VIRdb: a comprehensive database for interactive analysis of genes/proteins involved in the pathogenesis of vitiligo

Priyansh Srivastava, Alakto Choudhury, Mehak Talwar, Sabyasachi Mohanty, Priyanka Narad and Abhishek Sengupta

Amity Institute of Biotechnology, Amity University, Uttar Pradesh, India

Corresponding authors
Priyanka Narad, pnarad@amity.edu
Abhishek Sengupta,
asengupta@amity.edu

ABSTRACT

Vitiligo is a chronic asymptomatic disorder affecting melanocytes from the basal layer of the epidermis which leads to a patchy loss of skin color. Even though it is one of the neglected disease conditions, people suffering from vitiligo are more prone to psychological disorders. As of now, various studies have been done in order to project auto-immune implications as the root cause. To understand the complexity of vitiligo, we propose the Vitiligo Information Resource (VIRdb) that integrates both the drug-target and systems approach to produce a comprehensive repository entirely devoted to vitiligo, along with curated information at both protein level and gene level along with potential therapeutics leads. These 25,041 natural compounds are curated from Natural Product Activity and Species Source Database. VIRdb is an attempt to accelerate the drug discovery process and laboratory trials for vitiligo through the computationally derived potential drugs. It is an exhaustive resource consisting of 129 differentially expressed genes, which are validated through gene ontology and pathway enrichment analysis. We also report 22 genes through enrichment analysis which are involved in the regulation of epithelial cell differentiation. At the protein level, 40 curated protein target molecules along with their natural hits that are derived through virtual screening. We also demonstrate the utility of the VIRdb by exploring the Protein–Protein Interaction Network and Gene–Gene Interaction Network of the target proteins and differentially expressed genes. For maintaining the quality and standard of the data in the VIRdb, the gold standard in bioinformatics toolkits like Cytoscape, Schrödinger's GLIDE, along with the server installation of MATLAB, are used for generating results. VIRdb can be accessed through "http://www.vitiligoinfores.com/".

INTRODUCTION

Vitiligo is a chronic asymptomatic disorder affecting melanocytes from the basal layer of the epidermis which leads to a patchy loss of skin color (*Rodrigues et al., 2017*). It affects 0.5–2% of the world's population (*Ezzedine & Silverberg, 2016*). The lesions can appear in different shapes and sizes on any visible part of the body (*Tarlé et al., 2014*). Even though it is one of the neglected diseases condition people suffering from vitiligo are more

prone to psychological disorders and the risk of developing hearing problems (*Ezzedine et al., 2015*). The auto-immune inflammation theory remains to be the most acceptable, accounting for its development via an adaptive immune response (*Boniface et al., 2017*). According to biochemical theory, accumulation of Reactive Oxygen species (ROS) in the melanocytes remains the pioneer, for activation of CD8+ T cells against melanocytes (*Wang, Li & Li, 2019*; *Deng et al., 2019*). Previous studies have also reported that HIF-1-alpha, F2RL-1 and PIK3CB over-expression in CD8+ cells (*Deng et al., 2019*). HIF-1-alpha plays a crucial part in ROS accumulation response, whereas the F2RL-1 and PIK3CB are responsible for an immune response with CD8+ T cells (*Deng et al., 2019*). Due to loss of melanocytes in individuals affected by vitiligo, one can also expect an increased incidence of non-melanoma cancer and actinic keratose (*Zhang et al., 2016*). Arresting the progression of lesions and repigmentation of formed lesions remains to be a crucial strategy for treatment (*Bishnoi & Parsad, 2018*). AK-STAT inhibitors have demonstrated promising outcomes in the treatment of vitiligo, including repigmentation in topical applications. JAK inhibitor, tofacitinib and JAK-1,2 inhibitor, ruxolitinib, have been tested positive in causing repigmentation (*Bishnoi & Parsad, 2018*).

To understand the complexity of the vitiligo, we envisage a comprehensive online repository focusing on the systematically derived and manually curated potential protein targets along with their prospective hits derived from the natural compounds databases. To fully understand the genetic complexity of the vitiligo we also aim to present a genetic interaction network of the differentially expressed genes, in Lesional Vitiligo, Peri-Lesional Vitiligo, and Non-Lesional Vitiligo against healthy controls. The predicted differentially expressed genes across samples were then validated using gene enrichment analysis with Kyoto Encyclopedia of Gene and Genome (KEGG) and Gene Ontology (GO) (*Kanehisa et al., 2020*; *The Gene Ontology Consortium, 2007*). The Vitiligo Information Resource (VIRdb) systematically harbors all the information relating to vitiligo in terms of Potential protein targets and differentially expressed genes in a user-friendly database. As more than 50% of the approved drugs in DrugBank are natural compounds, therefore, we took the liberty of virtually screening all the natural compounds present in Natural Product Activity and Species Source Database (NPASS) against the potential protein targets (*Wishart et al., 2017*; *Zeng et al., 2017*). For maintaining, the quality and standard of the data in the VIRdb, the gold standard in bioinformatics toolkits like Cytoscape, Schrödinger's GLIDE, along with the server installation of MATLAB are used (*Shannon, 2003*; *Schrödinger, 2020*; *MATLAB, 2019*). The screened library can be used to carry wet-lab experiments based on potential hits which would thereby lead to considerably less consumption of time and other resources to carry out the protocol for drug development. We envision that VIRdb will be pertinent for the researchers and clinicians engaged in drug development for vitiligo.

