Scientific Research Publishing

Voice of the Publisher, 2024, 10, 461-477
https://www.scirp.org/journal/vp
ISSN Online: 2380-7598
ISSN Print: 2380-7571

Seasonal DISDB: A Novel Computational Database for Chemical Profiling of Seasonal Diseases

Venu Paritala[1*], Harsha Thummala[2], Sukesh Kalva[2], Rajashekhar Reddy Shagamreddy[3], Jagadish Mandava[3]

[1]Department of Management with Data Analytics, Indiana Wesleyan University, Marion, IN, USA
[2]Department of BioTechnology, Vignan's Foundation for Science, Technology and Research, Guntur, AP, India
[3]Department of Health Informatics and Bioinformatics, Grand Valley State University, Grand Rapids, MI, USA
Email: *vvenuparitala@gmail.com

How to cite this paper: Paritala, V., Thummala, H., Kalva, S., Shagamreddy, R. R., & Mandava, J. (2024). Seasonal DISDB: A Novel Computational Database for Chemical Profiling of Seasonal Diseases. *Voice of the Publisher, 10,* 461-477.
https://doi.org/10.4236/vp.2024.104036

Received: September 11, 2024
Accepted: December 28, 2024
Published: December 31, 2024

Abstract

Systems biology, an emerging field, investigates the interactions of genes and drug actions across various levels of complexity, ranging from molecular and cellular to tissue and organism levels. Seasonal diseases, often overlooked by the global population, pose a significant threat to public health, potentially as dangerous as, if not more than, many chronic diseases. Every year, around a billion cases of seasonal diseases occur worldwide, with 3 to 5 million resulting in severe illness, and an estimated 290,000 to 650,000 deaths annually. Recognizing the critical need to bridge the expanding realm of biological knowledge with vaccine biology to combat these diseases, we introduce an international Seasonal Database (SD). This database integrates genomic, proteomic, and therapeutic resources, encompassing disease-associated genes, proteins, epitopes, and molecular phylogeny. Our compilation includes over 3000 genes annotated for 2000 proteins, providing a comprehensive resource for in-depth exploration of seasonal diseases. SD facilitates in silico evaluations to predict epitopes capable of triggering T-cell responses, and offers a wealth of disease-related information, including symptoms, precautions, treatments, and approved drugs. Additionally, the database aids in selecting new compounds based on their activity profiles against a wide range of biological targets, including those linked to adverse drug events. The user-friendly interface of SD incorporates embedded bioinformatics tools, enabling precise in silico experiments, and can be accessed at https://seasonaldatabase.shinyapps.io/Seasonaldis/).

Keywords

Seasonal, Phylogenetic, Genomic, Proteomic, Target, GenBank,

Pubmed, Disease, Bioinformatics

1. Introduction

The global population frequently ignores seasonal diseases, even though they pose a serious threat to public health and may even be more dangerous than many chronic diseases. Every year, around a billion cases of seasonal diseases occur worldwide, with 3 to 5 million of these resulting in severe illness. Alarmingly, seasonal diseases are responsible for an estimated 290,000 to 650,000 deaths annually. To address this growing concern, the Seasonal Database (SD) was developed as a comprehensive resource to assist researchers in studying and mitigating the impact of these diseases. The database integrates crucial information on proteins, drugs, genes, and epitopes, allowing researchers to explore molecular interactions, identify potential drug targets, and predict immune responses.

The traditional drug design paradigm, where drugs selectively interact with one or two proteins, is being challenged by recent studies revealing the prevalence of polypharmacology, where drugs interact with multiple targets. For instance, celecoxib, initially considered a selective cyclooxygenase-2 non-steroidal anti-inflammatory drug (NSAID), has been found to act on additional targets like carbonic anhydrase II and 5-lipoxygenase. Similarly, rosiglitazone, used in type II diabetes treatment, not only stimulates peroxisome proliferator-activated receptor g but also blocks interferon gamma-induced chemokine expression in Graves disease. Polypharmacology, while offering potential benefits, often leads to side effects, as seen with withdrawn drugs like cisapride and astemizole due to the risk of fatal cardiac arrhythmia linked to their blockade of the hERG potassium ion channel, an unintended "anti-target." Understanding the dynamic nature of targets and anti-targets, as exemplified by H1R antagonists, has prompted new strategies to predict and characterize drug-target associations, aiming to enhance the efficacy, reduce toxicity, and minimize adverse effects in current drug discovery paradigms (Taboureau et al., 2010).

