulfilment of the requirement for the reward 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/her under my guidance.

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

Dr. Yasha Hasija   **Prof. D Kumar**

(Project Mentor)   Head of Department

Department of Biotechnology   Department of Biotechnology

Delhi Technological University   Delhi Technological University

<u>DECLARATION</u>

I, Mottadi Shiva (2K15/BIO/09) declare that M.Tech dissertation entitled **"*MethSRD, SNiPSRD & InteractSRD: A computational database for genes associated with dermatological disorders*",** submitted in partial fulfilment 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 my own work carried out under the guidance of Dr. Yasha Hasija.

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

Date: Name:

Place: Signature:

<u>ACKNOWLEDGEMENT</u>

I would like to express my sincere   and deep regards to various people who helped and supported me for the building up of this project. I would like to extend my regards and express my greatest appreciation to my guide and project mentor **Dr. Yasha Hasija** Assistant Professor, Department of Biotechnology for her belief in my abilities and her exemplary guidance at each level of the project. I am very grateful to her for giving me this opportunity to carry out such a novel and valuable work.

I also take this opportunity to express a deep sense of gratitude to Dr. D Kumar, Professor, HOD (Department of Biotechnology) for providing me the opportunity to carry out my project successfully by providing all the necessary resources and guidance.

Lastly, I would also like to thank my faculty, fellow colleagues and staff of Department of Biotechnology for helping me out when I faced problems in the project.

Mottadi Shiva
2K15/BIO/09

ABBREVIATIONS USED

DDMGD   Dragon Database for Methylated Genes and Diseases

HGNC   HUGO Gene Nomenclature Committee

SNP   Single Nucleotide Polymorphism

DBCAT   Database of CpG islands and Analytical Tool

DNMT   DNA methyl-transferases

MeSH   Medical Subject Headings

HTML   Hypertext Markup Language

CSS   Cascading Style Sheets

JSON   JavaScript Object Notation

CONTENTS

MethSRD, SNiPSRD & InteractSRD: A computational database for genes associated with dermatological disorders

Mottadi Shiva

Delhi Technological University, Delhi, India

raoshiva10@gmail.com

ABSTRACT

The most prevalent and utmost visible organ of the body is skin. Epigenetic involves in regulation of cell growth & differentiation deprived of inducing various changes in the gene sequence. The major epigenetic mechanisms are DNA methylation. Latest studies specify that inconsistent DNA methylation is a most common piece of many human disorders, including heart diseases, skin diseases, cancer and others. The reversible addition of methyl groups at CpG dinucleotides are elucidated as DNA methylation.

For various diseases to diagnose and for better treatment decisions, it is necessary to gather information that makes an association between methylated genes and diseases. As the data about the methylation, single-nucleotide polymorphism and PPI data of various genes related to dermatological disordersis scattered across a large number of electronic publications, several specialized databases, and other repositories, in the present work, we have developed a manually curated online database to enable the extraction of the detailed information related to methylation, single-nucleotide polymorphisms, and interaction of genes by contributing resources from published literature and available databases.

To the best of our knowledge, this is the first online repository containing information related to genes methylated in dermatological disorders. We believe that the database will be useful to the biomedical community for identifying novel mechanisms associated with genodermatoses.

Keywords: Epigenetics, DNA methylation, SNP, database, Genetic skin disorders. .

2. INTRODUCTION

The most prevalent and utmost visible organ of the body is skin. It reflects the health of the body & performs as a barrier against injury and bacteria. Appropriately, at one time or another, ages like infants, children, teenagers, adults and the elderly nearly everyone has some skin disease. Skin disorders differ prominently in severity and symptoms; it can be categorized by four means 1) permanent, 2) temporary, 3) painful or 4)painless. Some of the diseases may have caused situationally, else genetically. Some of the skin conditions are lightly severe, whereas some can be life threatening. The cause of the disorder is not known.

Epigenetic is a mechanisms that are involved in regulation of various differentiation & cell growth without inducing changes in the gene sequence. The major epigenetic mechanisms comprise microRNA, DNA methylation and histone modification. Recent studies had indicated that inconsistent DNA methylation is a furthermost common feature of many human disorders that includes cancer, heart diseases, skin diseases, autoimmune diseases and others. It comprised of various diseases that have a complex pathogenesis & etiology, which includes factors such as genetics and acquired under which environment and diet comes. These acquired factors repeatedly have pathogenic effects over modification of DNA and histones, out of which DNA methylation is the most common mechanism.

Aberrant DNA methylation demonstrated in skin diseases; Herein we review the role of DNA methylation in the pathogenesis of skin diseases. The study of genetic, we simply aim at studying the genotype responsible for a quantifiable phenotype, which points out the relation of a gene with the protein expressed and further causing normal or abnormal body function. Epigenetic states to the changes in the expression of the genotype, which are transmissible, but there is no change in the contributing DNA sequence affecting the change in the phenotype. The most widely studied epigenetic change is DNA Methylation Epigenetic changes are reversible and play an important role in regulating various cellular processes.

DNA methylation is an inheritable and stable epigenetic mark, which are exclusively located in the 5th position of the pyrimidine ring of cytosines (5mC) adjacent to a guanine. These sites are referred to as **CpG sites**. DNAm helps to regulate gene expression heritably, which is catalyzed by enzyme called DNA methyltransferase (DNMTs) that improves and maintains addition of a methyl group to 5′ position of the cytosine ring to form 5′ methyl-cytosine [1].

DNAm plays a significant role in the development of the disease.[2] Also, guides cellular differentiation and the manifestation for some cancers.[3] DNAm modifies mammalian genomes that occurs almost exclusively on cytosine residues that pave the way of guanine known as CpG dinucleotide. This CpG dinucleotide is generally about 60%–90% methylated. Although non-CpG cytosines are similarly methylated in certain tissues such as in embryonic stem cells[4]. However, CGIs with the highest CpG density shows the lowest levels of DNAm [5].

Frequent studies have explored the biological importance of differential methylation at CGIs [6]. As CGIs are, the regions in DNA that consist more than 500 bp that are characterized by approximately 50% of GC content. Out of which 60% of CGIs, are related with annotated gene promoters and are unmethylated permitting the transcription of the correlated gene [7]. Often, hypomethylation is linked with gene activation whereas hypermethylation is related with gene repression, although there are numerous exclusions to this trend [8]. A SNP is a sequence variation in DNA, occurs when a single nucleotide, which can be either of bases [a) A-adenine, b) T- thymine, c) G-guanine, d) C-cytosine] that differs between a members of a species or paired chromosomes in an individual. SNP fall within either coding region or non-coding regions of genes, or in the intergenic regions (regions between genes).

As the data about the methylation, single-nucleotide polymorphism and interaction data of various genes are scattered across a great number of electronic publications, several specialized databases and other repositories, because of which searching for this information manually is not an easy task. Thus, the objective of my study was to create an efficient, effective and accurate system to enable the extraction of the detailed information related to methylation, single-nucleotide polymorphism, and interaction of genes, by contributing resources from various databases.

The aim of MethSRD, SNiPSRD, and InteractSRD is to provide a freely interactive database that is accessible easily for the relationships of methylation information, human single nucleotide polymorphisms (SNiPs) and Interaction with various dermatological disorders along with supporting evidence. By doing so, my database hopes to facilitate access to, and analysis of the relationships asserted between methylation information, human single nucleotide polymorphisms (SNPs) and Interaction with the observed disease conditions.

3. **REVIEW OF LITERATURE**

3.1 Skin Disorder

The most prevalent and utmost organ of the body that is visible is skin, that acts as a barrier against many injury and bacteria and also reflects the health of the body. Unfortunately, at one time or another, nearly everybody has some skin disease ranging from the age of infants, to the elderly. The range of the skin disorders varies broadly based on the type of symptoms and severity; it can be classified by, as by duration i.e. temporary or permanent duration, painless/painful or causes can be situational/genetic.

3.2 Role of Epigenetic in Skin

The mechanisms of Epigenetic are intricate for regulating cell growth & to differentiation cell without inducing deviations of the sequence of a gene. The major of the epigenetic mechanisms embrace methylation of DNA, modification in histone & microRNA. A recent study specify that unusual DNA methylation is one of the common features of many human disorders that include cancer, autoimmune diseases, heart diseases, skin diseases, and many others. Skin diseases consist of various diseases that abstain a complex etiology & pathogenesis, including genetics and acquired factors such as diet and environment. The acquired factors regularly have pathogenic effects through the modification of DNA and histones, for which DNA methylation is the greatest mutual mechanism. Skin diseases caused due to DNA methylation include skin tumors and autoimmune-related skin disorders, thus we reviewed the role of DNA methylation in the pathogenesis for many skin diseases. [9]

3.3 Epigenetic

Epigenetics can be explained as changes take place heritably in gene expression deprived of the changes in the sequence of DNA [10]. The study of genetic modestly aims at studying the genotype accountable for a quantifiable phenotype. It points to the relation of a gene with the protein expressed and auxiliary causing normal or abnormal body function. Any genetic is frequently stable and irreversible. In other words, Epigenetic refers to the changes in the expression of the genotype, which are heritable, causing no change in the contributing DNA sequence, which hereby causes the change in the phenotype. [11]. It involves the DNA modifications and surrounds structures by either DNA methylation, chromatin modification, or

noncoding RNA. The most extensively studied epigenetic change is DNA Methylation [12]. Epigenetic changes are mostly reversible and plays an important role for regulating many cellular processes [13]

3.4 DNA Methylation

DNA methylation is an inheritable and stable epigenetic mark, which are exclusively located in the fifth position of the pyrimidine ring of cytosines (5mC) adjacent to a base of guanine, which are referred to as **CpG sites**. DNAm acts for regulating the gene expression heritably, which is catalyzed by DNA methyl-transferase (DNMTs) that adds and maintains addition of a methyl group to 5′ position of the cytosine ring to form 5′ methyl-cytosine. [14] DNAm acts as a significant role for the development of the disease. [15] Also, guides cellular differentiation and the manifestation for some cancers. [16] DNAm modification in genomes occurs exclusively on (C) cytosine residues which pave the way of guanine known as CpG dinucleotide. This CpG dinucleotide is generally about 60%–90% methylated.

Although non-CpG cytosines are similarly methylated in certain tissues such as in embryonic stem cells. [17] However highest CpG density are inverse to levels of DNAm and shows lowest density. [18] Frequent studies have explored the biological importance of differential methylation at CGIs [19]. As CGIs are, the regions of DNA, consist of more than 500bp that are at least 50% of GC content. Out of which 60% are related with promoters of annotated gene and unmethylated permitting the transcription of the correlated gene. [20] Often, hypomethylation is linked with gene activation whereas hypermethylation is linked with gene repression, although there are numerous exclusions to this trend [21].

3.5 List of Skin Disorder and their associated Gene

As the data about the Skin Disorder is scattered around a great number of electronic publications, several specialized databases, and other repositories, thus, by contributing various online as well as offline resources from various databases the list of the dermatological disorder was extracted and comprehensive repository found through text mining.

3.5.1 Genetic Skin Disorders (OUP, 2010) Book

The glossed bibliographies comprise many review papers, those from enthusiastically available journals. The bibliographies of the cited articles can direct to extensive sources.

Descriptions of histopathology, at mutually the light and electron microscopic levels, are the attributed only when the original observations are from articles of general interest. The book Genetic Skin Disorders by Virginia P. Sybert will be useful to study dermatological disorder and are useful for generalists, pediatricians, dermatologists, and geneticists. For each disorder, it includes the answers to the questions that are most often. [22]

3.5.2 Comparative Toxicogenomics Database

The URL for the database is *http://:ctd.base.org*, which is a resource, publically available for toxicogenomic information that is manually curated from the various scientific literature that provides the information for various chemicals interactions with various gene products and concludes their effect on disease. [23] It provides a detailed information about the interactions between chemicals/drugs, genes, proteins, Gene Ontology annotations, taxa, pathways and interaction modules and their relationships to diseases and its phenotype. It was curated by professional bio-curators who leverage controlled vocabularies, ontologies and structured notation to code a triad of core interactions describing gene-disease relationships [24], which are then integrated internally to generate and inferred networks between chemical, gene and disease.

3.5.3 DisGeNET: a discovery platform containing collections of genes associated with human diseases

The wide data regarding the basis genetic of human diseases deceits as the core of drug discovery precision. However, several problems, such as heterogeneity, fragmentation availability and different conceptualization of the data that must be overcome. Data for integration is mainly from various curated expert repositories, GWAS catalogs and many of the scientific literature. DisGeNET data are community-driven ontologies and homogeneously annotated with controlled vocabularies. The information can be accessible through various web domain interface such as Cytoscape App, R package, RDF SPARQL endpoint and scripts in several programming languages. It is a tremendous platform to address a variety of questions related to genetic basis of human diseases. [25, 26]

3.5.4   Genetic Association Database

A database that serves data of genetic association for various complex diseases and disorders. The main aim of the GAD is collection, standardization and archiving the genetic association study and making it easily accessible to numourous scientific community. [27]

3.6   Gene-Methylation

3.6.1   Database of CpG islands and Analytical Tool (DBCAT)

The URL for the database is *http://db.cat.c.gm.ntu.edu.tw/*, which was developed for the characterization for the profiles of DNA methylation in human cancers. It is basically an application that is web-based, comprising numerous tools (convenient) for the investigating of epigenetic regulation occurred in human diseases. It is an online analytical tool for methylation. The tool can identify the queried genes and from microarray it quickly provides the methylated regions data, and the changes in methylation status amongst plenty of arrays and functional are related. [28]. It contributes information such as Gene, Chromosome, Entrez Gene ID, RefseqID, Description, Biological Process, Molecular Function, KEGG Pathway and Genome Position for my database.

3.6.2   MethHC: a database of DNA methylation and gene expression in human cancer

The URL for the database is *http://Meth.HC.m.bc.nctu.edu.tw*, which comprises a systematic integration of a huge collection of DNA methylation. It provides information such as DNA methylation, p-value, and Transcript. [29]

3.6.3   Dragon Database for Methylated Genes and Diseases(DDMGD)

The URL for the database is *http://www.cb.rc.ka.ust.edu.sa/ddmgd/;* it provides a wide-ranging repository for the information associated with methylation of genes in diseases that can be retrieved through various text mining. The text mining system extracts associated methylated genes in different diseases from certain PubMed abstracts or PubMed Central articles. It is a user friendly interface facilitating repossession for the connotations that are ranked according to confidence scores. [30] It provides information such as Methylation, Gene Expression, Gene ID, Disease Progression and PubMed ID to the database.

3.6.4 DiseaseMeth a human disease methylation database

The URL for the database is *http://bio.info.hrb.mu.edu.cn/*; it mainly focuses on the statistical analysis & efficient storage of data sets regarding DNA methylation from various diseases. From around 14,000 entries and 175 throughput data sets from a wide number of high sources, the experimental information has been collected and incorporated into the database. It supports many search options such as by disease name and gene ID. It also provides an integrated methylation data for gene that are based on cross-data set for the analysis of disease and normal samples, that gives the detailed identification of gene–disease relationship.[31] It provides information such as Gene Name, Chromosome No., Genome Position Start, Genome Position End, Methylation %, Gene ID and Probe name to the database.

3.6.5 Methy.Cancer the database of human DNA

The URL for the database is *http://methy.cancer.geno.mics.org.cn*, which is used to study the interplay of gene expression, DNA methylation, and cancer. It mainly hosts highly integrated data of both cancer-related gene as well as DNA methylation information from public resources. This database provide user-friendly access to mostly all the data and their connections. [32]. It provides information related to Gene Symbol, Alias, Karyotype and Ensembl ID to the database.

3.6.6 ENSEMBLE Database

ENSEMBLE is genome browser that supports research in transcriptional regulation, variation in sequence, evolution & comparative genomics. It interprets genes and computes multiple alignments to predict the regulatory function and collects data related to various diseases. It also includes tools like Basic Local Alignment Search Tool, BLAT Variant Effect Predictor (VEP) and Biomart for all supported species. It provides information related to percentage GC content to the database.

3.7 Gene-SNP data

3.7.1 RegulomeDB

RegulomeDB is a database that escorts the interpretation for controlling variants in human genome; it contains high throughput data sets that was collected experimentally from

ENCODE and various sources. In addition, it estimates computationally and annotations were performed manually to recognize putative regulatory potential and predicts the functional variants. A powerful tool for combining the data source that gives the variants scores to support each separate functional variants from an extent of large pool and at last providing a minor set of reputed sites with theories that are testable as to achieve up to their function. [33]

3.7.2   HumSavar

The URL for **HumSavar** is *http://www.uni.prot.org/support./docs/humsavar.html*, in which the variations of sequence are distinguished for various human proteins, which are recorded in a text named humsavar with an .txt extension, which grants the sequence amino acid substitution and its position, it also mentions the variation type into two categories either polymorphism or disease mutation. [34] It provides information related to Gene, Entrez ID, dbSNP, Swiss-Prot-AC, AA-change, KEGG Link and dbSNP Link to the database.

3.8   Basic Gene Information

3.8.1   HGNC

It provides information related to RefseqID, Entrez Gene ID, Chromosome, HGNC ID, Enzyme IDs, Synonyms, Gene Family ID, Ensembl Gene ID, PubMed IDs, UCSC ID, OMIM ID, Vega ID, UniProt ID and Accession Numbers to the database.

3.8.2   Ensemble

It provides information related to Gene name, HGNC ID, NCBI Gene ID, PDB ID, Vega gene ID, ILLUMINA Human Methylation probe, Karyotype band, Gene Start (bp) and Gene End (bp) to the database.