METHODS

Microarray gene expression analysis

Gene Expression Omnibus was manually queried using "("vitiligo" (MeSH Terms) OR vitiligo (All Fields)) AND ("humans" (MeSH Terms) OR "*Homo sapiens*" (Organism) OR

homo sapiens (All Fields))". The returned results were manually curated for microarray experiments having healthy controls against affected individuals. We selected GSE65127 as it comprises of 10 healthy controls against 10 conditions (Lesional, Peri-lesional, Non-lesional). The CEL files of 40 samples having expression data of Lesional, Peri-lesional, Non-lesional against healthy control on Affymetrix GPL570 platform were isolated (*Clough & Barrett, 2016*; *Regazzetti et al., 2015*) data. Combinations of {Healthy vs (Lesional, Peri-lesional, Non-lesional)} were processed using the server installation of *MATLAB (2019)* by applying the GC Robust Multichip Averaging (GCRMA) method with Empirical Bayes (*MATLAB, 2019*; *Wu et al., 2004*). GCRMA empirical Bayes is the posterior mean estimate which is calculated through S, that represents a quantity proportional to the concentration of transcripts in the hybridization mixture (Eq. 1). After the raw reads from CEL files are converted to expression data through GCRMA, the probe sets with un-annotated gene symbols were dropped programmatically. A *t*-test is performed with each gene expression value to study the relevant changes in the expression values. False Discovery Rate (FDR) for each test situation was computed using Storey–Tibshirani procedure (*Storey & Tibshirani, 2003*). To compute the biological expression significance, −log10 of *p*-values was calculated and plotted against log fold change (2 folds).

$$s = E[s|S > 0, \text{PM}, \text{MM}], \quad \text{with } s = \log(S) \tag{1}$$

Gene–gene interaction network construction

For designing the gene–gene interaction network for differentially expressed genes, GeneMania plugin from Cytoscape was used (*Shannon, 2003*; *Mostafavi et al., 2020*). Initially, 3 distinct networks for differentially expressed genes (Expressed in lesional, Expressed in non-lesional, and Expressed in peri-lesional) were made. After that, all the networks are merged, with common interacting gene nodes in the Cytoscape environment (*Shannon, 2003*). The tangential nodes were manually deleted and only the highly connected nodes were taken for further analysis. The predicted co-expressed genes from GeneMania interaction network were also taken for enrichment analysis (*Mostafavi et al., 2020*).

Gene Ontology (GO) and pathway analysis

ShinyGO server was used for the pathway enrichment analysis of the single expression conditions (Lesional, Peri-Lesional and Non-lesional, each distinctly) and unique differentially expressed genes (significant expression at least once across the conditions) and predicted GGI genes from GeneMania (*Mostafavi et al., 2020*; *Ge, Jung & Yao, 2019*). The differentially expressed genes were enriched in Kyoto Encyclopedia of Gene and Genome database (KEGG) (*Kanehisa et al., 2020*). The false discovery rate of 0.01 was used as the cut off to elute out the most significant pathways. For GO enrichment analysis the GO database was recruited from the ShinyGO server and GO enrichment analysis was functioned with the false discovery rate of <0.01 (*The Gene Ontology Consortium, 2007*; *Mostafavi et al., 2020*).

Biological data collection

The Pubmed Engine was queried using "((Vitiligo) AND (Proteins)) AND (Protein–Protein Interaction)" for retrieval of literature relating to the protein targets (*Motschall & Falck-Ytter, 2005*). The query returned 4 abstracts, out of which study that includes the systemic approach for the construction of Protein–Protein Interaction Network (PPI) was taken into consideration (*Malhotra et al., 2017*). PPI network of 215 proteins was revitalized using Cytoscape 3.7 and String Database (*Shannon, 2003*; *Mering, 2003*; *Malhotra et al., 2017*). A total of 50 protein targets were eluted out of the set of reported 64 vitiligo core proteins in the study (*Malhotra et al., 2017*). The sub-network of 50 proteins was created manually based on coherence of centrality among 64 nodes (*Malhotra et al., 2017*). Finally considering the on the availability of X-ray crystallographic structures in Protein Data Bank 40 protein crystal structures were taken for molecular docking out of 50 nodes (Appendix I) (*Berman, 2000*; *Malhotra et al., 2017*). All the potential protein targets were annotated manually with their respective UniProt ID, PDB ID, STRING ID, KEGG ID, encoding Gene Symbol and target active amino acids from the scientific literature (*Kanehisa et al., 2020*; *Mering, 2003*; *Berman, 2000*; *The UniProt Consortium, 2016*). Highest-resolution structures were manually curated and confiscated for Molecular docking studies for the protein structures that comprises multiple deposits in the Protein Data Bank (*Wlodawer et al., 2013*). For identification of the active amino acids, each PDB profile was manually curated in reference to its literature (Appendix I).