The assembly of large-scale chemical bioactivity databases is expanding our understanding of polypharmacology (high affinity bioactivity across related targets) and promiscuity (low affinity across multiple families) of chemicals, thereby broadening the chemical space for potential druggable targets. While many studies focus on specific protein families like G-protein coupled receptors, nuclear receptors, and kinases, there is also a consideration of the global pharmacology profiles of chemicals (Stajich et al., 2011). Recent advancements in chemoinformatics, exemplified by tools such as iPHACE and Similarity Ensemble Approach (SEA), facilitate polypharmacology data mining, enabling pharmacological space navigation for small molecule drugs. Additionally, biological information for a vast array of chemical compounds can be accessed through databases like PubChem, CheBI, and ChEMBL.

Simultaneously, the National Centre for Biotechnology Information (NCBI), accessible at https://www.ncbi.nlm.nih.gov/, serves as a comprehensive repository of bioinformatics resources, featuring an interconnected database network (Aurrecoechea et al., 2016). In parallel, the Seasonal Database, available at seasonaldb.org, stands as a valuable online resource dedicated to data mining and phylogenetic research. This freely accessible platform provides an intuitive and interactive interface for delving into various aspects, including gene exploration, phylogeny, computational data encompassing targets, proteomics, and phenomic information for specified species. Additionally, the Database offers insights into blast pathways, epitope prediction, clinical variants, and findings derived from numerous bioinformatics analyses. The integration of these databases contributes significantly to the advancement of bioinformatics research and aids researchers in accessing a wealth of interconnected information for a holistic understanding of biological data. The IMMUNE DB (https://venuparitala.shinyapps.io/immune_db/) is a resource dedicated to providing the most up-to-date and accurate information about immune illnesses. The databases listed are developed independently of other databases and websites. To determine the gene function of a specific insect's genomic data, the gene above function identification database may be utilized separately to annotate the positions of all genes in the species (Wildeman, 2016). The techniques necessary for data mining are highly complicated, complex, and time-consuming. They cannot adequately fulfill the objectives of data mining researchers, and the cost is too expensive.

In response to these challenges, researchers from bio-information technology companies have developed specialized biological information cloud service platforms to cater to the evolving needs of researchers. Agri IT Solutions Co., Ltd has introduced Neglated DISDB, accessible at https://venuparitala.shinyapps.io/Neglated-diseasedb/. This platform focuses on neglected diseases and offers gene data mining capabilities, encompassing genes, proteins, and targets. It stands out for its exclusivity in covering neglected diseases precisely without addressing other conditions concurrently. However, it's essential to note that while Neglated DISDB provides phylogenetic insights, it comes at a cost and is not available for free use (Paritala et al., 2021). Users are required to pay a membership fee to access its features and benefits.

In light of numerous limitations in existing databases, our team has developed SD (Seasonal databases) at http://www.seasonaldb.org.cn/. This comprehensive service database uniquely integrates blast and phylogeny analyses. Unlike others, it facilitates simultaneous exploration of the evolutionary relationships across various species, offering a user-friendly interface. SD goes beyond conventional use by allowing users to define genes, targets, and disease causes based on age factor statistical analysis. It stands as a pioneering database, providing a thorough and cohesive platform for comprehensive analysis.

Data Sources

The Seasonal Database (SD) categorizes 25 diseases based on seasonal organiza-

tion. Disease information is sourced from reputable entities such as the Centers for Disease Control and Prevention (CDC) (https://www.cdc.gov/flu/about/season/index.html) (Pratt et al., 2011) and the European Centre for Disease Prevention and Control (ECDC) (Zarb et al., 2012), with additional insights gathered from Pubmed (https://pubmed.ncbi.nlm.nih.gov/) (Clark et al., 2015). Targets are retrieved from ChEMBL (https://www.ebi.ac.uk/chembl/) (Freedman et al., 2006), proteins are sourced from UniProt (https://www.uniprot.org/) (Gaulton et al., 2011; Bateman et al., 2022), and epitopes are obtained from IEDB (Vita et al., 2018) (https://www.iedb.org/). Phylogenetic analysis is conducted using Mega (Molecular Evolutionary Genetics Analysis) (Tamura et al., 2013), while sequence similarity identification is performed through BLAST (Basic Local Alignment Search Tool) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al., 1990). General information is collected from various literature sources, creating a comprehensive and multifaceted resource in the Seasonal Database.