3.9   Interaction (Network module identification)

The Protein Protein Interaction network for Dermatological related Diseases are very complex and contains a large number of 1) nodes and 2) edges, making it incomprehensible for the extraction of interactions that have a biologically meaning. As the cellular component and molecular functions are modular, thus the present work was to identify major modules or clusters that are enriched with various biological function or processes from the complex network of

dermatological diseases. This modular analysis gives a improved insight for knowing the relationship between various complex dermatological diseases.

The PPI network is built up using Cytoscape tool. The prediction of the functional behavior of a particular module are easier than that of an individual gene. To identify the biologically meaningful modules in the Dermatological related Diseases interaction network, assigning different weights to the interactions based on various type of experiment from the retrieved interaction. Among all the clustering algorithms, we chose Degree distributions.

The degree of a node denoted as "n", in undirected networks is given by the number of edges; node n is linked to. A node is counted as a self-loop of like the node two edges are considered as a degree [35]. The nodes count with a given degree h where h = 0,1,… gives the node degree distribution. The node with an in-degree n in directed networks is said to be the count for the incoming edges whereas the out-degree is the count for all outgoing edges. Related to networks that are undirected, there exists an in and out-degree distribution as said by Barabási and Albert who used the node degree distribution to distinguish between random (as defined by Erdős and Rényi) [36] and scale-free network topologies [37].

3.9.1 Cytoscape

Cytoscape is a freely available and open source software that integrates various bio-molecular networks that interacts with high-throughput data of expression. It is most powerful when used in conjunction with large databases of protein–DNA, protein–protein and various interactions involved genetic that are progressively available for humans and model organisms. It provides functionality that is basic to the design and queries network, allows integration visualizing the network with phenotypes and profiles of expression; and creates a link to the network of functional annotations databases. [38]

There are basically a variety of different modeling environments that has been developed for simulating biochemical reactions and gene transcription kinetics [39], cellular physiology [40] and metabolic control [41].

MATERIALS AND METHODS

The MethSRD (Human DNA Methylation database for genes associated with skin-related disorders), SNiPSRD (Single nucleotide polymorphisms for genes associated with Skin Related Disorders) and InteractSRD (Interaction database for genes associated with skin-related disorders): A computational database for genes associated with dermatological disorders will be a database. The data for various genes associated with Skin Related Disorders was extracted across a large number of published literature, several specialized databases and other repositories that give an insight data regarding methylation data of DNA.

The MethSRD database will be contain around 11 column namely Gene ID, Chromosome, gene-disease association, transcript for respective gene, Methylation Type, Gene Expression, PubMed ID, Methylation percentage, ILLUMINA Human Methylation probe, percentage of GC Content, Karyotype Ontology data like Biological Processes, Molecular Function, Cellular Component and KEGG Pathway for respective gene and diseases.

The SNiPSRD database will be containing column namely Gene name, Gene ID, Single Nucleotide Polymorphisms ID (dbSNP), Swiss-Prot-AC, and Amino Acid –change.

The InteractSRD database will be containing interaction between Gene or Disease, Gene or Pathway, Gene or Molecular Function, Gene or Methylation Type, Gene or Gene Expression, Disease or Pathway, Disease or Molecular Function, Disease or Methylation Type, and Disease or Gene Expression.

The web service was implemented using the HTML, CSS, JavaScript, Jquery and JSON format. The web service will be having the functionality for searching various aspects like Methylation search by Gene name, Disease name, Chromosome Number, SNiPs search by Gene name, Gene Ontology search by Gene name and Interaction search.

METHODOLOGY

A1. Data Collection and Data Curation for SNiPSRD and MethSRD database creation

For data collection and curation was perfromed and are mainly into four types of data included in the database: (i) Gene-Disease Association (ii) Gene Methylation data, (iii) SNiP and mutations data, (iv) Gene/Disease Ontology (Biological Process, Molecular Function, Cellular Component and KEGG Pathway) and (v) General Gene data.

(i)   Gene-Disease Association

To know the disease, the disease list was identified from Genetic Skin Disorders from an annotated bibliography contains many review papers named Genetic Skin Disorders [42]. To study the gene-disease association, the gene-disease association data was integrated with mainly three databases (i) **Comparative Toxicogenomics Database [43]**. (ii) **DisGeNET [44]** (iii) **Genetic Association Database [45]**

(ii)   Gene Methylation

To study the methylation of a gene, the methylation data for DNA, the data were integrated manly from databases like DBCAT, MethHC, DDMGD, DiseaseMeth, Ensemble, and MethylCancer. The Data from **DBCAT** contains information related to **Gene ID** and **Chromosome**. Data from **MethHC** be responsible for **p-value** between include gene and disease, and **Transcript** for respective gene. **DDMGD** provides information related to **Methylation Type, Gene Expression**, and **PubMed ID**. Whereas **DiseaseMeth** provides information related to **Methylation percentage. Ensemble** provides information related to **ILLUMINA Human Methylation probe and percentage of GC Content. MethyCancer** provides information related to **Karyotype**.

(iii)   SNiP and mutations data

The SNiP (Single Nucleotide Polymorphism) and mutations data are mainly from HumSavar, which provide information related to Gene Name, Gene ID, dbSNP, Swiss-Prot-AC, AA-change, KEGG Link, dbSNP Link

(iv) Gene/Disease Ontology (Biological Process, Molecular Function, Cellular Component and KEGG Pathway)

In order to study the ontology of a gene or disease, the ontology data were integrated mainly from DBCAT. The Data from DBCAT contains information related to Biological Process, Molecular Function, Cellular Component and KEGG Pathway and their respective GO ids.

(v) General Gene data

In general information regarding gene are mainly from **HGNC** and **Ensemble**, **HGNC** provides information related to RefseqID, Entrez Gene ID, Chromosome, Enzyme IDs, Synonyms, Gene Family ID, Ensembl Gene ID, PubMed IDs, UCSC ID, OMIM ID, UniProt ID, Accession Numbers. **Ensemble** provides information related to Gene name, HGNC ID, PDB ID, Vega Gene ID and Gene End and Start (bp).

A2. Data Integrity (Uniqueness)

The data collected from various resources were scattered which are heterogeneous in nature, thus Firstly there was creation for a uniform format for describing the diverse and improper data to a compressed and well distributed data by integrating the data with the help of mapping of the corresponding data with respect to the gene associated with Dermatological Disorder.

A3. Data Modelling (Categorizing)

Data modeling is a step for database designing to create a conceptual model of how 2data items relate to each other. Thus, here we did the data modeling by categorizing the integrated data into mainly five categories that listed as [i] **Gene-disease association** that contains column regarding Gene Name, Disease Name, Disease Class and Disease ID. [ii] **Gene Methylation** that provides information regarding Gene ID, Chromosome, p-value, Transcript, Methylation Type, Gene Expression, PubMed ID, Methylation percentage, Methylation probe, % of GC Content and Karyotype. [iii] **SNiP and mutation data** that gives information regarding the Gene Name, Gene ID, dbSNP, Swiss-Prot-AC, Amino Acid –change. [iv] **Ontology Data** gives information related to Biological Process, Molecular Function, Cellular Component and KEGG Pathway. [v]

General Gene related information, provides information such as OMIM ID, UniProt ID, Accession Numbers, Gene name, HGNC ID, PDB ID, Vega Gene ID, Gene End and Start (bp), RefseqID, GeneID, Chromosome, Enzyme IDs, Synonyms, Gene Family ID, Ensembl Gene ID, PubMed IDs and UCSC ID.

A4. Graphical User Interface for Database

The construction of the user interface was done by using the JSON, Jquery, JavaScript, CSS, and HTML. In total eight (.html) pages namely Home.html, SearchMethylation.html, SearchSNiP.html, SearchGeneInformation.html, GeneGO.html, DiseaseGO.html, information.html and Contact.html was constructed. Home.html page contains the Welcome page-containing introduction to MethSRD, Methylation and Data source. **SearchMethylation.html** Search for the Methylation related data for the queried Gene, and it contains three ways to search the database either by selecting gene name or chromosome umber or disease type. **SearchSNiP.html** and **SearchGeneInformation.html** Search the database for SNP's and general gene information by selecting gene name respectively. **GeneGO.html** and **DiseaseGO.html** Search for Pathway, Molecular Function, Biological Process and Cellular Component by selecting gene name and disease name respectively.

A5. InteractSRD database creation

a.   Interaction Networks using Cytoscape

b.   Graphical User Interface for InteractSRD

A5.a Interaction Networks using Cytoscape

Next step was to developed an network of interaction that provide a novel cues for the connectivity of Dermatological-related disorders and towards to bridge the proteins associated with multiple Dermatological-related disorders. Dermatological-related disorders and Proteins Bipartite Network: If each Dermatological-related disorders incline to have a unrelated and individual genetic origin, then the network named bipartite network would disconnect entire network into plenty of single nodes that are corresponding to specific disorders or might be grouped later to form cluster of small network which are closely related disorders.

However, in bipartite network of the SRDs-gene exhibits various connections between both individual dermatological disorders and 45 disorder classes. The InteractSRD database will be containing interaction between Gene or Disease, Gene or Pathway, Gene or Molecular Function, Gene or Methylation Type, Gene or Gene Expression, Disease or Pathway, Disease or Molecular Function, Disease or Methylation Type, and Disease or Gene Expression.

A5.b Graphical User Interface for InteractSRD

The construction of the user interface was done by using the JSON, Jquery, JavaScript, Cytoscape.js, CSS, and HTML. In total nine (.html) pages **GeneDisease.html** page contains the dropdown for selecting single or multiple disease name, after which the next dropdown pops up that contains gene associated with the disease selected from the disease list. **GenePathway.html** page contains the dropdown for selecting single or multiple pathway names, after which the next dropdown pops up that contains the gene associated with the pathway selected from the pathways list. **GeneMolecularFunction.html** page contains the dropdown for selecting single or multiple Molecular Function name, after which the next dropdown pops up that contains the gene associated with the Molecular Function selected from the Molecular Functions list. **GeneMethylationType.html** page contains the dropdown for selecting single or multiple Methylation type name, after which the next dropdown pops up that contains the gene associated with the Methylation type selected from the Methylation type list. **GeneExpression.html** page contains the dropdown for selecting single or multiple Expression names, after which the next dropdown pops up that contains the gene associated with the Expression selected from the Expression list. **DiseasePathway.html** page contains the dropdown for selecting single or multiple pathway names, after which the next dropdown pops up that contains the Disease associated with the pathway selected from the pathways list. **DiseaseMolecularFunction.html** page contains the dropdown for selecting single or multiple Molecular Function name, after which the next dropdown pops up that contains the Disease associated with the Molecular Function selected from the Molecular Functions list. **DiseaseMethylationType.html** page contains the dropdown for selecting single or multiple Methylation type name, after which the next dropdown pops up that contains the Disease associated with the Methylation type selected from the Methylation type list. **DiseaseGeneExpression.html** page contains the dropdown for

selecting single or multiple Expression names, after which the next dropdown pops up that contains the Disease associated with the Expression selected from the Expression list.

Figure 1- Flow of Methodology

RESULTS

MethSRD, SNiPSRD, and InteractSRD catalogs information about Methylation Patterns, Single Nucleotide Polymorphism, and interaction associated with dermatological related disorders. The database can be queried individually with the gene name. MethSRD currently has information on more than 1000 genes associated with 139 different type of dermatological disorder. The distribution of Skin related Disorder among various disease classes has been represented in Figure A.

The developed database is available as free online resources at *http://genomeinformatics.dtu.ac.in/dermameth/home.html*.   The result page mainly contains five different sections, the first section is regarding the **Methylation Data** that provides details regarding Gene Name, Disease Name, Disease Class and Disease ID, Gene ID, Chromosome No., p-value, Transcript, Methylation Type, Gene Expression, PubMed ID, Methylation percentage, Methylation probe, % of GC Content and Karyotype. The next section is **SNiP and mutation data** that gives information regarding the Gene Name, Gene ID, dbSNP, Swiss-Prot-AC, Amino Acid –change. The third section is **Ontology Data** that gives information related to Biological Process, Molecular Function, Cellular Component and KEGG Pathway. The fourth section provides the **General Gene related information** that provides information such as OMIM ID, UniProt ID, Accession Numbers, Gene name, HGNC ID, PDB ID, Vega ID, End and Start (bp), RefseqID, GeneID, Chromosome, Enzyme IDs, Synonyms, Gene Family ID, Ensembl Gene ID, PubMed IDs and UCSC ID. The last section is regarding the **Interaction Data** that gives information regarding the interaction between Gene or Disease, Gene or Pathway, Gene or Molecular Function, Gene or Methylation Type, Gene or Gene Expression, Disease or Pathway, Disease or Molecular Function, Disease or Methylation Type and Disease or Gene Expression.

USER INTERFACE

MethSRD and SNiPSRD provides a user-friendly interface for querying comprehensive information on each gene associated with the dermatological disorder (Figure. 2). Users can query the database for either retrieving Methylation data, SNiP data, General Gene Information data or for Gene Ontology data through either by Genes name or chromosome number or disease class.   In total eight (.html) pages namely Home.html, SearchMethylation.html,