Virtual screening against NPASS

As more than 50% of the approved drugs in DrugBank are natural compounds, therefore, we took the liberty of virtually screening all the natural compounds present in Natural Product Activity and Species Source Database (NPASS) against the potential protein targets (*Wishart et al., 2017*; *Zeng et al., 2017*). The NPASS database consists of 25,041 natural compounds (*Zeng et al., 2017*). Every structure was neutralized at the physiological pH and a library of 36,229 potential ligands was created including the stereoisomers and tautomers (*Srivastava, Pal & Misra, 2018*). Target protein structures files were prepared using Schrödinger Protein Preparation Wizard by removing excess waters and were minimized using OPLS 2005 force field (*Clark et al., 2019*; *Sweere & Fraaije, 2017*). Receptor grid of dimensions of $15 \times 15 \times 15$ angstroms was generated around the active amino acids. Glide HTVS pipeline was used to carry out virtual screening of the natural compound library to each target protein. Ligands were sorted and ranked based on the Glide score as it is empirical of various computed scores and is also quite accurate in wet-lab experimentation (*Greenwood et al., 2010*).

Architecture and design

Vitiligo Information Resource was formulated using Apache HTTP server on Linux Platform. MySQL was used to store the information as it offers fast and easy query processing because of its RDBMS architecture (*Mathur et al., 2016*). HTML 5 and Cascading Style Sheets were used to design the front-end interface of the VIRdb. PHP 7

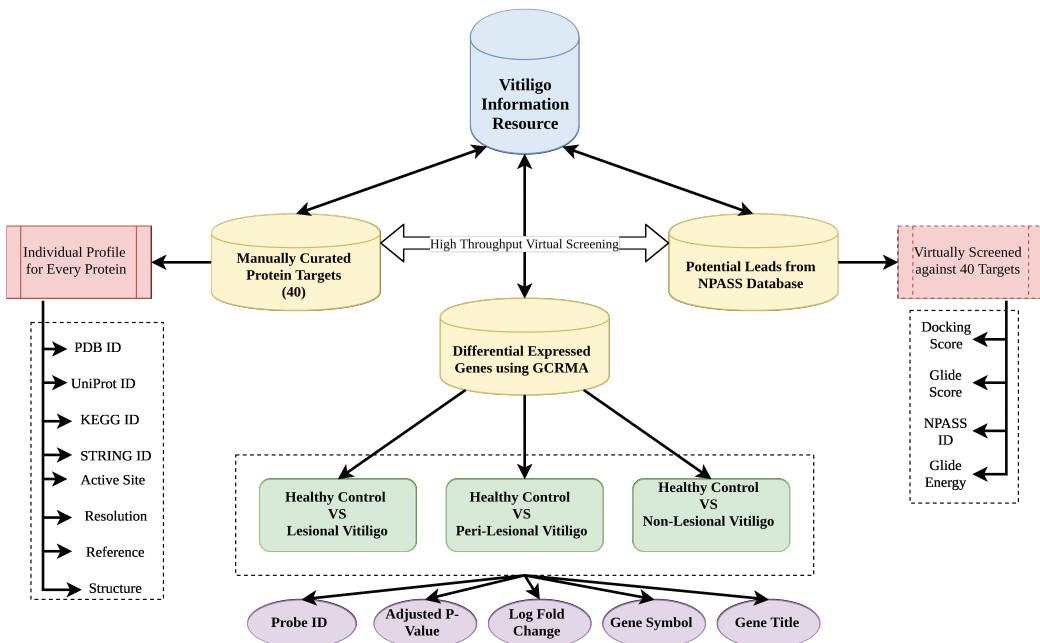



Figure 1 **Database architecture.** Architecture of VIR database.

was used for maintaining the dynamic connectivity between the front-end and back-end. Each entry in the database was connected to their parent database via html anchors. The complete architectural schema of vitiligo information resource db is depicted in Fig. 1.