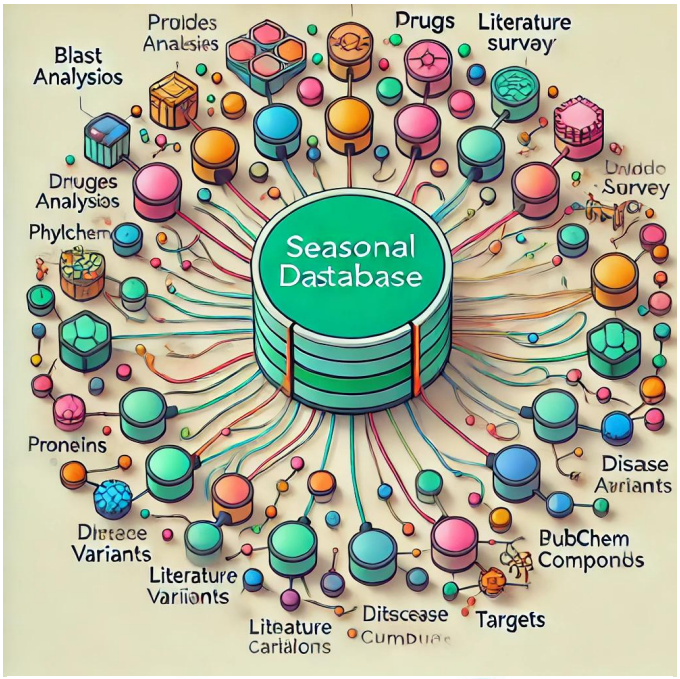
2. Implementation

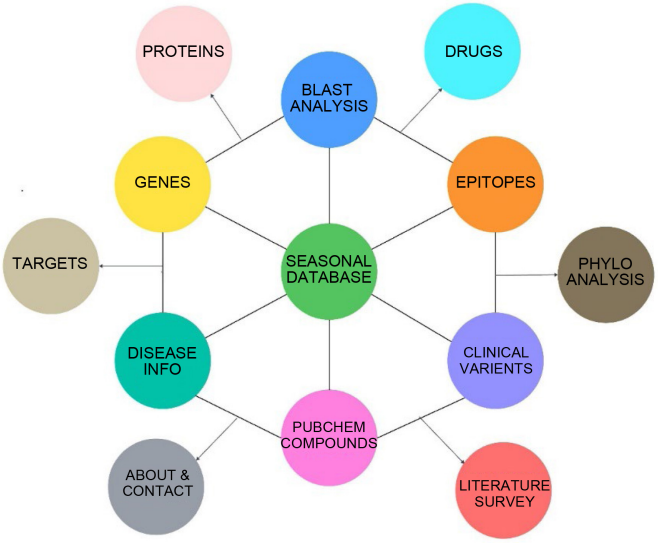
The Seasonal Database (SD) was created through the implementation of R for Reproducible Research, utilizing R Studio as outlined in the Second Edition from 2021. Hosted on a cloud platform, the development of the analytical pipeline employed a variety of R packages (Li, 2021), including shiny, shiny dashboard, shiny themes, dashboard themes, slick, DT, ggplot2, multicolor, gg network, shiny css loaders, shiny custom loader, HTML tools, and HTML widgets. This comprehensive selection of R packages is instrumental in constructing both the user interface and server components of the SD. The interactive user interface is further enriched using HTML and CSS, delivering a dynamic and visually appealing experience for users engaging with the Database. The incorporation of these technologies and packages ensures the creation of a robust and user-friendly platform, facilitating seamless access and analysis of seasonal disease data.

3. Results

3.1. Layout and Content of SD

The SD (Seasonal Database) homepage is divided into two parts: the header (**Figure 1**) and the central green core with a large panel. A drop-down menu labeled "Search for Data Types" is conveniently located on the left side of the header, providing quick access to various features such as Disease Information, Genes, Blast Analysis, Epitopes, Clinical Variants, PubChem Compounds, Proteins, Targets, Drugs, Phylogenetic Analysis, and Literature Survey Search (see **Figure 1**). Directly below the search boxes, the close menu offers easy navigation to information related to the SD project, login and registration options, social media links, a YouTube tutorial channel, and a Contact Us email form. For user support, inquiries can be directed to seasonaldatabase@gmail.com.