SearchSNiP.html, SearchGeneInformation.html, GeneGO.html, DiseaseGO.html, information.html and Contact.html was constructed.
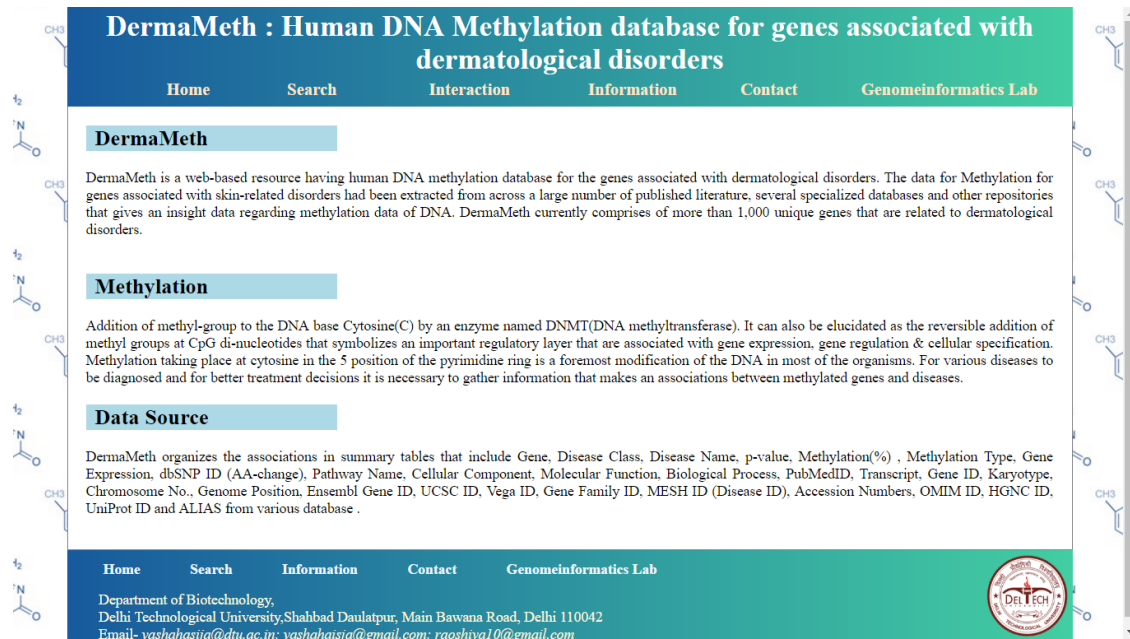
Figure 2- Home.html page contains the Welcome page-containing introduction to MethSRD, Methylation and Data Source
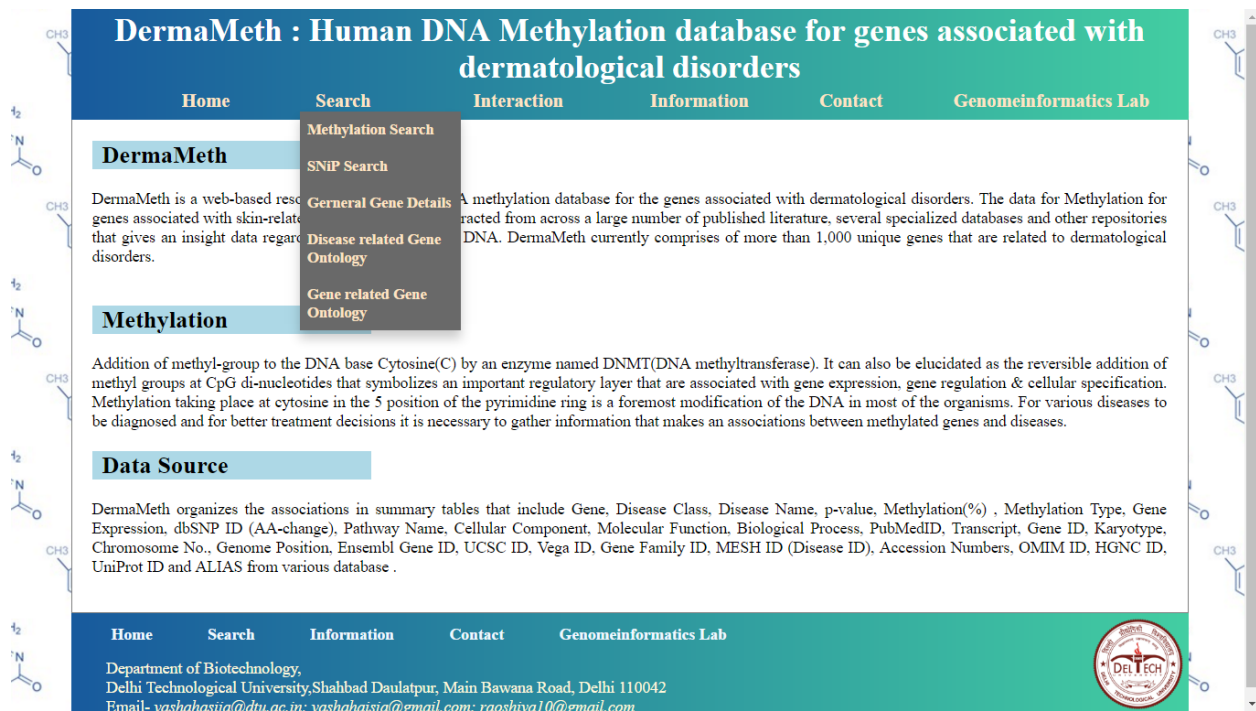
Figure 3- A user can search for either Methylation, SNiP, General Information, Ontologies with respect to either Gene or Diseases.
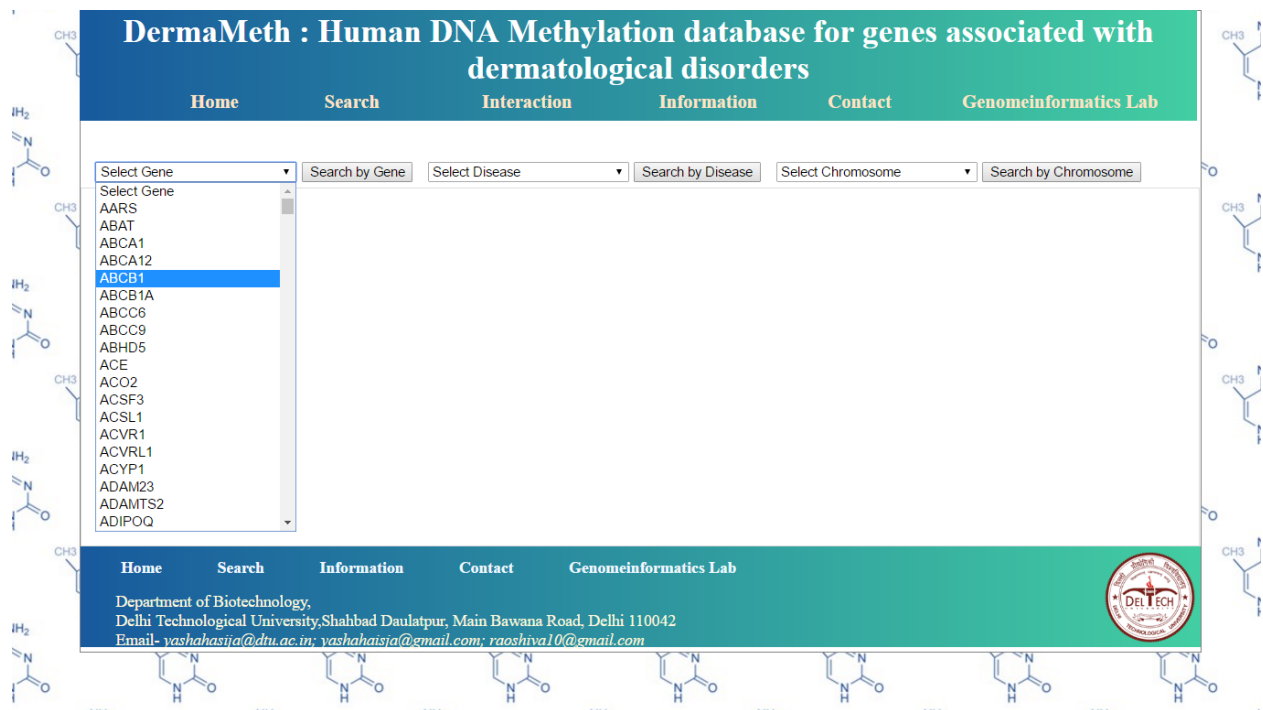
Figure 4- User can either use three ways to retrieve the methylation data by selecting Gene Name, by Chromosome No., or by Disease name respectively.
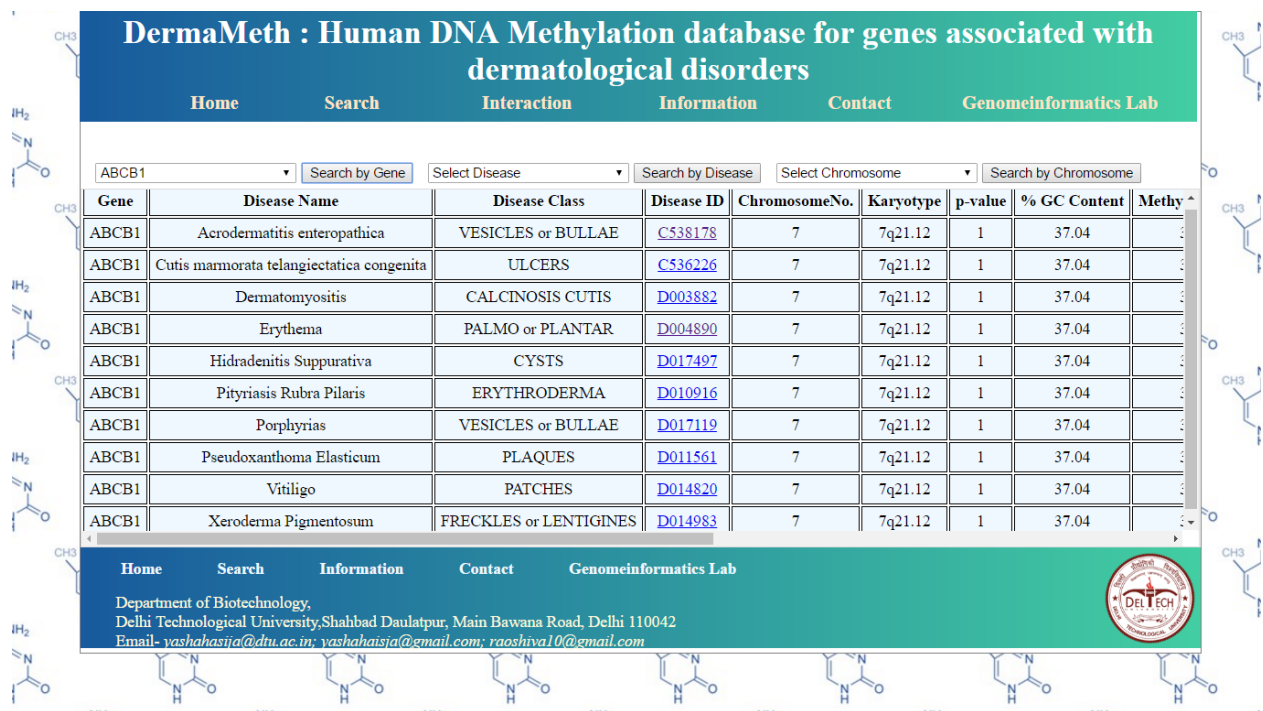
Figure 5- Result for Methylation data retrieved from gene search

Figure 6- Remaining column for result of Methylation data retrieved from gene search

Figure 7- SearchSNiP.html and SearchGeneInformation.html Search the database for SNP's and general gene information by selecting gene name respectively.

Figure 8- Search Result for SNiP data by selecting gene name.

Figure 9- Search result for General Gene information data by selecting gene name.

Figure 10- DiseaseGO.html Search for Pathway, Molecular Function, Biological Process and Cellular Component by selecting gene name and disease name respectively.

Figure 11- Search result page for Pathway by selecting disease name respectively.

Figure 12- Search result page for Molecular Function by selecting disease name respectively.

Figure 13- Search result page for Biological Process by selecting disease name respectively.

Figure 14- Search result page for Cellular Component by selecting disease name respectively.

Figure 15- GeneGO.html Search for Pathway, Molecular Function, Biological Process and Cellular Component by selecting gene name and disease name respectively.

Figure 16- Search result page for Pathway by selecting gene name.

Figure 17- Search result page for Molecular Function by selecting gene name.

Figure 18- Search result page for Biological Process by selecting gene name.

Figure 19- Search result page for Cellular Component by selecting gene name.

InteractSRD is a user-friendly interface provided for querying detailed information on each gene associated with the dermatological disorder (Figure. 2). Users can query the Interaction database for retrieving Interaction data by selecting the first column, which can be either Disease, Pathway, Molecular Function, Gene Expression or Methylation Type after which the next second column dropdown pops up that contains either gene/disease. After selecting the submit button, the user gets the interaction network.

In total nine (.html) pages namely GeneDisease.html, GenePathway.html, GeneMolecularFunction.html, GeneMethylationType.html, GeneExpression.html, DiseasePathway.html, DiseaseMolecularFunction.html, DiseaseMethylationType.html, and DiseaseGeneExpression.html was constructed.
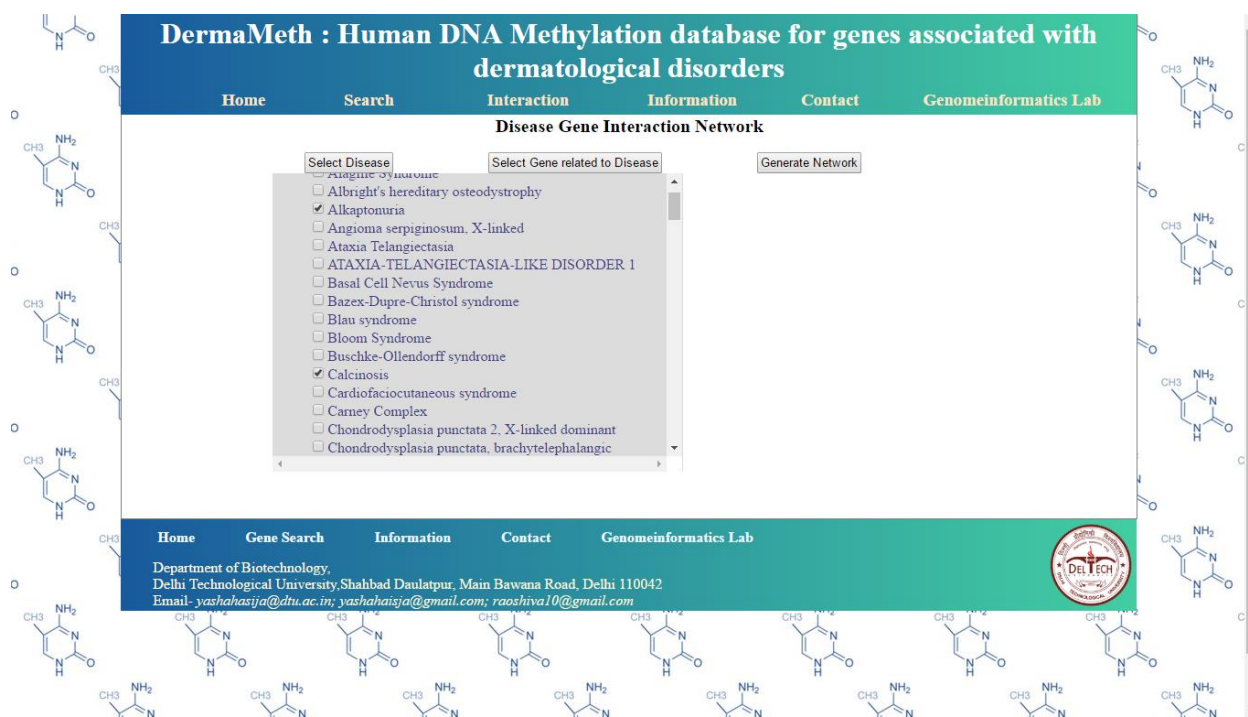
Figure 20- Search page for Gene-Disease by selecting multiple disease name.
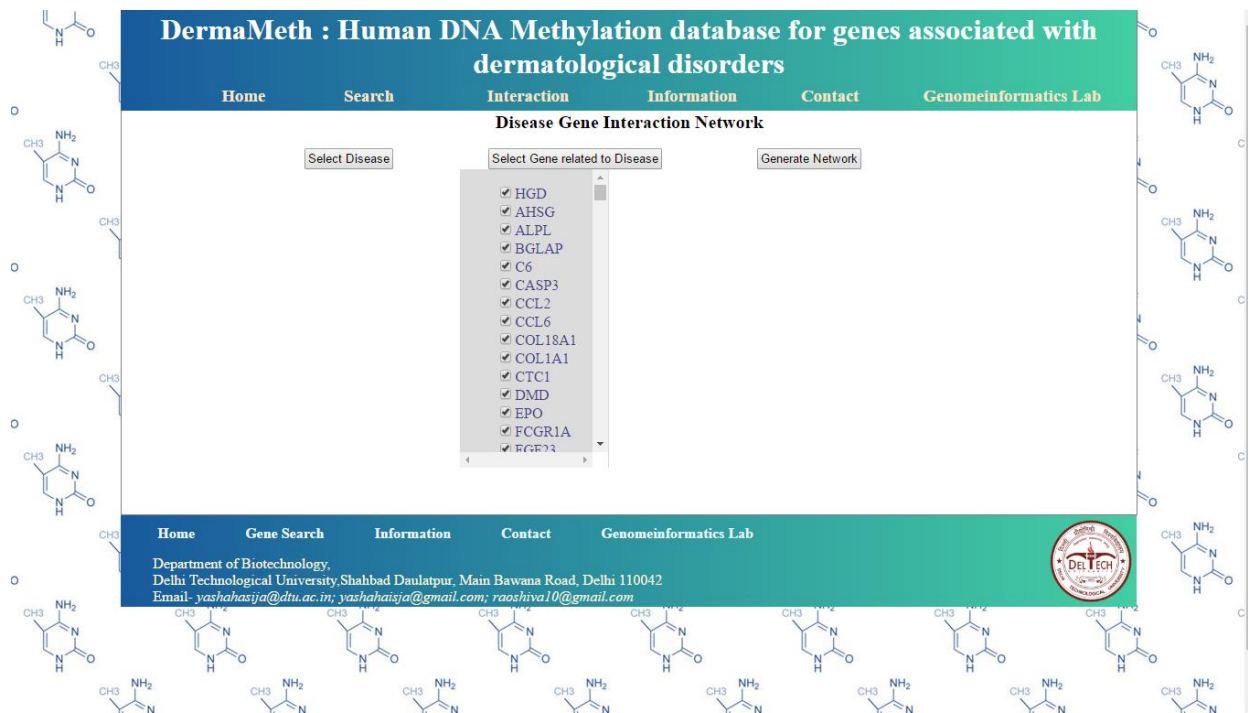
Figure 21- Search page for Gene-Disease by selecting multiple gene name for respective Diseases.
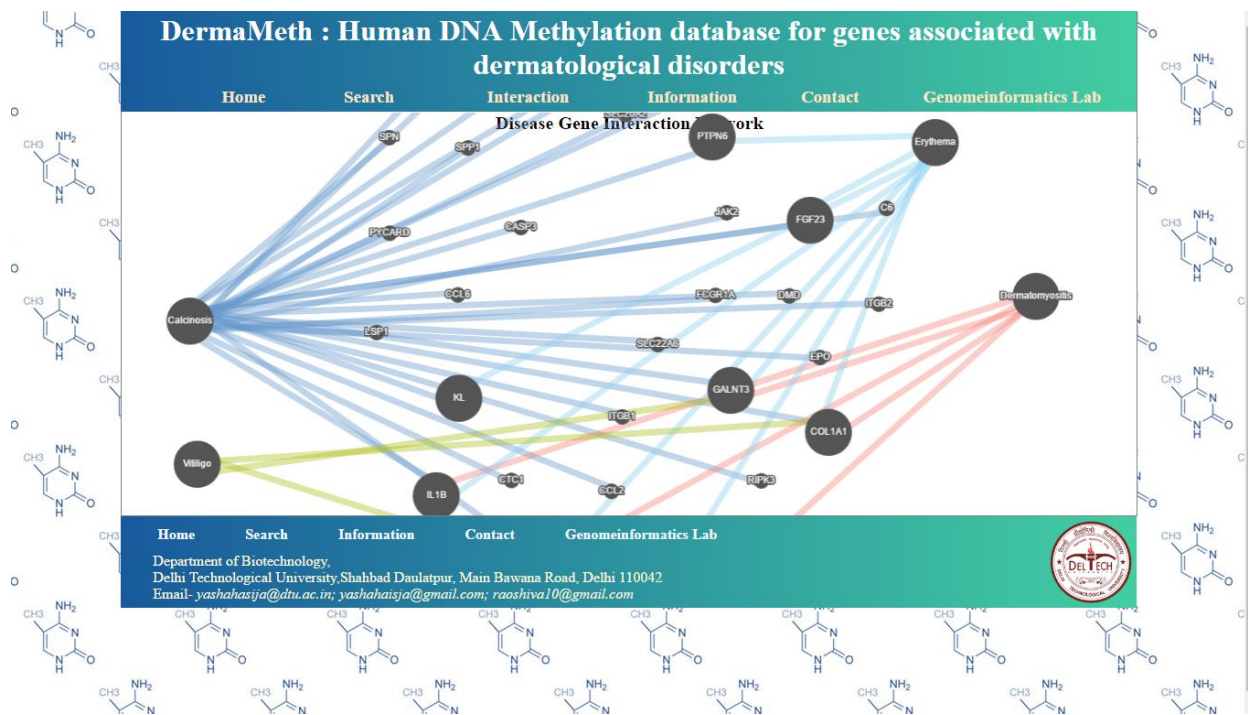
Figure 22- Result page for Gene-Disease interaction with different color indicating different Disease.

Figure 23- Search page for Gene-Pathway by selecting multiple pathway names.

Figure 24- Search page for Gene-Pathway by selecting multiple gene names for respective Pathway.

Figure 25- Result page for Gene-Pathway interaction with different color indicating different Pathway.

Figure 26- Search page for Gene-Molecular Function by selecting multiple Molecular Function name.