RESULTS

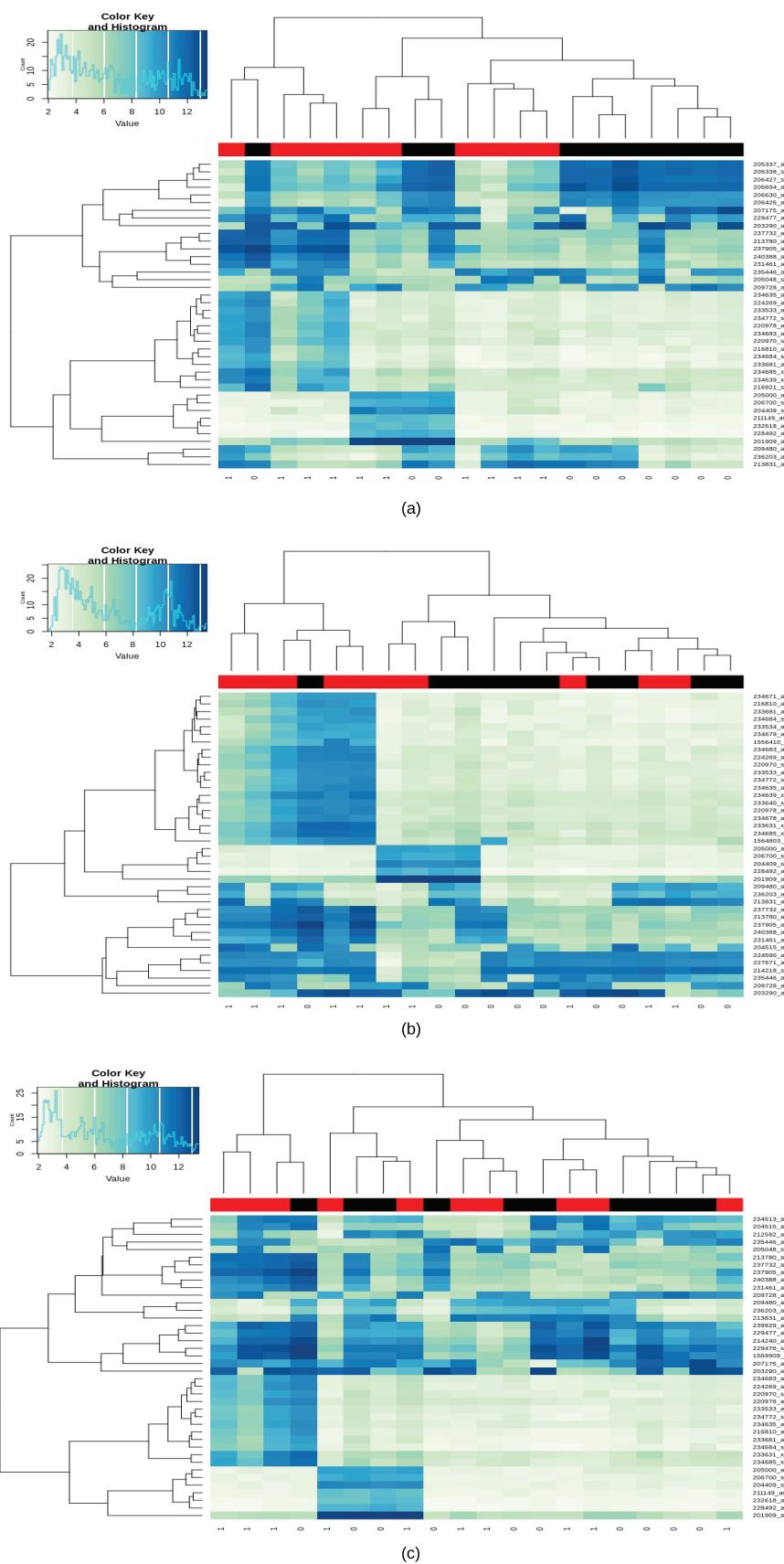
Differentially expressed genes

A total of 123 differential genes for lesional vitiligo against healthy controls. Out of 123 genes, 52 genes were predicted to be as under-expressed and 70 genes were predicted to be over-expressed in accordance with the LogFc filter (Appendix II; Fig. 2A). In the case of Non-Lesional vs healthy control, a total of 27 upregulated differential genes were identified (Appendix III; Fig. 2B). For Peri-Lesional vitiligo, a total of 32 upregulated genes were curated (Appendix IV; Fig. 2C). Upon, a union of all the sets (Differential Genes, Lesional, Non-Lesional, and Peri-Lesional) as a total of 129 differential genes have been curated. Out of which 52 were under-expressed and 76 were over-expressed (Appendix V).

GO and pathway enrichment analysis

The DCT, EDNRB, TYR, TYRP1, and WNT6 that were from the set of 129 unique differentially expressed genes across the samples were enriched in Melanogenesis pathway with an FDR of 0.0038 in KEGG (*Kanehisa et al., 2020*) (Appendix VI). Upon GO enrichment 6 genes (GPR143, OCA2, TYR, DCT, TYRP1, SLC45A2) were enriched in Melanosome membrane (Cellular Compartment), 7 genes (TYR, DCT, TYRP1, OCA2, SLC45A2, PMEL, CDH3) were enriched in Melanin metabolic process (Biological Process) (*The Gene Ontology Consortium, 2007*) (Appendix VII). Surprisingly, 2 genes were

(a)

(b)

(c)
Figure 2 Heatmap for lesional, non-lesional and peri-lesional conditions. Heatmap of differentially expressed gene expression values with black bars representing the Healthy/control samples and red being the case/diseased samples; (A) heatmap of differentially expressed gene expression values in lesional vitiligo against Healthy controls; (B) heatmap of differentially expressed gene expression values in non-lesional vitiligo against healthy controls; (C) heatmap of differentially expressed gene expression values in peri-lesional vitiligo against healthy controls.

enriched with monophenol monooxygenase activity (molecular function) and 2 genes were enriched in oxidoreductase activity (molecular function), which could account for the oxidative stress-induced melanocyte degradation (*Wang, Li & Li, 2019*; *Deng et al., 2019*) (Appendix VII) (For enrichment analysis of the single expression conditions (Lesional, Peri-Lesional and Non-lesional, each distinctly) see Supplements 1–3).

Gene–gene interaction network

We upvote 22 genes that were predicted while construction of common network (ACYP1, APLP2, ARNTL, CEBPG, CPEB2, DIP2A, HNRNPA3, HNRNPA3P1, HTATSF1, JKAMP, MAFG, MFN2, MKL1, ORC2, PAGR1, PDPK1, PEX5, PRKX, PRKY, TLE4, TMEM70, USP47) with GeneMania (*Shannon, 2003*, Fig. 3; Appendix VIII). Upon gene enrichment analysis of these 22 genes with GO db MAFG, PRKX, ARNTL were enriched in Regulation of epithelial cell differentiation as the biological process (*The Gene Ontology Consortium, 2007*). These genes were never reported in the literature for these types of vitiligo therefore, might be relevant in understanding the molecular aspects of vitiligo.