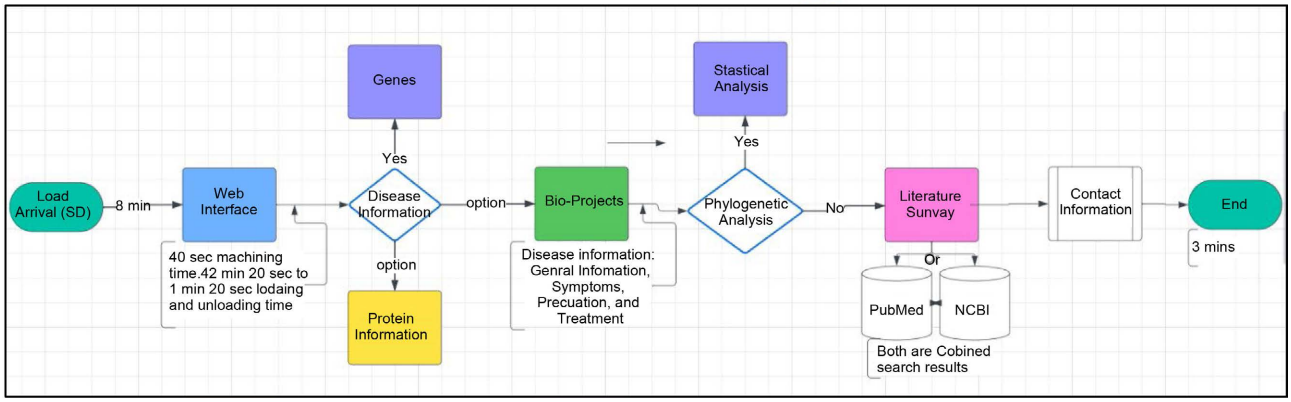


Figure 1. The structure of the seasonal disease database (SEASONAL DISDB), with its key sections such as Disease Info, Genes, BLAST analysis, epitopes, clinical variants, PubChem compounds, proteins, targets, drugs, phylogenetic analysis, and literature survey.

3.2. Exploring SD (Seasonal Database) Records: Disease Name Searches and Disease Record Pages

The Seasonal Database (SD) provides an organised interface for exploring disease records based on seasonal variations. Users can search for diseases, which are categorised under different seasons such as Monsoon, Summer, and Winter. The image showcases the Winter Season section, listing diseases with detailed information, including disease name, general information, symptoms, precautions, treatment, and organism type. For instance, "Common Cold" is featured with descriptions of respiratory illness symptoms like a stuffy and runny nose, sneezing, and mild hacking cough. Precautionary measures, such as washing hands frequently and avoiding close contact with infected individuals, are suggested, alongside treatments that emphasise symptom relief (Figure 2). Diseases in the database are classified according to the season in which they are most prevalent, providing a comprehensive resource for users seeking disease-specific guidance and preventive strategies.

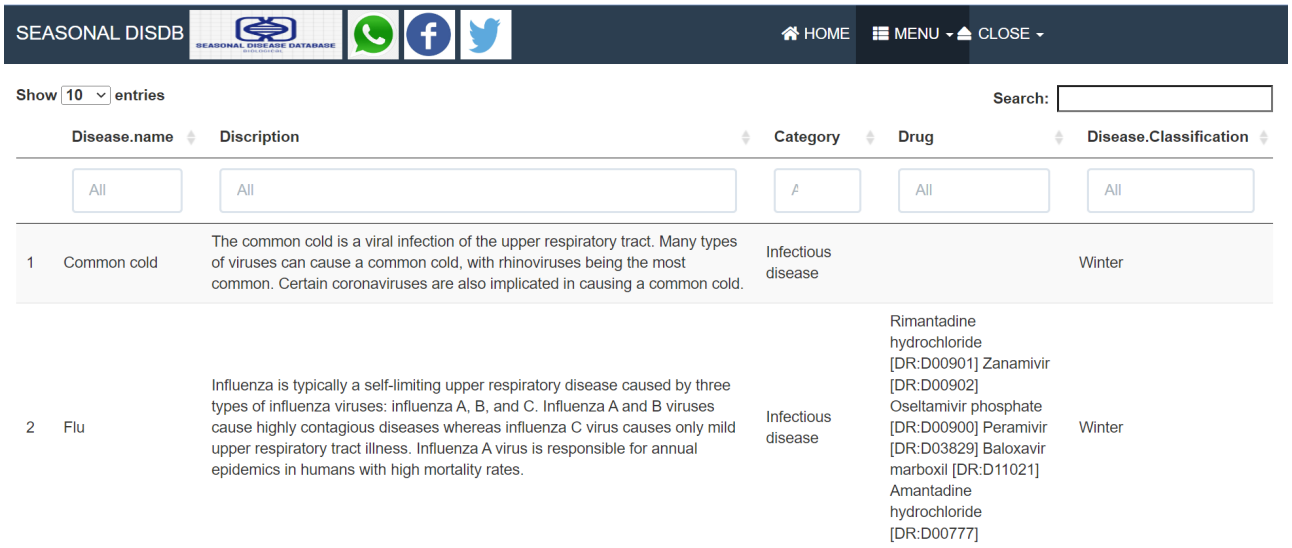


Figure 2. Navigating seasonal disease records and search features in the seasonal database.