Figure 27- Search page for Gene-Molecular Function by selecting multiple gene names for respective Molecular Function.

Figure 28- Result page for Gene-Molecular Function interaction with different color indicating different Molecular Function.
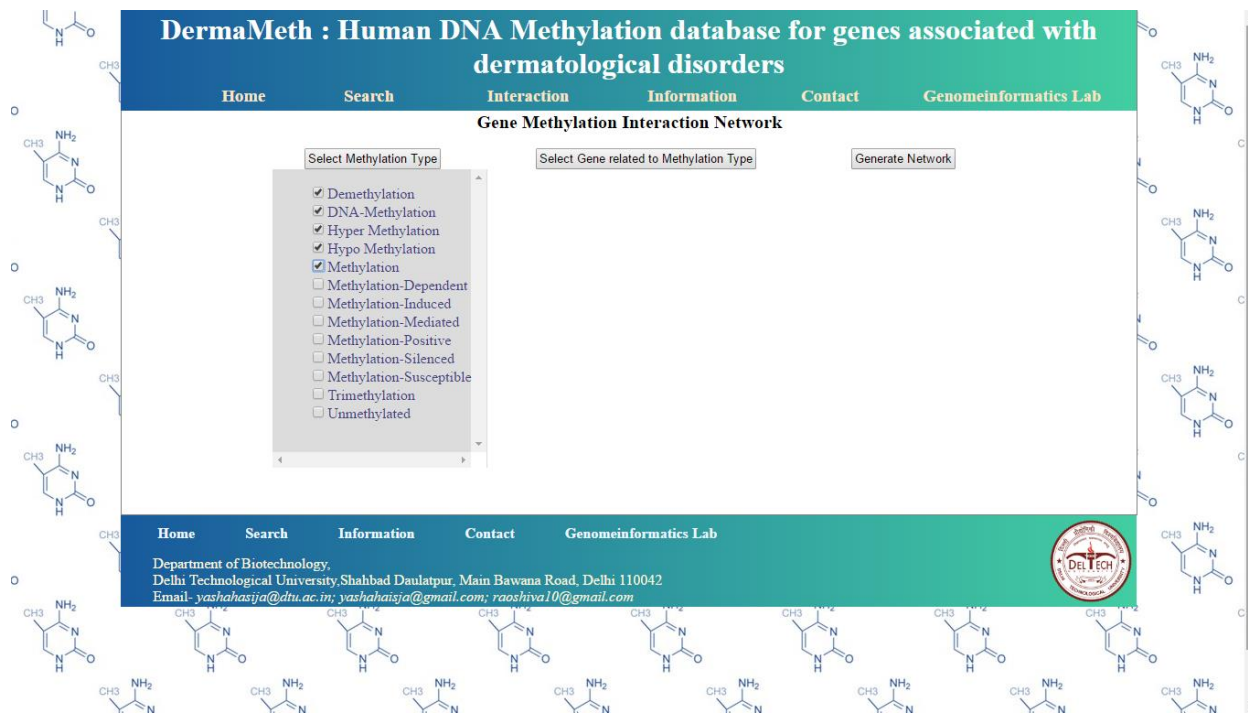
Figure 29- Search page for Gene-Methylation Type by selecting multiple Methylation Type name.
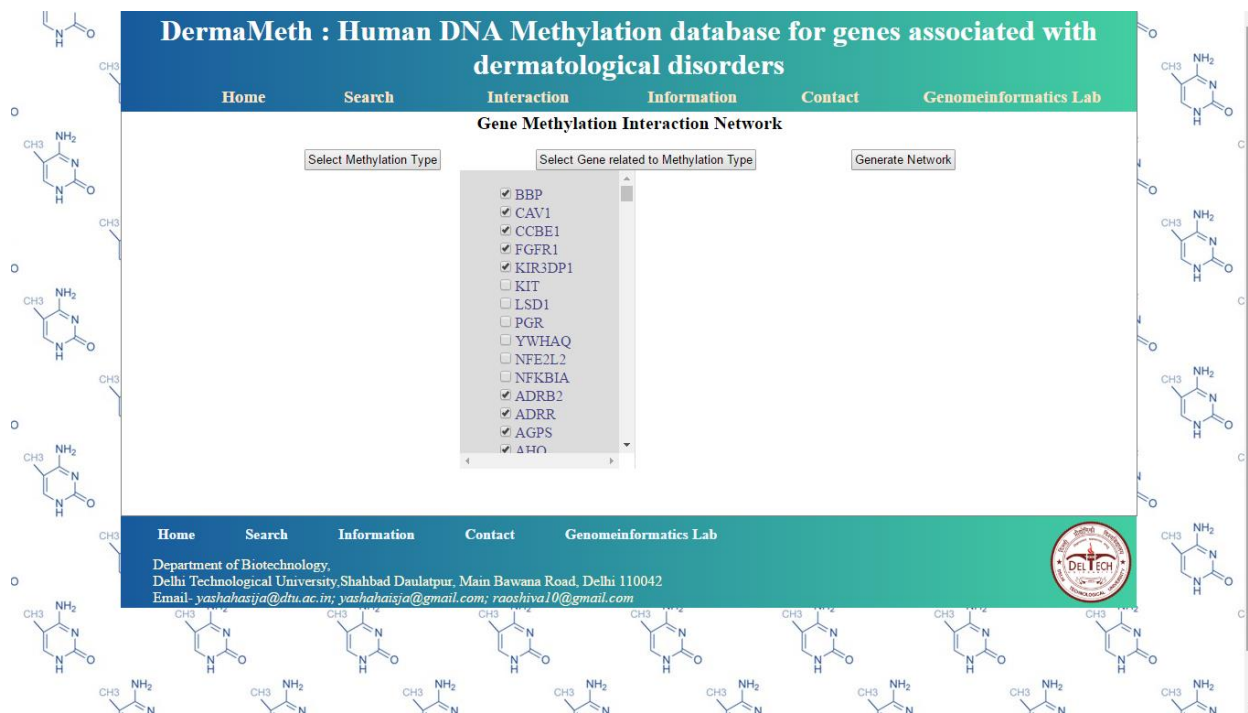
Figure 30- Search page for Gene-Methylation Type by selecting multiple gene names for respective Methylation Type.
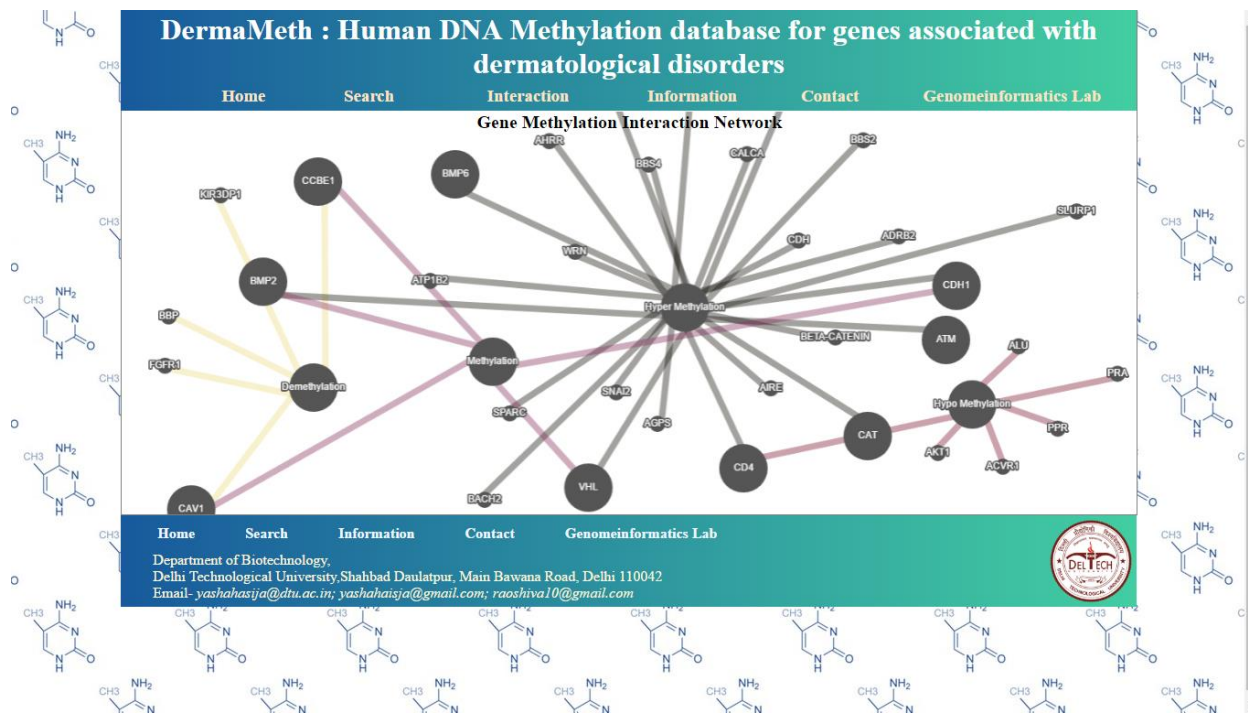
Figure 31- Result page for Gene-Methylation Type interaction with different color indicating different Methylation Type.

Figure 32- Search page for Gene- Expression by selecting multiple Gene Expression name.

Figure 33- Search page for Gene- Expression by selecting multiple gene names for respective Gene Expression.

Figure 34- Result page for Gene- Expression interaction with different color indicating different Gene Expression.
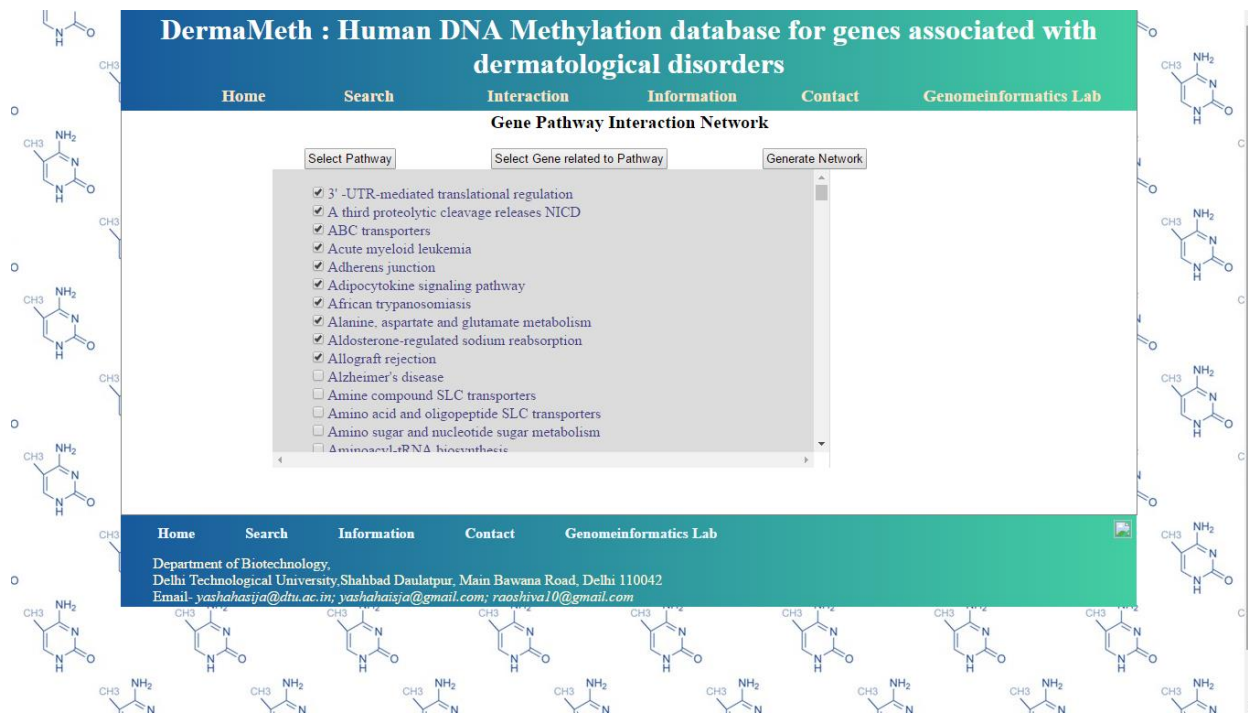
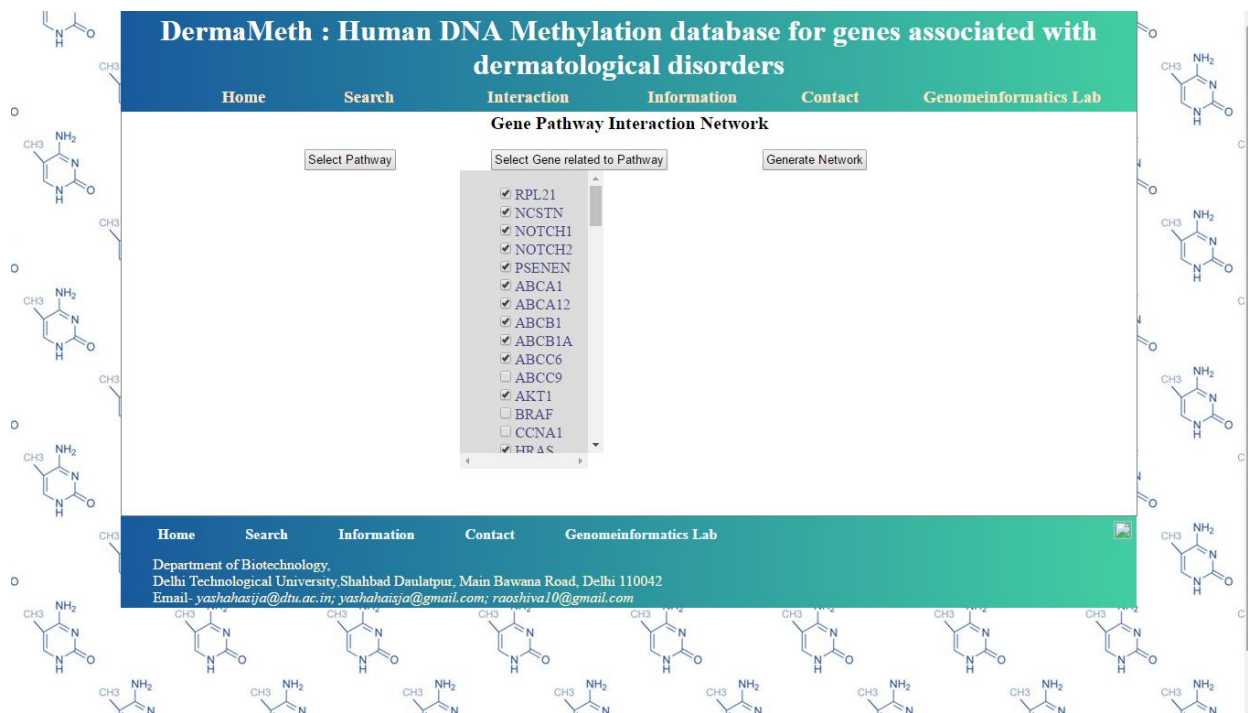Figure 35- Search page for Disease-Pathway by selecting multiple pathway names.

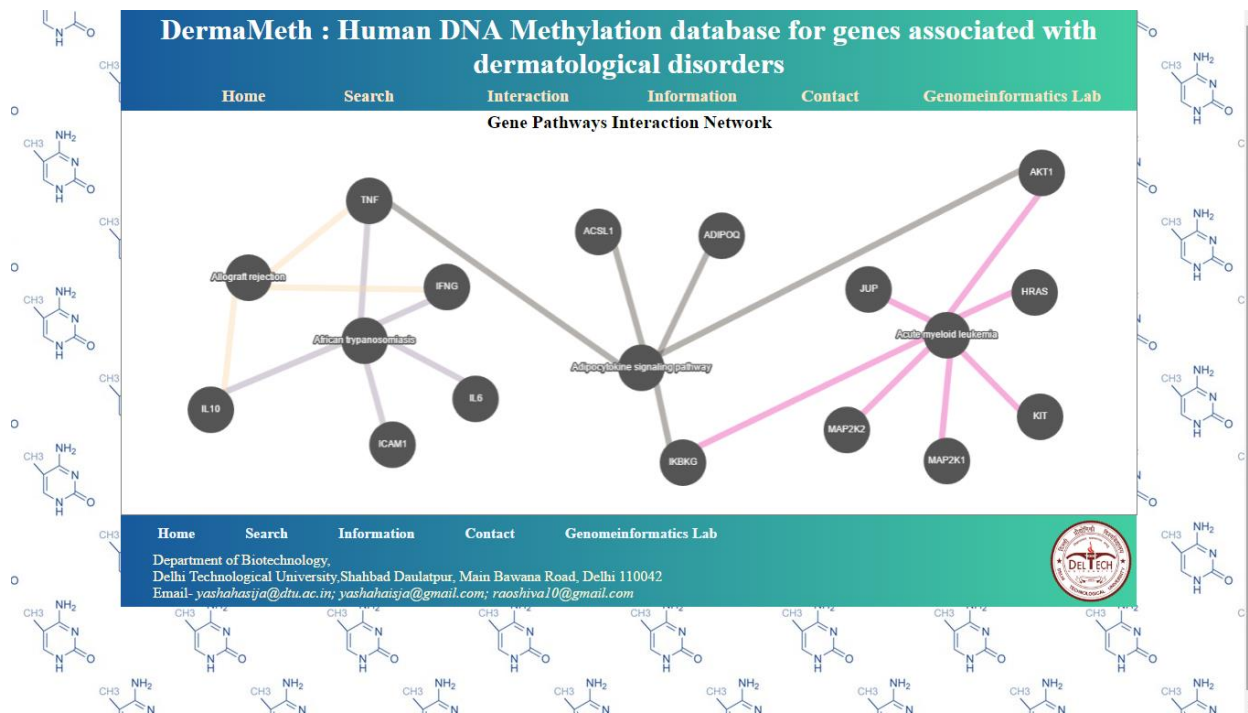Figure 36- Search page for Disease-Pathway by selecting multiple Disease name for respective Pathway.