Data browsing and web interface

The main browsing window of the VIRdb holds the links to Target proteins, Potential hits and differential genes (Fig. 4A). The end-user can either browse any of the sections or can directly download the data in CSV file format. The network files are only available for download and can be downloaded in SIF of the Cytoscape 3.7. The target protein toggle directs the user to the comprehensive list of proteins that were used as targets for the molecular docking studies (*Shannon, 2003*). The other two toggles of potential hits and differential genes direct the user towards the dynamic tables where the user can select from the types of vitiligo and target specific drugs. The target proteins have individual profiles as shown in Fig. 4B, which comprises their active site residue references along with the potential hits against it. The identifiers (STRING ID, KEGG ID, PDB Id, SwissProt Id) of the target proteins are connected to their respective databases through the HTML anchors (*Kanehisa et al., 2020*; *Mering, 2003*; *Berman, 2000*; *The UniProt Consortium, 2016*). The encoding genes for the proteins are connected to the GeneCards database for maintaining the quality of data (*Safran et al., 2010*). The user can directly download the molecular docking results (NPASS ids, Docking Scores, Glide gscores, and Glide energy) in CSV file format from the individual protein profiles for further analysis (*Zeng et al., 2017*). The active site reference of the protein is depicted in a tabular format depicting the amino residue name, chain and position in the sequence. Potential hits can be browsed through dynamic tables and frames. The user can select the protein's PDB identifier from below the navigation pane and can dynamically see the tailored results for

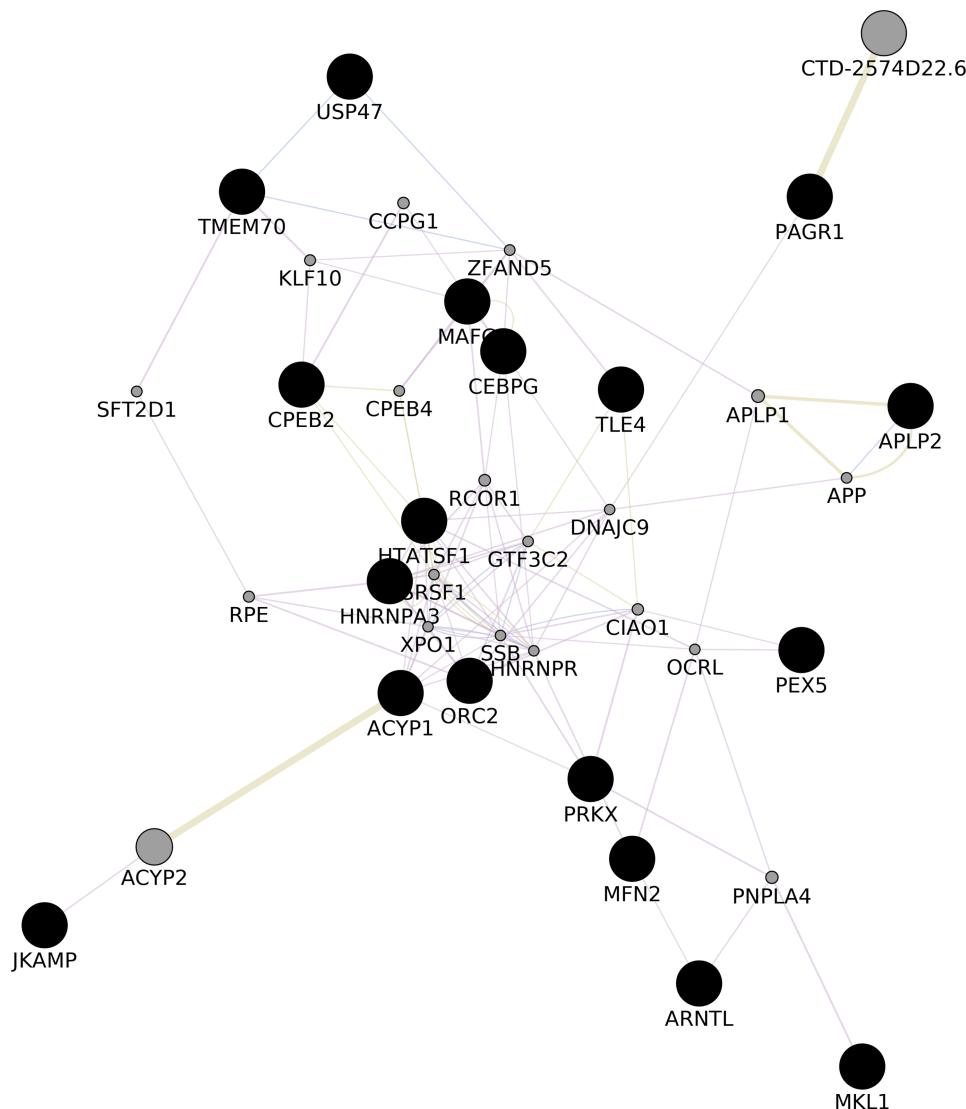



Figure 3 GGI Network. Gene–gene Interaction network of the predicted set of differential genes in lesional, non-lesional, and peri-lesional condition through the set of 129 unique genes; co-expressed genes are shown in gray nodes; constructed using Cytoscape 3.7 (*Shannon, 2003*).