3.3. Genes

Genes are essential in understanding biological processes and are critical for identifying disease-specific traits. The Seasonal Database (SD) integrates information for various diseases, with each disease represented by a corresponding gene tab panel. These panels provide crucial details, including Gene ID, gene name, and frequency information specific to each seasonal disease. This information is highlighted and can be utilised effectively by researchers to advance their studies.

Additionally, top genes with the highest frequency are showcased to aid in research prioritisation (Figure 3). Users can easily navigate the database using features such as the Gene ID search, which falls under the Annotation, Curation, and Identifiers category. For added convenience, gene records can be saved, book-

marked, or stored in a personal basket, facilitating seamless access and download options for further research.

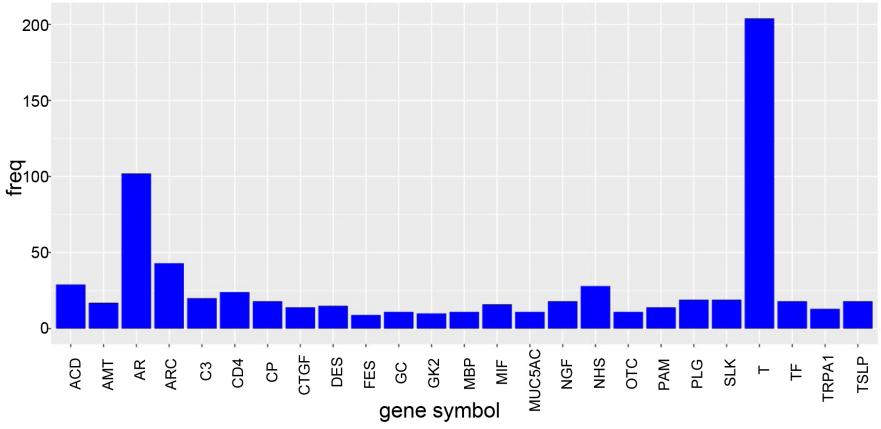


Figure 3. Highlighting top genes with highest frequency and navigation features in the seasonal database.

3.4. Blast (Basic Local Alignment Search Tool) Analysis

Blast analysis plays a pivotal role in bioinformatics by providing detailed insights into the closest species related to a particular organism responsible for causing a disease. To comprehend the intricacies of various conditions, we conduct blast analyses for each disease organism, following a seasonal approach. The analytical outcomes yield diverse data records, encompassing crucial information such as Description, Scientific Name, Common Name, Taxid, Max Score, Total Score, Query Cover, E value, and Per. Ident, Acc. Len, and Accession (Table 1). To enhance user accessibility, we have implemented a user-friendly search bar positioned atop each data record, facilitating easy filtering for users. This dataset proves invaluable for subsequent studies, particularly in the construction of phylogeny analysis, shedding light on the interconnectedness of different organisms.

3.5. Epitopes

The identification of antigens and antibodies, along with the presentation of epitopes, plays a crucial role in immunoinformatics studies. In this context, Seasonal Databases (SD) emerge as significant contributors, standing out as one of the largest biological databases. Within SD, a dedicated epitopes tab panel is featured, housing a wealth of data records vital for comprehensive analysis. These records include Epitope ID, Object Type, Description, Epitope Modified Residue(s), Epitope Modification(s), Starting Position, Ending Position, Non-peptidic epitope Accession, Epitope Synonyms, Antigen Name, Antigen Accession, Parent Protein, Parent Protein Accession, Organism Name, Parent Organism, Parent Organism ID, Epitope Comments, Epitope Relationship, Object Type, Description, Starting Position, Ending Position, Non-peptidic, object Accession, Synonyms, Antigen Name, Parent Protein, Organism Name, and Parent Organism. These data are from the Immune Epitope Database (IEDB); these records are

Table 1. Blast analysis of asthma with Description, Scientific Name, Common Name, Taxid, Max Score, Total Score, Query Cover, E-value, Percent Identity, Accession Length, and Accession.