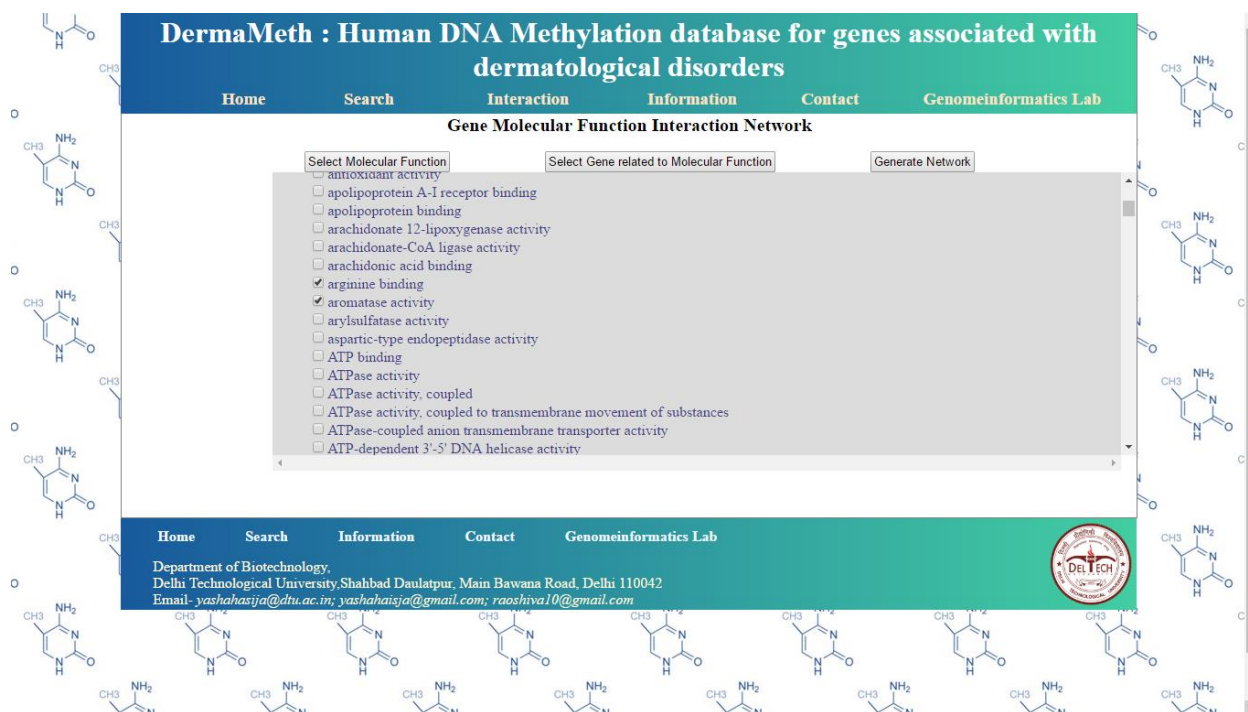Figure 37- Result page for Disease-Pathway interaction with different color indicating different Pathway.
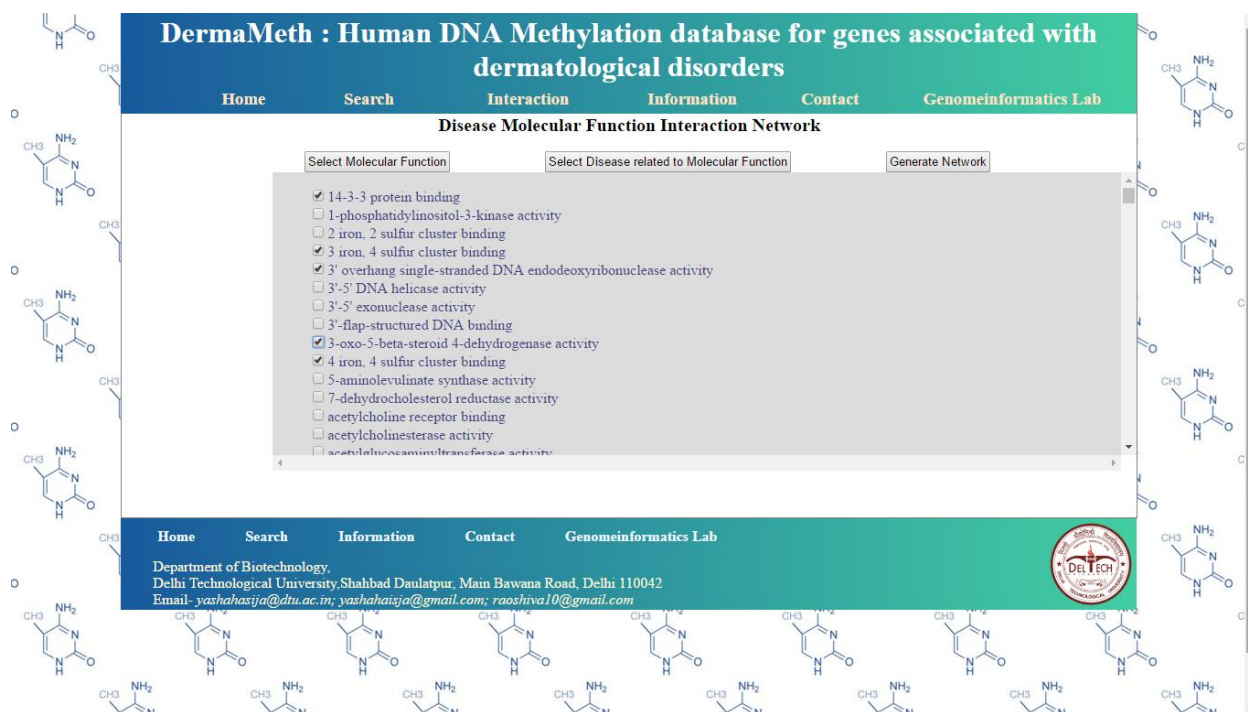
Figure 38- Search page for Disease-Molecular Function by selecting multiple Molecular Function name.
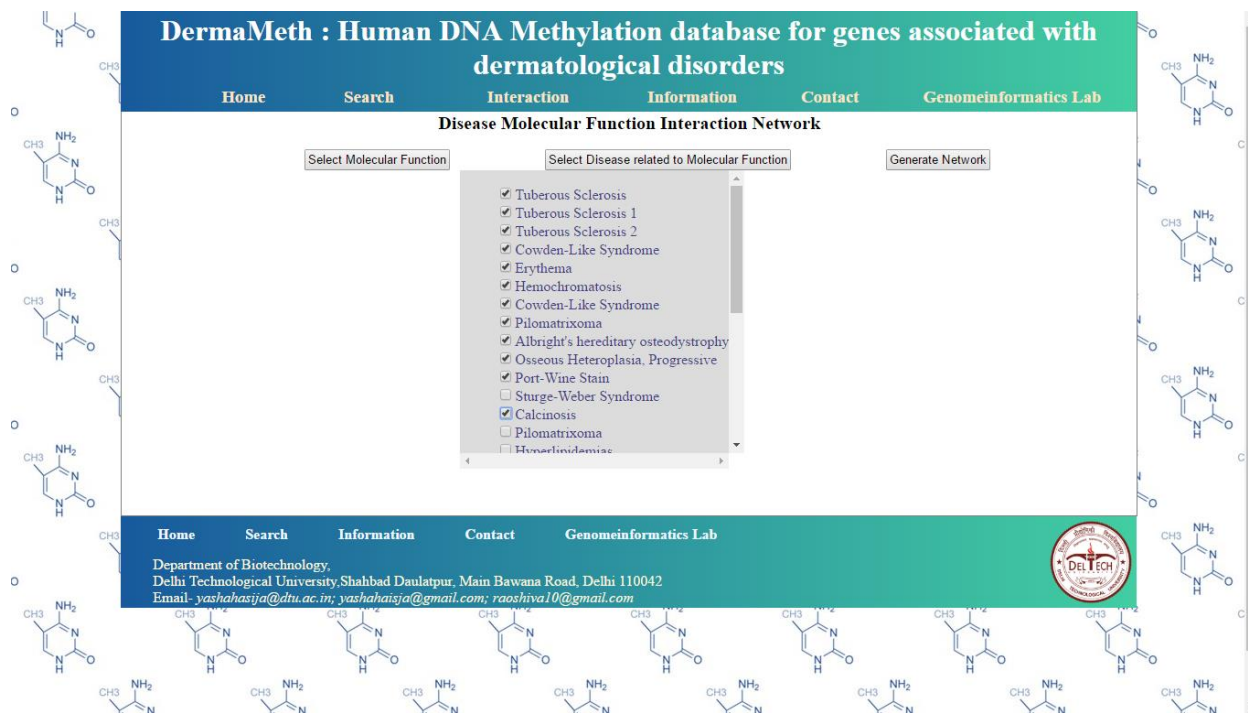
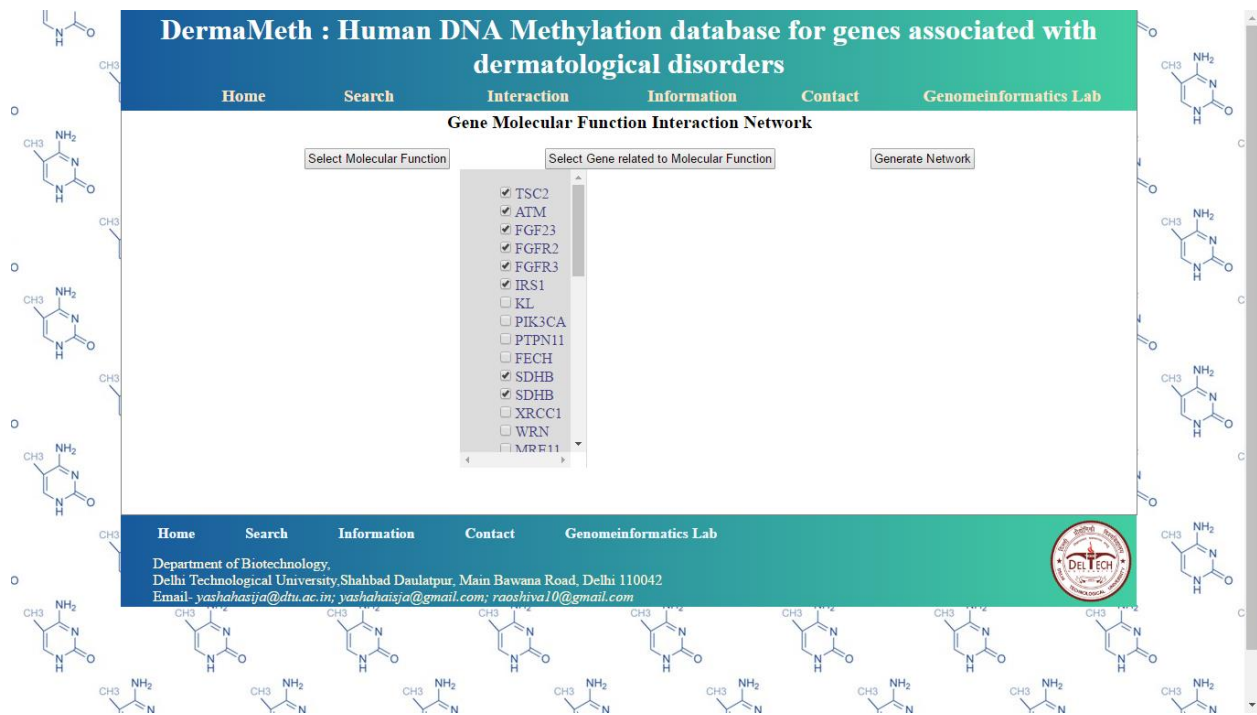Figure 39- Search page for Disease-Molecular Function by selecting multiple Disease name for respective Molecular Function.
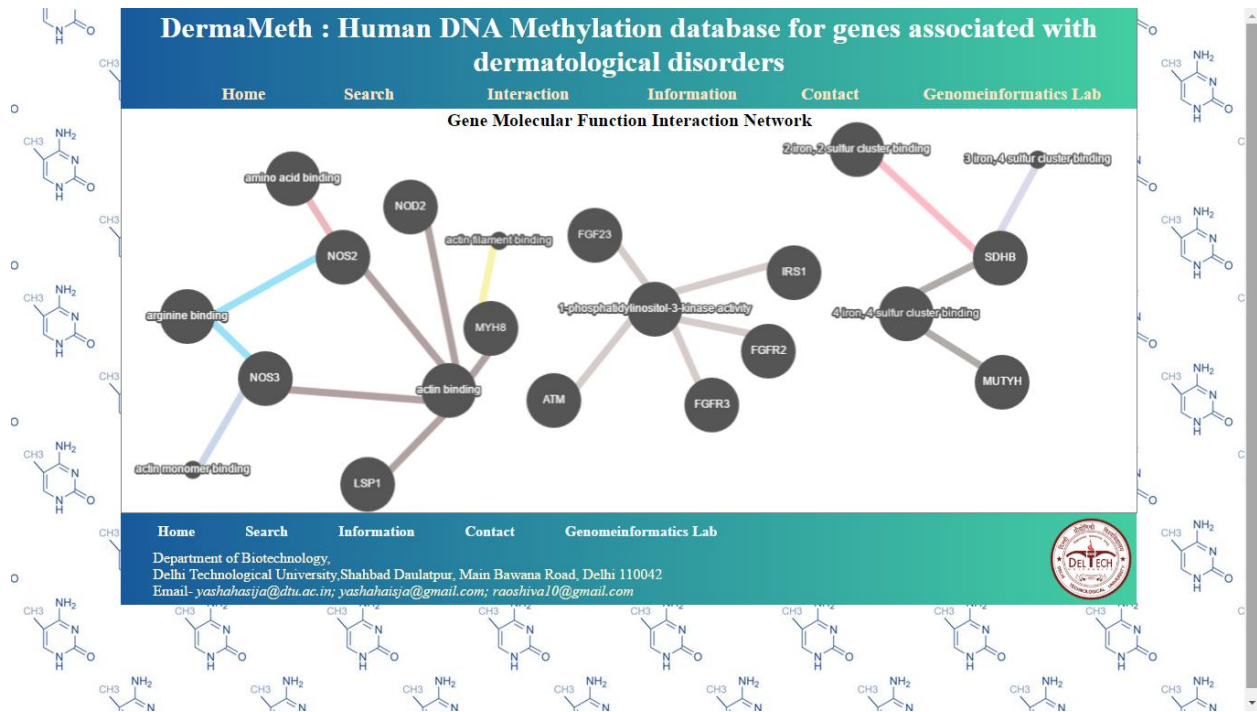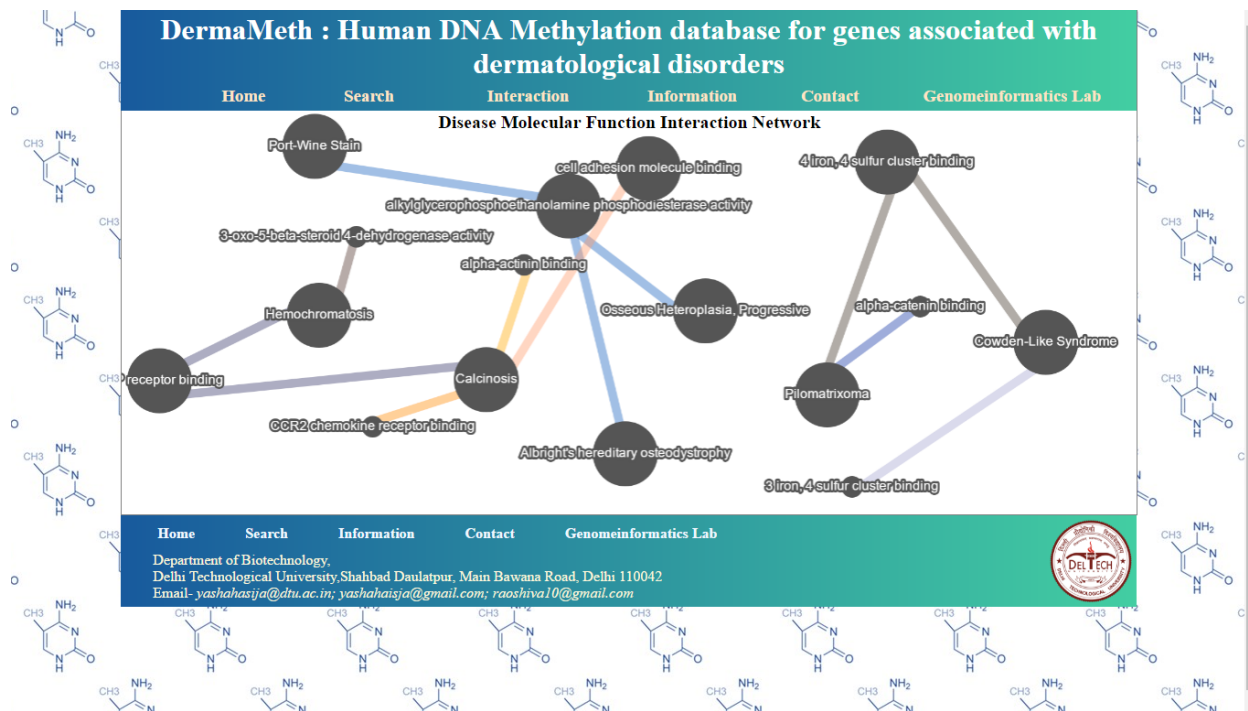