the potential hits against them. The tables contain information about the NPASS db identifier, Glide Gscore, Docking Score and Glide energy, that correlates to the binding affinity of the ligand with the target. Each NPASS identifier is connected to the NPASS db entry of the respective compound (*Zeng et al., 2017*). The user can also download the results from this page in CSV files formats for further analysis. Differential genes that are predicted using *MATLAB (2019)* are also projected in a dynamic fashion. The end-user can see the genes by selecting the appropriate toggle from below the navigation bar. The table consists of Probe Ids, LogFc, Gene Symbol, Adjusted *p*-values and gene descriptions. Each gene symbol is connected to the GeneCard entry through HTML anchors (*Safran et al., 2010*). LogFc values with negative signs denote the downregulated genes while the positive

Figure 4 VIRdb Interface. (A) Browsing Interface of VIRdb and (B) individual protein profile.

ones account for the upregulation. The probe IDs of the genes are connected to Harvard's Gene Enrichment Profiler (*Bateman et al., 2014*).

Discussion and community impact

As per the results of KEGG enrichment analysis, the differential genes were enriched in the pathway of Melanogenesis. WNT6 being one of the enriched genes has been previously reported for vitiligo therapeutics and wound healing through melanocyte stem cell differentiation (*Yamada et al., 2013*; *Sun et al., 2018*). Moreover, DCT, TYR, EDNRB and TYRP1 have also been reported previously playing a vital role in vitiligo (*Salinas-Santander et al., 2018*; *Jin et al., 2010*; *Takeo et al., 2016*). Therefore, other differential genes which were enriched along with WNT6, DCT, TYR, EDNRB and TYRP1 by GO could provide more insights on vitiligo prognosis (Appendix VII). We also report ACYP1, APLP2, ARNTL, CEBPG, CPEB2, DIP2A, HNRNPA3, HNRNPA3P1, HTATSF1, JKAMP, MAFG, MFN2, MKL1, ORC2, PAGR1, PDPK1, PEX5, PRKX, PRKY, TLE4, TMEM70, USP47 for these types of vitiligo (Lesional, Non-lesional, and Peri-Lesional) therefore might be relevant in understanding the molecular aspects of vitiligo. VIRdb, integrates both drug-target and systems approach to produce a comprehensive repository entirely devoted to Vitiligo. Having curated information at both protein level and gene level along with potential therapeutics hits makes VIRdb first and one of its kind disease-specific databases in the scientific community. VIRdb holds manually curated protein targets that are involved in the vitiligo based on the systems approach, these protein target structure files are minimized and optimized using OPLS 2005 force fields and can be downloaded from the database for carrying out in-silico analysis. VIR also offers a complete profile for these target protein having links to all major databases

(KEGG db, STRING db, GeneCards db, SwissProt db, NPASS db) making it a cross-functional database as well (*Kanehisa et al., 2020*; *Mering, 2003*; *The UniProt Consortium, 2016*; *Safran et al., 2010*).

CONCLUSION

Lesional Vitiligo, Peri-Lesional vitiligo, and Non-Lesional vitiligo being the most common types of vitiligo have been thoroughly studied for differentially expressed genes and are descriptively integrated into the compendium of VIRdb. All of which are validated through KEGG and GO db, having pathways in functions in melanogenesis (*Kanehisa et al., 2020*; *The Gene Ontology Consortium, 2007*). It is also the home for the potentially active natural compounds that are virtually screened against the potential protein targets. These compounds are curated from NPASS db (*Zeng et al., 2017*). It is an attempt to accelerate the drug discovery process and laboratory trials for vitiligo. We will update VIRdb iteratively at regular intervals in order to cover the domain of vitiligo as a diseased condition. We also provide users with an option to submit new data. Moreover, we will validate each new entry before its incorporation for maintaining a high level of quality of data. In future various improvizations on the VIRdb could be done like the inclusion of various other chemical databases like ZINC, DrugBank and PubChem (*Wishart et al., 2017*; *Irwin & Shoichet, 2005*; *Kim et al., 2015*).

ACKNOWLEDGEMENTS

P.N. and A.S. would like to acknowledge Dr. Ashok K. Chauhan, Founder President, Amity University, Uttar Pradesh for providing us the opportunity to conduct research and innovation. P.S., S.M., A.C., M.T., P.N. and A.S. would like to thank the Centre for Computational Biology and Bioinformatics, Amity Institute of Biotechnology, Amity University for providing them necessary resources. P.S., S.M., A.C. and M.T. would like to specifically acknowledge P.N. and A.S. for introducing the topic of the drug-target interaction network for vitiligo.

Funding

The authors received no funding for this work.

Competing Interests

The authors declare that they have no competing interests. Even though they have made an attempt to cover as much information as possible related to Vitiligo by manual curation and sophisticated bioinformatics approaches, it is possible that a few articles might not be incorporated with their search criteria.

Author Contributions

- Priyansh Srivastava conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Alakto Choudhury analyzed the data, prepared figures and/or tables, database Development, and approved the final draft.
- Mehak Talwar performed the experiments, prepared figures and/or tables, website Development, and approved the final draft.
- Sabyasachi Mohanty performed the experiments, prepared figures and/or tables, database Development, and approved the final draft.
- Priyanka Narad conceived and designed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Abhishek Sengupta conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The analysis results are available as Supplemental Files. The database is available at https://vitiligoinfores.com/.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.9119#supplemental-information.