Description	Scientific Name	Common Name	Taxid	Max Score	Total Score	Query Cover	E value	Per. ident	Acc. Len	Accession
Homo sapiens C-X-C motif chemokine receptor 2 (CXCR2), transcript variant 1, mRNA	Homo sapiens	human	9606	5269	5269	99%	0	100	2853	NM_001557.4
Homo sapiens interleukin 8 receptor beta (IL8RB) mRNA, complete cds	Homo sapiens	human	9606	5249	5249	99%	0	99.83	2856	L19593.1
Homo sapiens interleukin 8 receptor, beta, mRNA (cDNA clone MGC: 46215 IMAGE: 5752441), complete cds	Homo sapiens	human	9606	5212	5212	98%	0	100	2930	BC037961.1
PREDICTED: Pan paniscus C-X-C motif chemokine receptor 2 (CXCR2), transcript variant X1, mRNA	Pan paniscus	pygmy chimpanzee	9597	5156	5156	99%	0	99.03	3159	XM_003818594.4
Pan troglodytes C-X-C motif chemokine receptor 2 (CXCR2), mRNA	Pan troglodytes	chimpanzee	9598	5120	5120	99%	0	98.95	2864	NM_001102661.3
PREDICTED: Gorilla gorilla gorilla C-X-C chemokine receptor type 2 (LOC101151143), transcript variant X1, mRNA	Gorilla gorilla gorilla	western lowland gorilla	9595	5027	5027	99%	0	98.39	2859	XM_004033185.3

Continued

PREDICTED: Homo sapiens C-X-C motif chemokine receptor 2 (CXCR2), transcript variant X2, mRNA	Homo sapiens	human	9606	4791	4791	90%	0	100	2622	XM_017003990.1
PREDICTED: Homo sapiens C-X-C motif chemokine receptor 2 (CXCR2), transcript variant X1, mRNA	Homo sapiens	human	9606	4789	4882	91%	0	99.96	2642	XM_005246530.3
Homo sapiens C-X-C motif chemokine receptor 2 (CXCR2), transcript variant 2, mRNA	Homo sapiens	human	9606	4787	4787	90%	0	100	2656	NM_001168298.2

indispensable for the design of a multi-epitope vaccine targeting specific antigens, contributing significantly to advancements in immunotherapy.

3.6. Clinical Variants

Our team has successfully developed Seasonal Databases, a comprehensive platform that integrates clinical variants, providing a wealth of information crucial for understanding genetic implications. This innovative Database encompasses a range of critical data fields, including Disease Name, Name, Gene(s), Protein change, Condition(s), Clinical significance (Last reviewed), Review status, Accession, GRCh37Chromosome, GRCh37Location, GRCh38Chromosome, GRCh38Location, VariationID, AlleleID(s), dbSNP ID, and Canonical SPDI. This integration of clinical variants within Seasonal Databases serves as a valuable resource for researchers, clinicians, and geneticists, facilitating a deeper understanding of disease mechanisms and genetic factors associated with various conditions.

3.7. PubChem Compounds

In our ongoing efforts to advance computational chemistry and facilitate silicon vaccine design, we have developed Seasonal Databases that now include comprehensive information on PubChem Compounds. This dataset comprises essential data fields such as Compound ID, Compound Name, Compound Synonym,

Molecular Weight, Molecular Formula, Polar Area, Complexity, XlogP, Heavy Cnt, Hydrogen Bond Donor, Hydrogen Bond Acceptor, Rotbonds, InChI, Isosmiles, InChIKey, IUPAC Name, Mesh Headings, Annothits, Annothitcnt, AIDS, CIDCDate, SIDSRCName, DepCatg, and Annotation. The integration of PubChem Compounds into Seasonal Databases provides researchers with a robust resource for exploring chemical properties, aiding in the development of insulin vaccine designs and furthering our understanding of computational chemistry applications in the field of immunology.

3.8. Proteins

Expanding the capabilities of Seasonal Databases (SD), we have incorporated a dedicated Protein Tab Panel enriched with diverse data records essential for comprehensive biological investigations. The dataset within this tab includes critical information such as Entry, Entry Name, Status, Protein Names, Gene Names, Organism, and Length. Researchers and bioinformaticians can refer to this robust compilation to gain insights into the characteristics and attributes of proteins associated with various biological entities. The seamless integration of this Protein Tab Panel within SD enhances the Database's utility, providing a valuable resource for those exploring protein-related aspects in the realm of seasonal studies and beyond.

3.9. Targets

The databases have been enriched with a dedicated section on Targets, encompassing pivotal data records crucial for targeted studies. Within this domain, researchers can access essential information, including ChEMBL ID, Name, UniProt Accessions, Type, Organism, Compounds, Activities, Tax ID, and Species Group Flag. These pertinent details have been meticulously curated from ChEMBL databases, as outlined in the Materials and Methods section. This inclusion enhances the comprehensiveness of the databases, providing a valuable repository for those engaged in research and exploration of target-specific aspects. The data's origin from ChEMBL databases further ensures the reliability and relevance of the information, contributing to the robustness of the databases and supporting targeted investigations in various scientific domains.