Figure 40- Result page for Disease-Molecular Function interaction with different color indicating different Molecular Function.
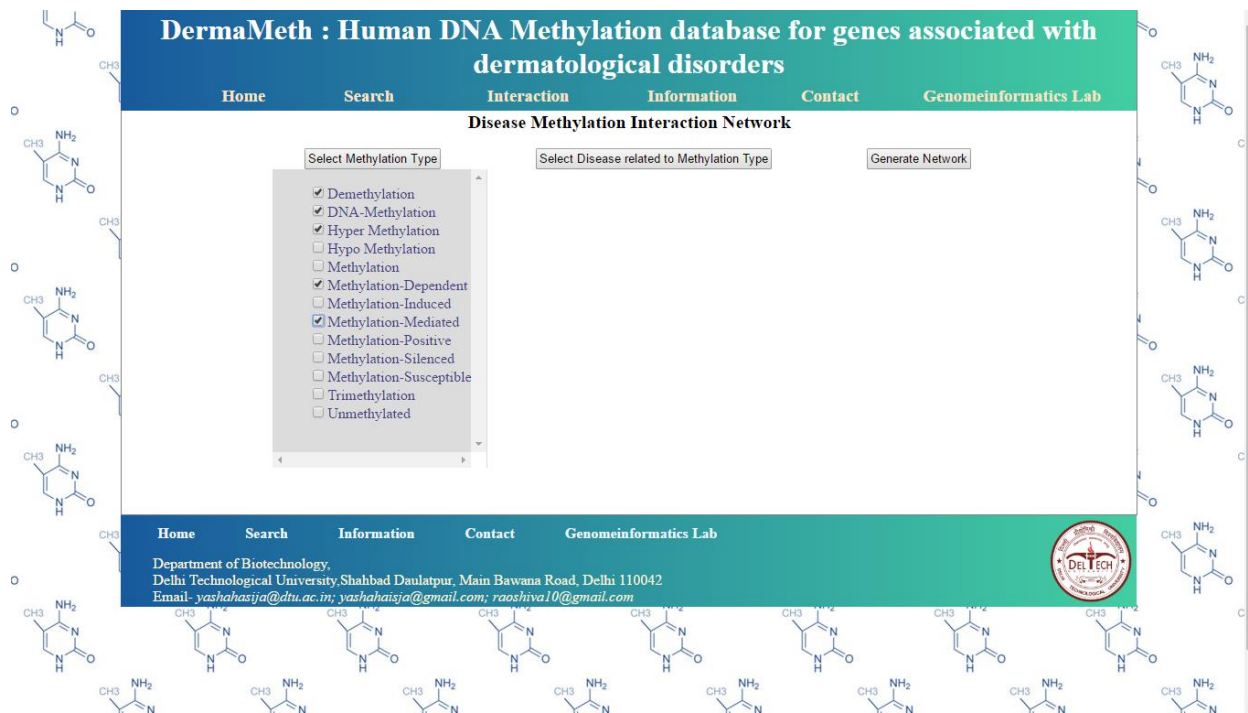
Figure 41- Search page for Disease-Methylation Type by selecting multiple Methylation Type name.
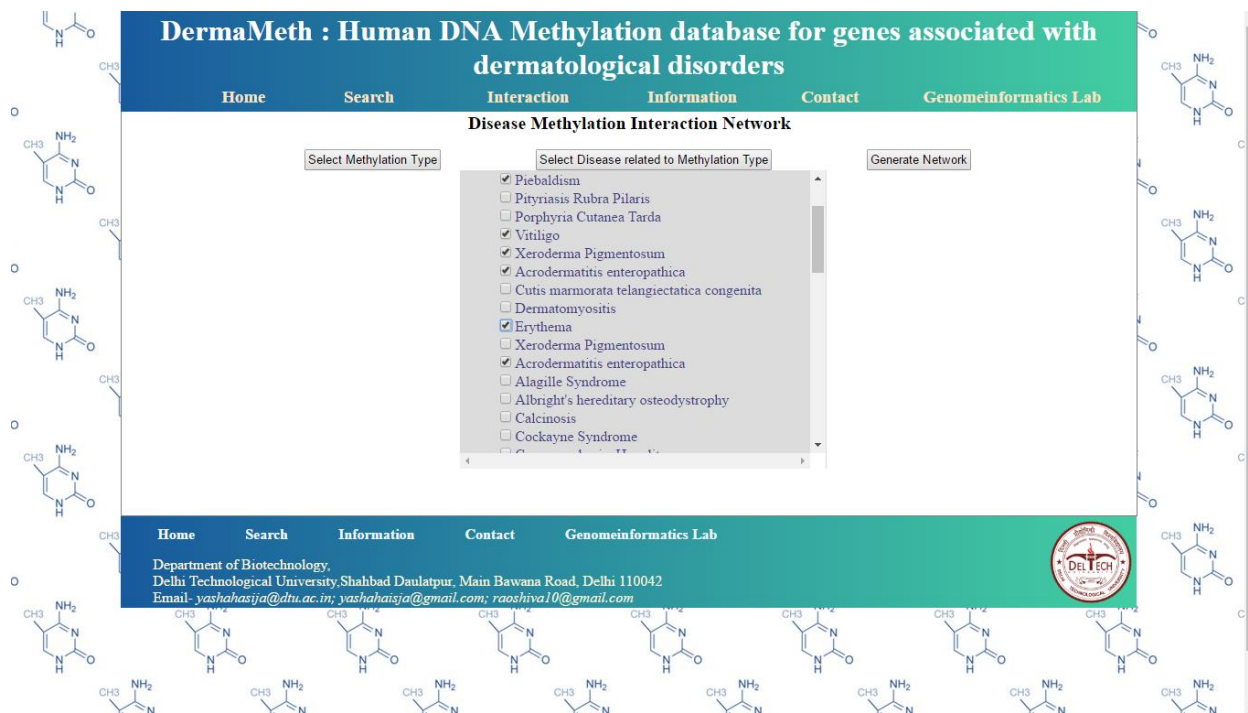
Figure 42- Search page for Disease-Methylation Type by selecting multiple Disease name for respective Methylation Type.
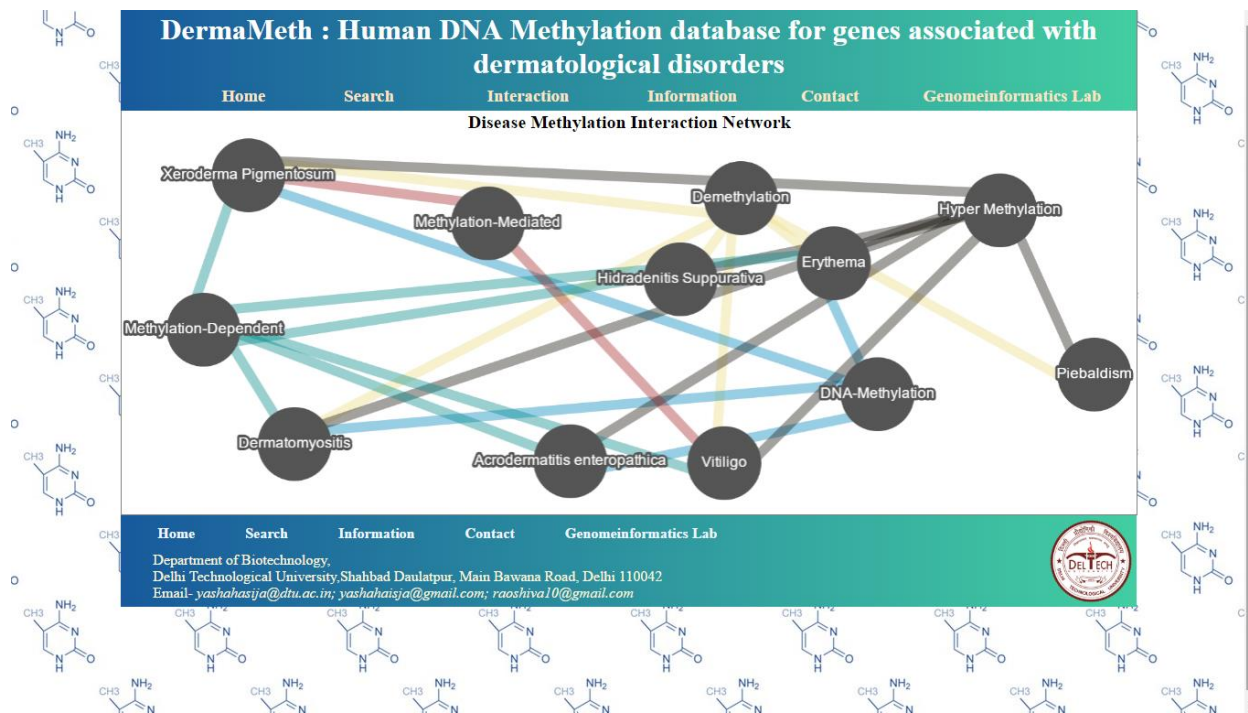
Figure 43- Result page for Disease-Methylation Type interaction with different color indicating different Methylation Type.
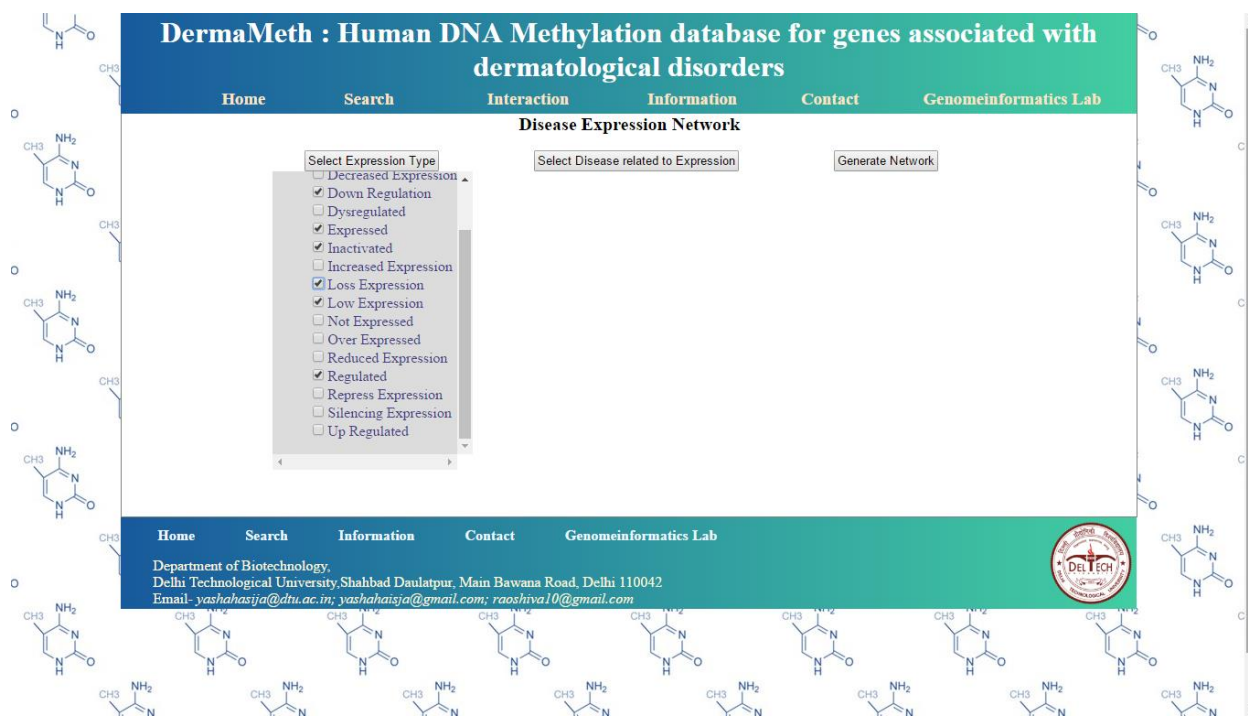
Figure 44- Search page for Disease- Expression by selecting multiple Disease Expression name.
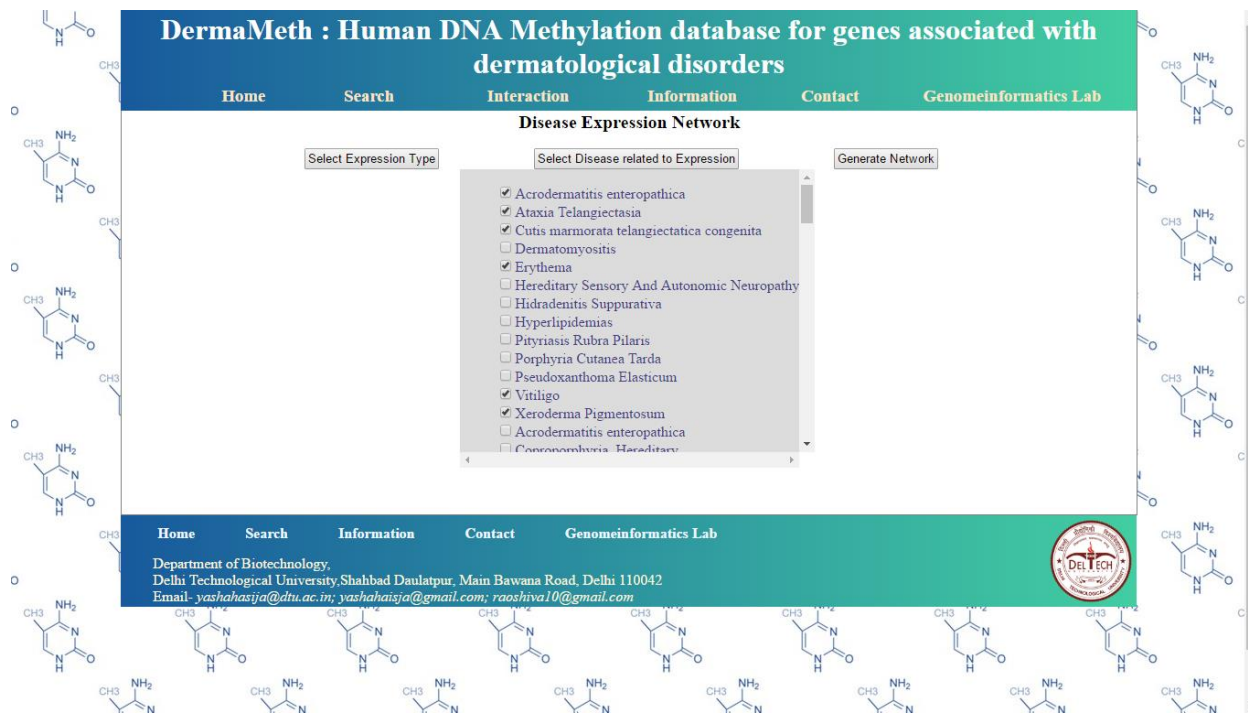
Figure 45- Search page for Disease- Expression by selecting multiple Disease name for respective Disease Expression.
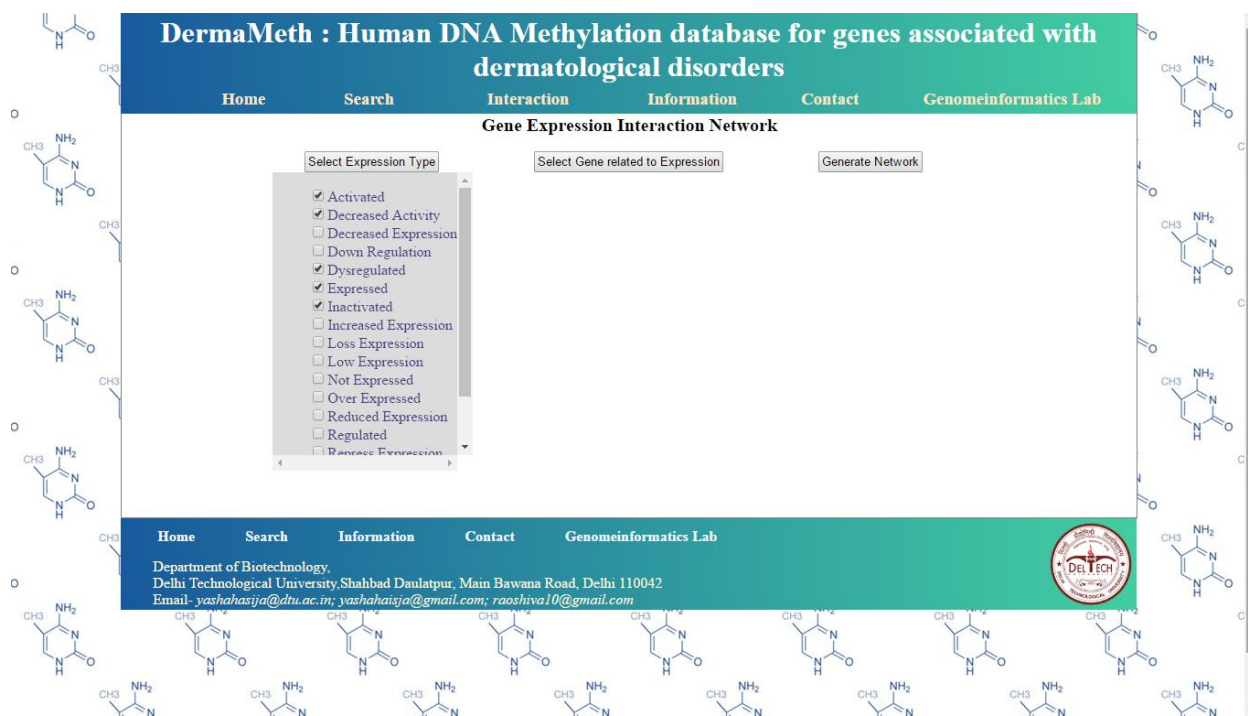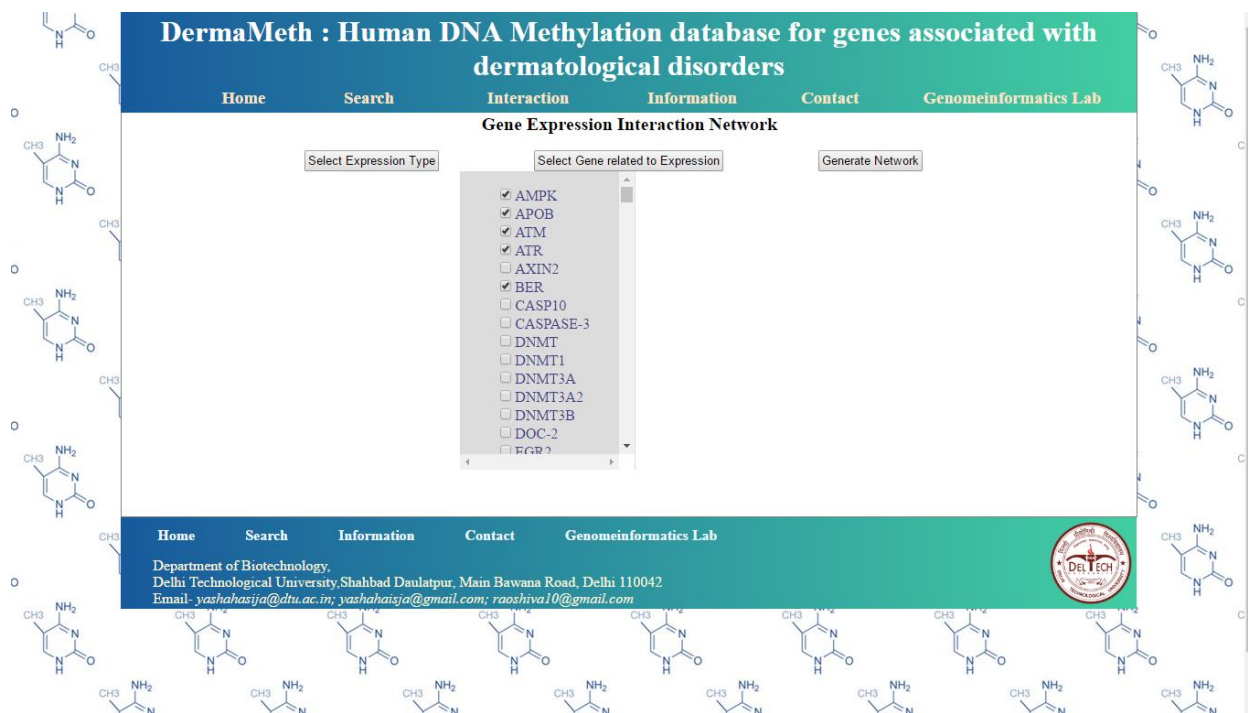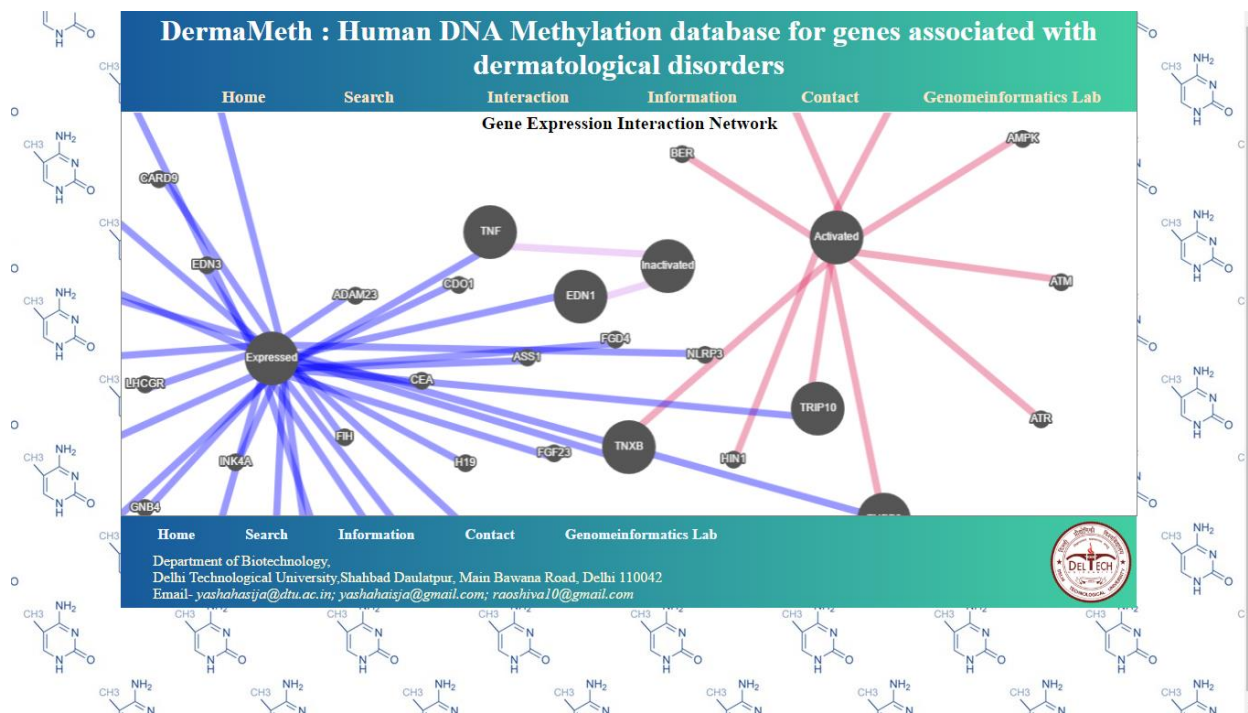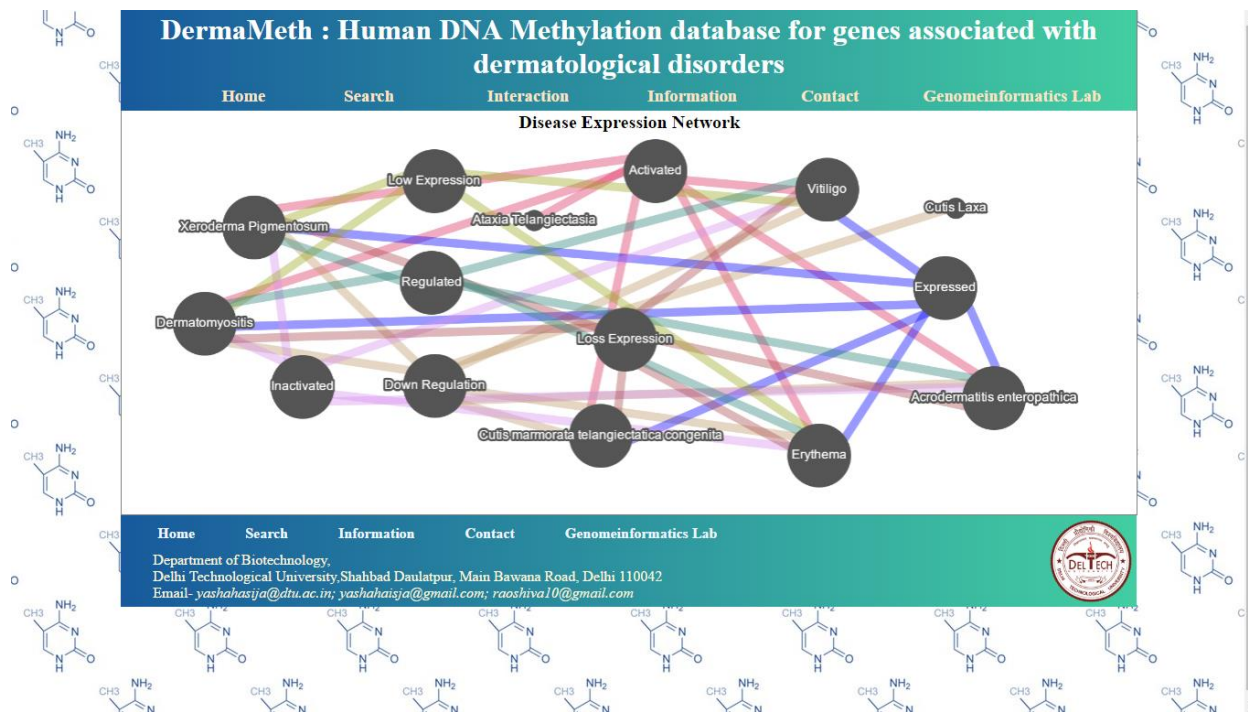