REFERENCES

Bateman A, El-Hachem N, Beck A, Aerts H, Haibe-Kains B. 2014. Importance of collection in gene set enrichment analysis of drug response in cancer cell lines. *Scientific Reports* **4(1)**:1675 DOI 10.1038/srep04092.

Berman H. 2000. The protein data bank. *Nucleic Acids Research* **28(1)**:235–242 DOI 10.1093/nar/28.1.235.

Bishnoi A, Parsad D. 2018. Clinical and molecular aspects of vitiligo treatments. *International Journal of Molecular Sciences* **19(5)**:1509 DOI 10.3390/ijms19051509.

Boniface K, Seneschal J, Picardo M, Taïeb A. 2017. Vitiligo: focus on clinical aspects, immunopathogenesis, and therapy. *Clinical Reviews in Allergy & Immunology* **54(1)**:52–67 DOI 10.1007/s12016-017-8622-7.

Clark A, Negron C, Hauser K, Sun M, Wang L, Abel R, Friesner R. 2019. Relative binding affinity prediction of charge-changing sequence mutations with FEP in protein–protein interfaces. *Journal of Molecular Biology* **431(7)**:1481–1493 DOI 10.1016/j.jmb.2019.02.003.

Clough E, Barrett T. 2016. The gene expression omnibus database. In: Mathé E, Davis S, eds. *Statistical Genomics. Methods in Molecular Biology.* Vol. 1418. New York: Humana Press, 93–110.

Deng Q, Wei J, Zou P, Xiao Y, Zeng Z, Shi Y, Zhan Y, Zhang H, Tang B, Zeng Q, Xiao R. 2019. Transcriptome analysis and emerging driver identification of CD8+ T cells in patients with vitiligo. *Oxidative Medicine and Cellular Longevity* **2019(3)**:1–15 DOI 10.1155/2019/2503924.

Ezzedine K, Eleftheriadou V, Whitton M, Van Geel N. 2015. Vitiligo. *Lancet* **386(9988)**:74–84 DOI 10.1016/S0140-6736(14)60763-7.

Ezzedine K, Silverberg N. 2016. A practical approach to the diagnosis and treatment of vitiligo in children. *Pediatrics* **138(1)**:e20154126 DOI 10.1542/peds.2015-4126.

Ge S, Jung D, Yao R. 2019. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics. Bioinformatics* **36(8)**:2628–2629 DOI 10.1093/bioinformatics/btz931.

Greenwood J, Calkins D, Sullivan A, Shelley J. 2010. Towards the comprehensive, rapid, and accurate prediction of the favourable tautomeric states of drug-like molecules in aqueous solution. *Journal of Computer-Aided Molecular Design* **24(6–7)**:591–604 DOI 10.1007/s10822-010-9349-1.

Irwin J, Shoichet B. 2005. ZINC: a free database of commercially available compounds for virtual screening. *ChemInform* **36(16)**:16215 DOI 10.1002/chin.200516215.

Jin Y, Birlea S, Fain P, Gowan K, Riccardi S, Holland P, Mailloux C, Sufit A, Hutton S, Amadi-Myers A, Bennett D, Wallace M, McCormack W, Kemp E, Gawkrodger D, Weetman A, Picardo M, Leone G, Taïeb A, Jouary T, Ezzedine K, Van Geel N, Lambert J, Overbeck A, Spritz R. 2010. Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. *New England Journal of Medicine* **362(18)**:1686–1697 DOI 10.1056/NEJMoa0908547.

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2020. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* **45(D1)**:D353–D361 DOI 10.1093/nar/gkw1092.

Kim S, Thiessen P, Bolton E, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker B, Wang J, Yu B, Zhang J, Bryant S. 2015. PubChem substance and compound databases. *Nucleic Acids Research* **44(D1)**:D1202–D1213 DOI 10.1093/nar/gkv951.

Malhotra A, Jha M, Singh S, Pandey K. 2017. Construction of a comprehensive protein-protein interaction map for vitiligo disease to identify key regulatory elements: a systemic approach. *Interdisciplinary Sciences: Computational Life Sciences* **10(3)**:500–514 DOI 10.1007/s12539-017-0213-z.

Mathur D, Prakash S, Anand P, Kaur H, Agrawal P, Mehta A, Kumar R, Singh S, Raghava G. 2016. PEPlife: a repository of the half-life of peptides. *Scientific Reports* **6(1)**:36617 DOI 10.1038/srep36617.

MATLAB. 2019. *MATLAB for Artificial Intelligence.* Natick: The MathWorks Inc. Available at https://in.mathworks.com/.

Mering C. 2003. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Research* **31(1)**:258–261 DOI 10.1093/nar/gkg034.

Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. 2020. GeneMANIA: a real-time multiple association network integration algorithms for predicting gene function. *Genome Biology* **9(Suppl. 1)**:S4 DOI 10.1186/gb-2008-9-s1-s4.

Motschall E, Falck-Ytter Y. 2005. Searching the MEDLINE literature database through PubMed: a short guide. *Oncology Research and Treatment* **28(10)**:517–522 DOI 10.1159/000087186.