3.10. Drugs

In tandem with our commitment to comprehensive data integration, the databases now feature an inclusive section on Drugs, incorporating diverse data records essential for a holistic understanding of therapeutic interventions. These encompass critical information such as Disease Name, Disease Classification, Drugs, and Category. This curated dataset, designed to be user-friendly, serves as a valuable resource for a broad audience, including researchers, clinicians, and pharmaceutical professionals. The inclusion of drugs within the same databases enhances accessibility and facilitates the exploration of the intricate relationship between

diseases and their corresponding therapeutic agents. This amalgamation of pertinent information provides a robust foundation for informed decision-making and advances the collective understanding of drug-disease interactions across various domains.

3.11. Phylogenetics Analysis

Phylogenetic analysis, serving as a fundamental pillar in biological research, plays a pivotal role in unraveling the intricate tapestry of evolutionary relationships among organisms, thereby influencing a myriad of scientific disciplines. This indispensable tool examines genetic and morphological variations within and between species, enabling the reconstruction of ancestral relationships and offering insights into divergence patterns that trace the evolutionary history of life on Earth. In the context of our Seasonal Databases, phylogenetic analysis is seamlessly integrated as a cornerstone, enriching research endeavors across diverse scientific domains. These databases meticulously curate crucial data records, including Disease Name, Disease Classification, Organism Retriever, and Sequence Type, providing a comprehensive foundation for understanding disease dynamics. Utilizing the Neighbor-Joining method, the databases present phylogenetic trees that elucidate evolutionary relationships among pathogens, offering enhanced insights into their origins and transmission patterns (**Figure 4**). Complemented by statistical analyses, manifested as informative bar graphs plotting diseases against organisms, this holistic approach not only facilitates taxonomy and classification but also empowers researchers to discern evolutionary patterns. The Seasonal Databases thus emerge as invaluable tools, housing critical data and providing analytical frameworks to advance phylogenetic research, particularly in the context of seasonal diseases.

3.12. Literature Survey

At the culmination of our research article, we present a distinctive feature termed "Literature Survey," designed to enhance the accessibility and efficiency of navigating through relevant scholarly works. This invaluable resource compiles a wealth of data meticulously extracted from Pubmed databases, employing the robust capabilities of R programming. The Literature Survey encapsulates essential information, including PubMed ID (PMID), Title, Authors, Citation, First Author, Journal/Book, Publication Year, creation date, PMCID, NIHMS ID, and DOI (**Figure 5**). To optimize user experience, each data category is accompanied by an intuitive search option, empowering users to seamlessly filter and extract specific information tailored to their research needs. This comprehensive and user-friendly feature serves as a cornerstone, facilitating a more streamlined and insightful exploration of the extensive body of literature underpinning our research.

In addition to the wealth of information provided, our databases boast an additional navigation menu that enhances the user experience—the "Theme and