Figure 46- Result page for Disease- Expression interaction with different color indicating different Disease Expression.

Table 1 Disease categorized into different Class

Disease Class	Disease Name
ABSCESSES	Granulomatous Disease, Chronic
	Chronic Granulomatous Disease
	Granulomatous Disease, Chronic, X-Linked
	Job Syndrome
	Tyrosine Kinase 2 Deficiency
ACANTHOSIS NIGRICANS	Crouzon Syndrome With Acanthosis Nigricans
	Cutis Gyrata Syndrome of Beare And Stevenson
	Prader-Willi Syndrome
ALBINISM	Hermansky Pudlak syndrome 2
	Oculocutaneous albinism type 2
ANGIOKERATOMA	Fucosidosis
ANKYLOBLEPHARON FILIFORME ADNATUM	Popliteal Pterygium Syndrome
CAFÉ-AU-LAIT SPOTS	NOONAN SYNDROME 7
	Noonan Syndrome
	NOONAN SYNDROME-LIKE DISORDER
	Noonan syndrome 3
	Noonan Syndrome 6
	Noonan Syndrome 2
	Noonan Syndrome 5
	NOONAN SYNDROME-LIKE DISORDERHAIR
	Noonan Syndrome 4
CALCINOSIS CUTIS	Dermatomyositis
CALCINOSIS CUTIS/OSTEOMA CUTIS	Pilomatrixoma
	Osseous Heteroplasia, Progressive
COLLODION MEMBRANE	Ichthyosis, Lamellar
	Lamellar ichthyosis, type 2
	Gaucher Disease
	Lamellar ichthyosis, type 3
	Gaucher Disease, Type Iiic
	Gaucher Disease, Perinatal Lethal
	Ichthyosis, Lamellar, 5
	Trichothiodystrophy Syndromes
	Turner Syndrome
CYSTS	Hidradenitis Suppurativa
ECZEMA DERMATITIS	Methylmalonic acidemia
	Ichthyosis Vulgaris
	Wiskott-Aldrich Syndrome
EDEMA/LYMPHEDEMA	Lymphedema, Hereditary, II
	LYMPHEDEMA, HEREDITARY, IC
	Lymphedema, Hereditary, IB
	Melkersson-Rosenthal Syndrome
	Erythromelalgia
ELASTIC SKIN	Cutis Laxa
ELASTOSIS PERFORANS SERPIGNOSA	Down Syndrome
	Marfan Syndrome
EROSIONS	Hay-Wells syndrome
ERYTHEMA	Hemochromatosis
	Port-Wine Stain
	Hemochromatosis, Type 2B
	Hemochromatosis, type 2
	Erythromelalgia
	Hemochromatosis, type 4
	Hemochromatosis, type 3
ERYTHRODERMA	Pityriasis Rubra Pilaris
FOLLICULAR	Chondrodysplasia punctata,
ATROPHODERMA	brachytelephalangic
	Chondrodysplasia punctata 2, X-linked dominant
	Chondrodysplasia Punctata, Rhizomelic
FRECKLES or LENTIGINES	Xeroderma Pigmentosum
GENERALIZED	Harlequin type ichthyosis
	Refsum Disease
GRANULOMAS	Blau syndrome
KERATOACANTHOMA	Muir-Torre Syndrome
KERATOSIS PILARIS	Cardiofaciocutaneous syndrome
KNUCKLE PADS	Dupuytren Contracture
LIPOMAS	MERRF Syndrome
LOCALIZED ABSENCE	Johanson Blizzard syndrome
MACULES	Carney Complex
	Fanconi Syndrome
	Tuberous Sclerosis
	Tuberous Sclerosis 1
	Tuberous Sclerosis 2
	WATSON SYNDROME
	LEOPARD Syndrome
	Proteus Syndrome
	Schimke immunoosseous dysplasia
MALIGNANCY, CUTANEOUS	Dyskeratosis Congenita
NA	NA
NODULES	Myotonic Dystrophy
	Xanthomatosis, Cerebrotendinous
	Albright's hereditary osteodystrophy
	Alagille Syndrome
	Steatocystoma Multiplex
	Cowden-Like Syndrome
PALMO or PLANTAR	Erythema
PAPULES	Hyalinosis, Systemic
	Darier Disease
	Costello Syndrome
	Bazex-Dupre-Christol syndrome
	Lipoid Proteinosis of Urbach and Wiethe
	Basal Cell Nevus Syndrome
	Pachyonychia Congenita
	Buschke-Ollendorff syndrome
	Focal Dermal Hypoplasia
	Keratosis Follicularis Spinulosa Decalvans
	Cowden-Like Syndrome
PATCHES	Vitiligo
	Klippel-Trenaunay-Weber Syndrome
	Angioma serpiginosum, X-linked
	Bloom Syndrome
	Gangliosidosis, GM1
	Sturge-Weber Syndrome
	Alkaptonuria
	Incontinentia Pigmenti
	Piebaldism
PHOTOSENSITIVITY	Porphyrias, Hepatic
	Porphyria, Acute Hepatic
	Porphyrias
	Porphyria Cutanea Tarda
	Coproporphyria, Hereditary
	Smith-Lemli-Opitz Syndrome

	Porphyria, South African type
	Porphyria, Acute Intermittent
	Porphyria, Erythropoietic
PLAQUES	Pseudoxanthoma Elasticum
	Epidermal Nevus
	Erythrokeratodermia Variabilis
	Werner Syndrome
	Hartnup Disease
SCARS	PROTOPORPHYRIA, ERYTHROPOIETIC, X-LINKED
	Protoporphyria, Erythropoietic
	Dysautonomia, Familial
	Dowling-Degos Disease
TELANGIECTASES	Ataxia Telangiectasia
	Cockayne Syndrome
	Hallermann's Syndrome
	ATAXIA-TELANGIECTASIA-LIKE DISORDER 1
TUMORS	Calcinosis
	Tumoral Calcinosis, Hyperphosphatemic, Familial