Regazzetti C, Joly F, Marty C, Rivier M, Mehul B, Reiniche P, Mounier C, Rival Y, Piwnica D, Cavalié M, Chignon-Sicard B, Ballotti R, Voegel J, Passeron T. 2015. Transcriptional analysis of vitiligo skin reveals the alteration of WNT pathway: a promising target for repigmenting vitiligo patients. *Journal of Investigative Dermatology* **135(12)**:3105–3114 DOI 10.1038/jid.2015.335.

Rodrigues M, Ezzedine K, Hamzavi I, Pandya A, Harris J. 2017. New discoveries in the pathogenesis and classification of vitiligo. *Journal of the American Academy of Dermatology* **77(1)**:1–13 DOI 10.1016/j.jaad.2016.10.048.

Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H, Sirota-Madi A, Olender T, Golan Y, Stelzer G, Harel A, Lancet D. 2010. GeneCards Version 3: the human gene integrator. *Database* **2010**:baq020 DOI 10.1093/database/baq020.

Salinas-Santander M, Trevino V, De La Rosa-Moreno E, Verduzco-Garza B, Sinchez-Dominguez C, Canti-Salinas C, Ocampo-Garza J, Lagos-Rodriguez A, Ocampo-Candiani J, Ortiz-Lipez R. 2018. CAPN3, DCT, MLANA and TYRP1 are overexpressed in skin of vitiligo vulgaris Mexican patients. *Experimental and Therapeutic Medicine* **15(3)**:2804–2811.

Schrödinger. 2020. *Schrödinger Suite. Available at* http://www.schrodinger.com.

Shannon P. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13(11)**:2498–2504 DOI 10.1101/gr.1239303.

Srivastava P, Pal R, Misra G. 2018. Comparative modelling and virtual screening to discover potential competitive inhibitors targeting the 30s ribosomal subunit S2 and S9 in Acinetobacter baumannii. In: *2018 International Conference on Bioinformatics and Systems Biology (BSB) 26–28 October 2018, Allahabad, India.*

Storey J, Tibshirani R. 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* **100(16)**:9440–9445 DOI 10.1073/pnas.1530509100.

Sun Q, Rabbani P, Takeo M, Lee S, Lim C, Noel E, Taketo M, Myung P, Millar S, Ito M. 2018. Dissecting Wnt signaling for melanocyte regulation during wound healing. *Journal of Investigative Dermatology* **138(7)**:1591–1600 DOI 10.1016/j.jid.2018.01.030.

Sweere A, Fraaije J. 2017. Accuracy test of the OPLS-AA force field for calculating free energies of mixing and comparison with PAC-MAC. *Journal of Chemical Theory and Computation* **13(5)**:1911–1923 DOI 10.1021/acs.jctc.6b01106.

Takeo M, Lee W, Rabbani P, Sun Q, Hu H, Lim C, Manga P, Ito M. 2016. EdnrB governs regenerative response of melanocyte stem cells by crosstalk with Wnt signaling. *Cell Reports* **15(6)**:1291–1302 DOI 10.1016/j.celrep.2016.04.006.

Tarlé R, Nascimento L, Mira M, Castro C. 2014. Vitiligo: part 1. *Anais Brasileiros de Dermatologia* **89(3)**:461–470 DOI 10.1590/abd1806-4841.20142573.

The Gene Ontology Consortium. 2007. The Gene Ontology project in 2008. *Nucleic Acids Research* **36(1)**:D440–D444 DOI 10.1093/nar/gkm883.

The UniProt Consortium. 2016. UniProt: the universal protein knowledge base. *Nucleic Acids Research* **45(1)**:D158–D169.

Wang Y, Li S, Li C. 2019. Perspectives of new advances in the pathogenesis of vitiligo: from oxidative stress to autoimmunity. *Medical Science Monitor* **25**:1017–1023 DOI 10.12659/MSM.914898.

Wishart D, Feunang Y, Guo A, Lo E, Marcu A, Grant J, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. 2017. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Research* **46(D1)**:D1074–D1082 DOI 10.1093/nar/gkx1037.

Wlodawer A, Minor W, Dauter Z, Jaskolski M. 2013. Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS Journal* **280(22)**:5705–5736 DOI 10.1111/febs.12495.

Wu Z, Irizarry R, Gentleman R, Martinez-Murillo F, Spencer F. 2004. A model-based background adjustment for oligonucleotide expression arrays. *Journal of the American Statistical Association* **99(468)**:909–917 DOI 10.1198/016214504000000683.

Yamada T, Hasegawa S, Inoue Y, Date Y, Yamamoto N, Mizutani H, Nakata S, Matsunaga K, Akamatsu H. 2013. Wnt/β-Catenin and kit signaling sequentially regulate melanocyte stem cell

differentiation in UVB-induced epidermal pigmentation. *Journal of Investigative Dermatology* **133(12)**:2753–2762 DOI 10.1038/jid.2013.235.

Zeng X, Zhang P, He W, Qin C, Chen S, Tao L, Wang Y, Tan Y, Gao D, Wang B, Chen Z, Chen W, Jiang Y, Chen Y. 2017. NPASS: natural product activity and species source database for natural product research, discovery and tool development. *Nucleic Acids Research* **46(D1)**:D1217–D1222 DOI 10.1093/nar/gkx1026.

Zhang Y, Cai Y, Shi M, Jiang S, Cui S, Wu Y, Gao X, Chen H. 2016. The prevalence of vitiligo: a meta-analysis. *PLOS ONE* **11(9)**:e0163806 DOI 10.1371/journal.pone.0163806.