ULCERS	Tumoral Calcinosis, Normophosphatemic, Familial
	Cutis marmorata telangiectatica congenita
	Hereditary Sensory And Autonomic Neuropathy
VESICLES or BULLAE	Porphyria Cutanea Tarda
	Acrodermatitis enteropathica
	Epidermolysis Bullosa Dystrophica
	Transient bullous dermolysis of the newborn
	Epidermolysis Bullosa Simplex
	Ichthyosis Bullosa of Siemens
	Rothmund-Thomson Syndrome
	Ichthyosis
WHEALS	MUCKLE-WELLS SYNDROME
	Netherton Syndrome
XANTHOMA/XANTHELASMA	Hyperlipidemias
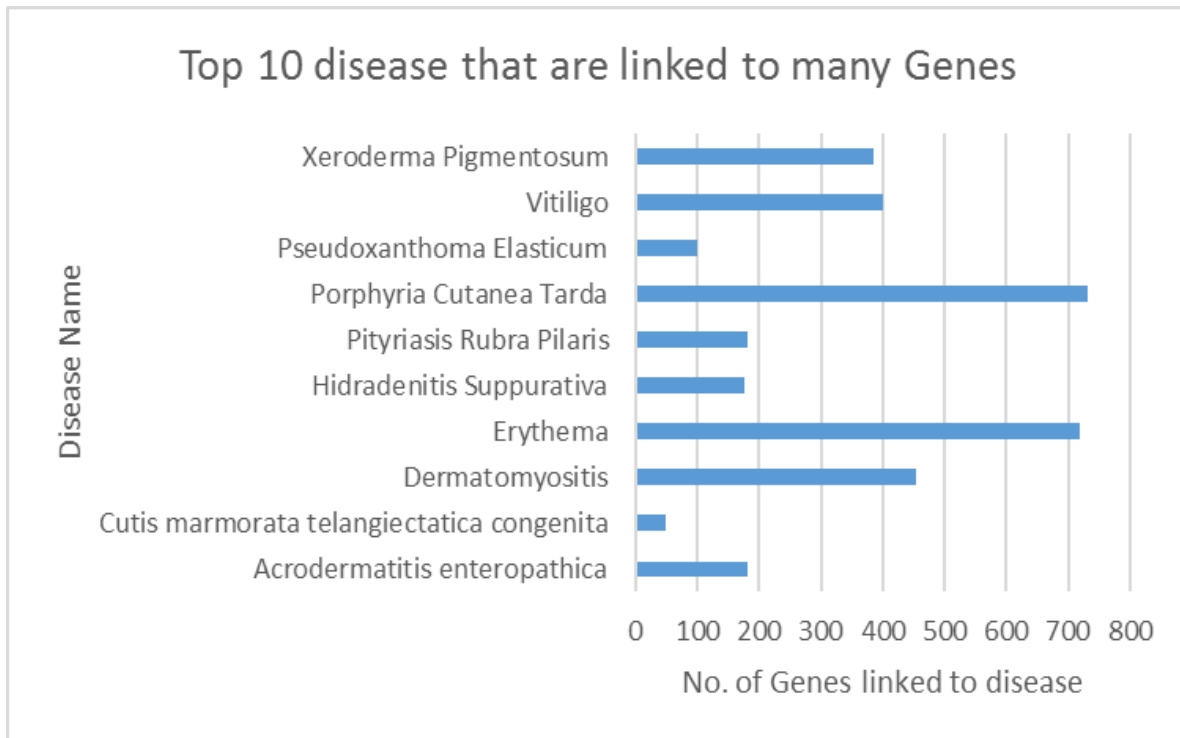
Figure 47- Top 10 diseases that are linked to various genes related to dermatological.
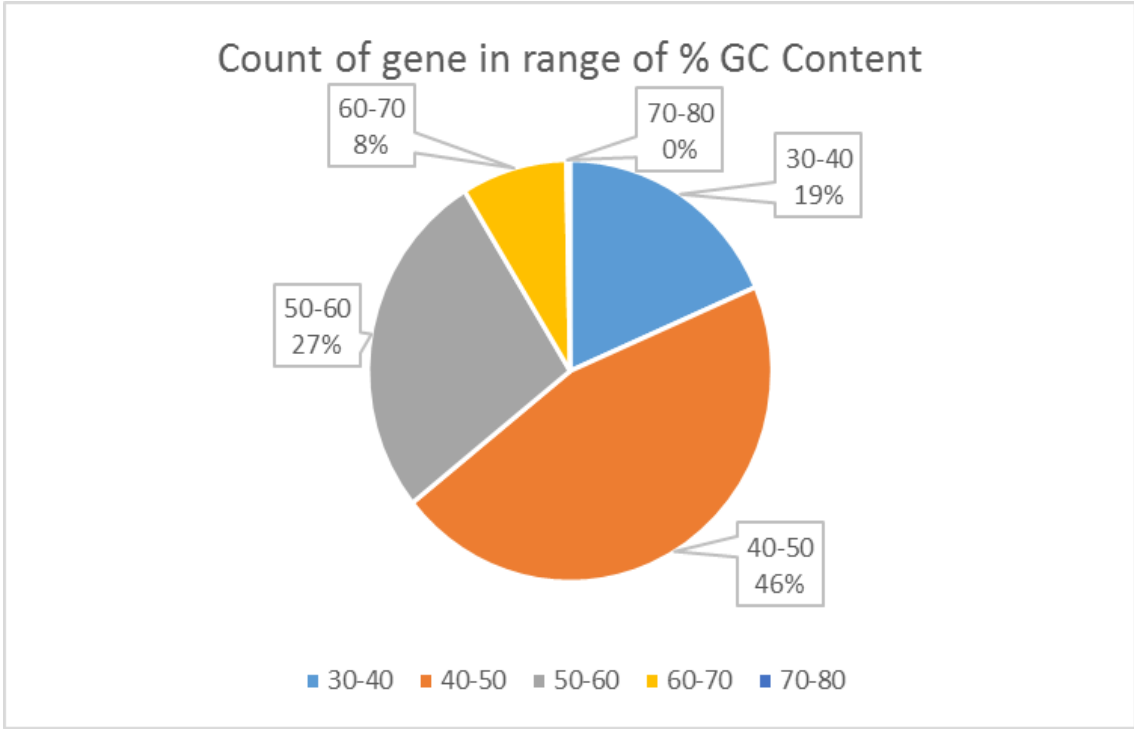
Figure 48- Count of the gene in a range of percentage GC content.
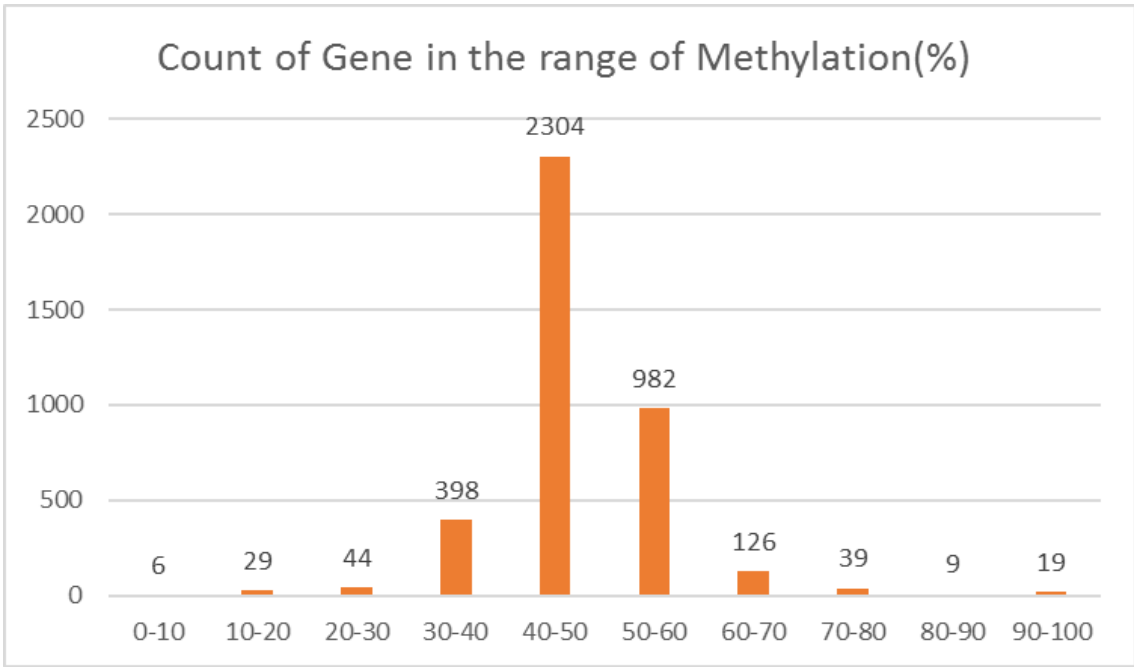
Figure 49- Count of Gene in the range of Methylation (percentage)

Figure 50- Top 30 pathway involved in various dermatological disorder

DISCUSSION AND CONCLUSION

The current database consists of more than 1000 genes that are associated with the dermatological disorder. These dermatological genes lead to various diseases such as Porphyria Cutanea Tarda, Erythema, Dermatomyositis, Vitiligo and Xeroderma Pigmentosum and many others. These diseases have been further classified into different 45 different disease classes.

MethSRD, SNiPSRD & InteractSRD is a first of its type repository that is developed, which is an user-interactive for giving the information of the genes associated with Dermatological diseases. This database enables the user to retrieve intricate methylation information of genes linked with dermatological diseases at the various level of genome-widespread level of systems. Mining for the database with biologically meaningful data is likely to divulge before indefinite evidences around the fundamental causes of disease related to dermatology.

For building up the database for DNA methylation that have a focus on human diseases to improve the methylation viewer for genomic and their functionality. For the importance behind the integrative analysis, we hereby will regularly collect data from new sources to enhance the analytical depths of the database. The database will be including more and more data sets and tools that can be used for the identification of disease or related DNA methylation that acts as a markers for specified genes using an integrated and differential identification tool for methylation prediction.

The principal data for the database MethSRD, SNiPSRD & InteractSRD represents as the association of methylation, SNiP with various Skin Related Disorders. Almost all the genetic disorders which are complex in nature are believed to be prevalent in skin which are considered as SRDs for the purpose of this work. The information on the database was entirely obtained from the reputed articles published in high-quality journals and various online medical forums discussing skin and their associated diseases. The data that were exists in the existing databases were carefully assessed, curated and corrected with proper reference to the original articles.

REFERENCES

1. Gal-Yam EN, Saito Y, Egger G, Jones PA (2008) 'Cancer epigenetics: modifications, screening, and therapy', Annual Review of Medicine, Vol.59, pp.267–280.

2. Portela A, Esteller M (2010) 'Epigenetic modifications and human disease', Nature Biotechnology, Vol. 28, pp.1057- 68. And Robertson KD (2005) 'DNA methylation and human disease', Nature Reviews Genetics Vol. 6, pp. 597-610.

3. Suzuki MM and Bird A (2008) 'DNA methylation landscapes: provocative insights from epigenomics', Nature Reviews Genetics, Vol. 9, pp. 465-476.

4. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, and Ye Z (2009) 'Human DNA methylomes at base resolution show widespread epigenomic differences', Nature Genetics, Vol. 462, pp.315-322

5. Bock C and Lengauer T (2008) 'Computational epigenetics. Bioinformatics', Vol. 24, No.1, pp.1–10

6. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Rongione M, and Webster M, et al. (2009) 'The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores' Nature Genetics, Vol. 41, pp.178-86

7. Johnson AA, Akman K, Calimport SR, Wuttke D, Stolzing A, de Magalhaes JP (2012) 'The role of DNA methylation in aging, rejuvenation, and age-related disease', Rejuvenation Research, Vol. 15, No.5, pp.483–494.

8. Fernandez AF, Assenov Y, Martin-Subero JI, Balint B, Siebert R, Taniguchi H, Yamamoto H, Hidalgo M, Tan A-C, Galm O, Ferrer I, Sanchez-Cespedes M, Villanueva A, Carmona J, et al (2012) 'A DNA methylation fingerprint of 1628 human samples', Genome Research, Vol. 2, pp.407-419

9. Yaping Lia, Amr H. Sawalhab, Qianjin Lua, (2009). Aberrant DNA methylation in skin diseases. Journal of Dermatological Science. Vol. 54, Issue 3, June 2009, P.No. 143-149.PMID: 19395242 DOI: 10.1016/j.jdermsci.2009.01.009

10. Riggs AD, Russo VE, Martienssen RA (1996). Epigenetic mechanisms of gene regulation. Plainview, N.Y.: Cold Spring Harbor Laboratory Press. ISBN 0-87969-490-4.

11. Bird, Adrian (2007). Perception of epigenetics. NATURE, Vol. 477, 24 May 2007, DOI: 10.1038/nature05913

12. Jone, Peter A, Takai Daiya (2001). The Role of DNA Methylation in Mammalian Epigenetics. Science 293, 1068 (2001). DOI: 10.1126/science.1063852

13. Simmons, D. (2008), Epigenetic influence and disease. Nature Education1(1):6)

14. Gal-Yam EN, Saito Y, Egger G, Jones PA (2008) 'Cancer epigenetics: modifications, screening, and therapy', Annual Review of Medicine, Vol.59, pp.267–280

15. Portela A, Esteller M (2010) 'Epigenetic modifications and human disease', Nature Biotechnology, Vol. 28, pp.1057- 68. And Robertson KD (2005) 'DNA methylation and human disease', Nature Reviews Genetics Vol. 6, pp. 597-610

16. Suzuki MM and Bird A (2008) 'DNA methylation landscapes: provocative insights from epigenomics', Nature Reviews Genetics, Vol. 9, pp. 465-476

17. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, and Ye Z (2009) 'Human DNA methylomes at base resolution show widespread epigenomic differences', Nature Genetics, Vol. 462, pp.315-322

18. Bock C and Lengauer T (2008) 'Computational epigenetics. Bioinformatics', Vol. 24, No.1, pp.1–10.

19. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Rongione M, and Webster M, et al. (2009) 'The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores' Nature Genetics, Vol. 41, pp.178-86.

20. Johnson AA, Akman K, Calimport SR, Wuttke D, Stolzing A, de Magalhaes JP (2012) 'The role of DNA methylation in aging, rejuvenation, and age-related disease', Rejuvenation Research, Vol. 15, No.5, pp.483–494

21. Fernandez AF, Assenov Y, Martin-Subero JI, Balint B, Siebert R, Taniguchi H, Yamamoto H, Hidalgo M, Tan A-C, Galm O, Ferrer I, Sanchez-Cespedes M, Villanueva A, Carmona J, et al (2012) 'A DNA methylation fingerprint of 1628 human samples', Genome Research, Vol. 2, pp.407-419

22. Virginia P. Sybert, M.D. (2010) Genetic Skin Disorders, Second Edition, Oxford University Press.

23. Davis,A.P., Murphy,C.G., Saraceni-Richards,C.A., Rosenstein,M.C., Wiegers,T.C. and Mattingly,C.J. (2009) Comparative Toxicogenomics Database: a knowledgebase and discovery tool for chemical-gene-disease networks. Nucleic Acids Res., 37, D786–D792

24. Davis,A.P.,Wiegers,T.C., Rosenstein,M.C., Murphy,C.G. and Mattingly,C.J. (2011) The curation paradigm and application tool used for manual curation of the scientific literature at the Comparative Toxicogenomics Database. Database, doi:10.1093/database/bar034

25. Janet Pinero, Alex Bravo, Nuria Queralt-Rosinach, Alba Gutierrez-Sacristan,Jordi Deu-Pons, Emilio Centeno, Javier Garcia-Garcia, Ferran Sanz and Laura I. Furlong* (2017). Nucleic Acids Research, 2017, Vol. 45, Database issue D833–D839 doi: 10.1093/nar/gkw943

26. Liu H., Liu W., Liao Y., Cheng L., Liu Q., Ren X., Shi L., Tu X., Wang Q.K. and An-Guo Y.; CADgene: a comprehensive database for coronary artery disease genes. Nucleic Acids Research. 39: 991-996; 2010.

27. Kevin G Bexker, Kathleen C Barnes, Tiffani J Bright and S Alex Wand. The Genetic Association Database is a database of genetic association data from complex diseases and disorders. (2014). Nature Genetics Vol. 36, No. 5, May 2004, P.No. 431-32

28. Kuo HC, Lin PY, Chung TC, Chao CM, Lai LC, Tsai MH, Chuang EY.(2011).DBCAT: database of CpG islands and analytical tools for identifying comprehensive methylation profiles in cancer cells. Journal Computational Biology 2011 Aug;18(8):1013-7. doi: 10.1089/cmb.2010.0038. PMID:21214365

29. Wei-Yun Huang, Sheng-Da Hsu, Hsi-Yuan Huang, Yi-Ming Sun, Chih-Hung Chou, Shun-Long Weng and Hsien-Da Huang (2015). MethHC: a database of DNA methylation and gene expression in human cancer. Nucleic Acids Research, 2015, Vol. 43, D856–D861, doi: 10.1093/nar/gku1151.

30. Arwa Bin Raies, Hicham Mansour, Roberto Incitti and Vladimir B. Bajic (2015). DDMGD: the database of text-mined associations between genes methylated in diseases from different species. Nucleic Acids Research, 2015, Vol. 43, Database issue D879–D886 doi: 10.1093/nar/gku1168

31. Jie Lv, Hongbo Liu, Jianzhong Su, Xueting Wu, Hui Liu, Boyan Li, Xue Xiao, Fang Wang, Qiong Wu,and Yan Zhang (2012). DiseaseMeth: a human disease methylation database. Nucleic Acids Research, 2012, Vol. 40, D1030–D1035 doi:10.1093/nar/gkr1169

32. Ximiao He, Suhua Chang, Jiajie Zhang, Qian Zhao, Haizhen Xiang, Kanthida Kusonmano, Liu Yang, Zhong Sheng Sun, Huanming Yang and Jing Wang (2008). MethyCancer: the database of human DNA methylation and cancer. Nucleic Acids Research, 2008, Vol. 36, D836–D841, doi:10.1093/nar/gkm730

33. Alan P. Boyle, Eurie L. Hong, Manoj Hariharan, Yong Cheng, Marc A. Schaub, Maya Kasowski, Konrad J. Karczewski, Julie Park, Benjamin C. Hitz, Shuai Weng, J. Michael Cherry, and Michael Snyder (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Research 22:1790–1797, 2012, Published by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/12.

34. Cathy H. Wu, Rolf Apweiler, Amos Bairoch, Darren A. Natale, Winona C. Barker (2006).The Universal Protein Resource (UniProt): an expanding universe of protein information. Nucleic Acids Research, 2006, Vol. 34, Database issue D187–D191 doi:10.1093/nar/gkj161

35. Diestel, R.: Graph theory. Springer-Verlag, Heidelberg (2005), ISBN 3-540-26182-6

36. Erdős, P., Renyi, A.: Publ Math Inst Hung Acad Sci **5** (1960) 17

37. Barabási, A.L., Oltvai, Z.N.: Network biology: understanding the cell's functional organization. Nat Rev Genet **5** (2004) 101-113

38. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks Paul Shannon,1 Andrew Markiel,1 Owen Ozier,2 Nitin S. Baliga,1 Jonathan T. Wang,2 Daniel Ramage,2 Nada Amin,2 Benno Schwikowski,1,5 and Trey Ideker2,3,4,5 1 Institute for Systems Biology, Seattle, Washington 98103, USA; 2 Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA; 3 Department of Bioengineering, University of California–San Diego, La Jolla, California 92093, USA

39. Endy, D. and Brent, R. 2001. Modelling cellular behaviour. Nature409: 391–395.

40. Loew, L.M. and Schaff, J.C. 2001. The Virtual Cell: A software environment for computational cell biology. Trends Biotechnol. 19: 401–406

41. Mendes, P. 1997. Biochemistry by numbers: Simulation of biochemical pathways with Gepasi 3. Trends Biochem.Sci. 22: 361–363.

42. Virginia P. Sybert, M.D. (2010) Genetic Skin Disorders, Second Edition, Oxford University Press

43. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, King BL, McMorran R, Wiegers J, Wiegers TC, Mattingly CJ. The Comparative Toxicogenomics Database: update 2017

44. Janet Pinero, Alex Bravo, Nuria Queralt-Rosinach, Alba Gutierrez-Sacristan,Jordi Deu-Pons, Emilio Centeno, Javier Garcia-Garcia, Ferran Sanz and Laura I. Furlong* (2017). Nucleic Acids Research, 2017, Vol. 45, Database issue D833– D839 doi: 10.1093/nar/gkw943

PUBLICATIONS

Mottadi Shiva, Yasha Hasija. (2017). SRDomics: Human DNA Methylation database for genes associated with dermatological-related disorders. IEEE/ACM Transactions on Computational Biology and Bioinformatics. (**Communicated**)