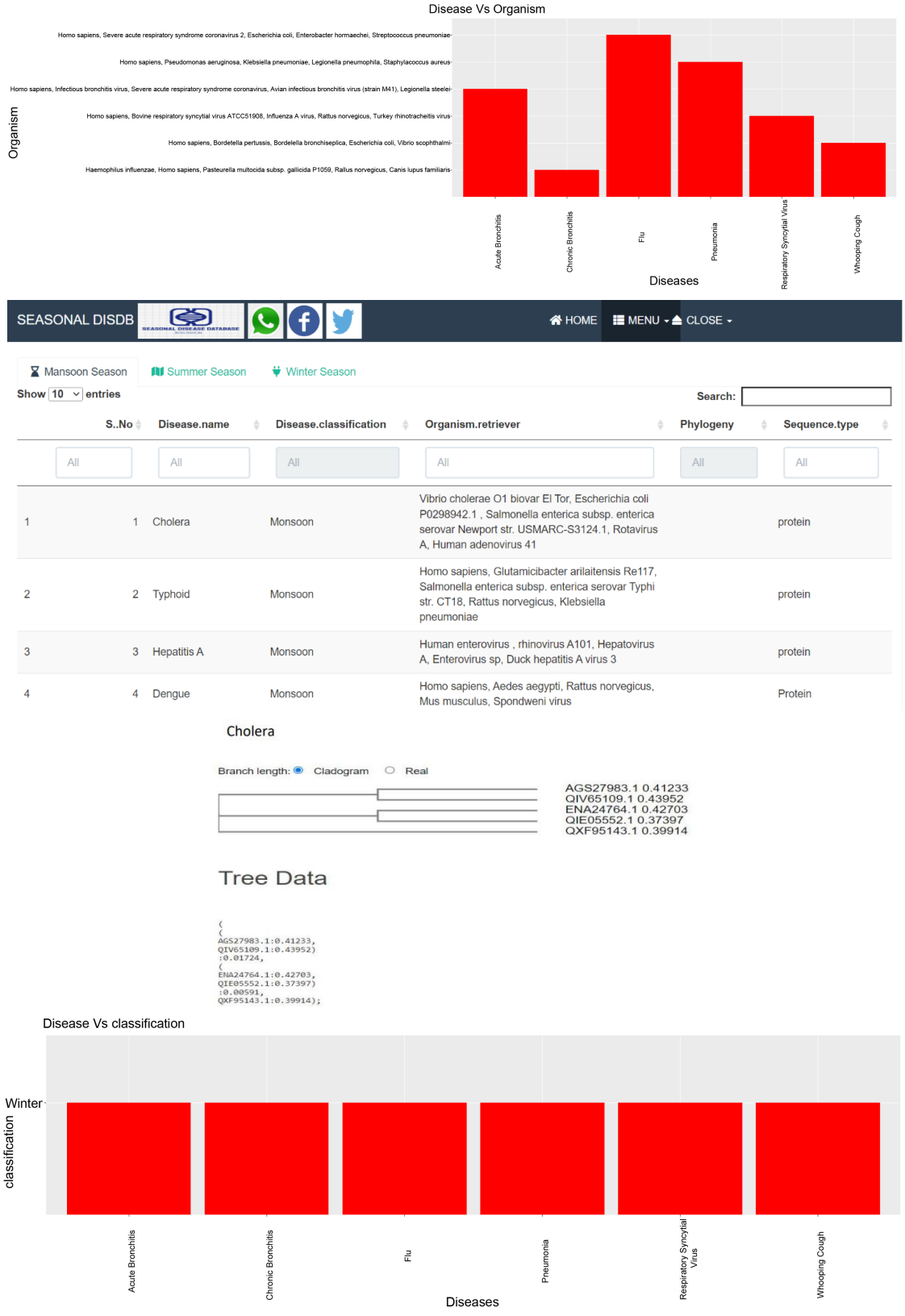
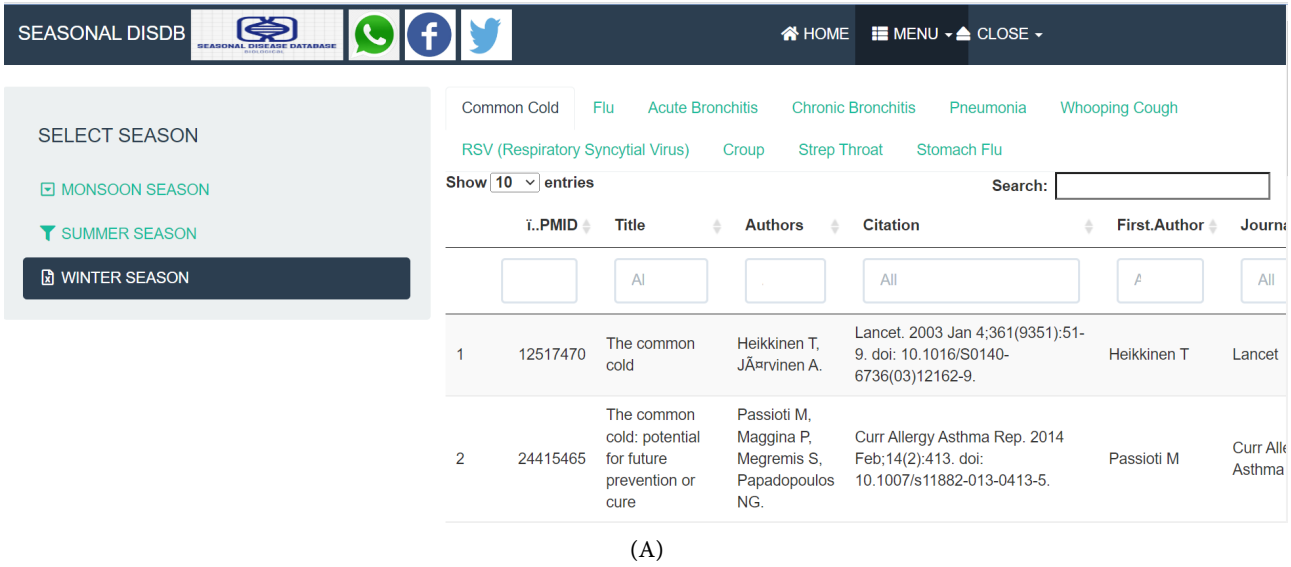
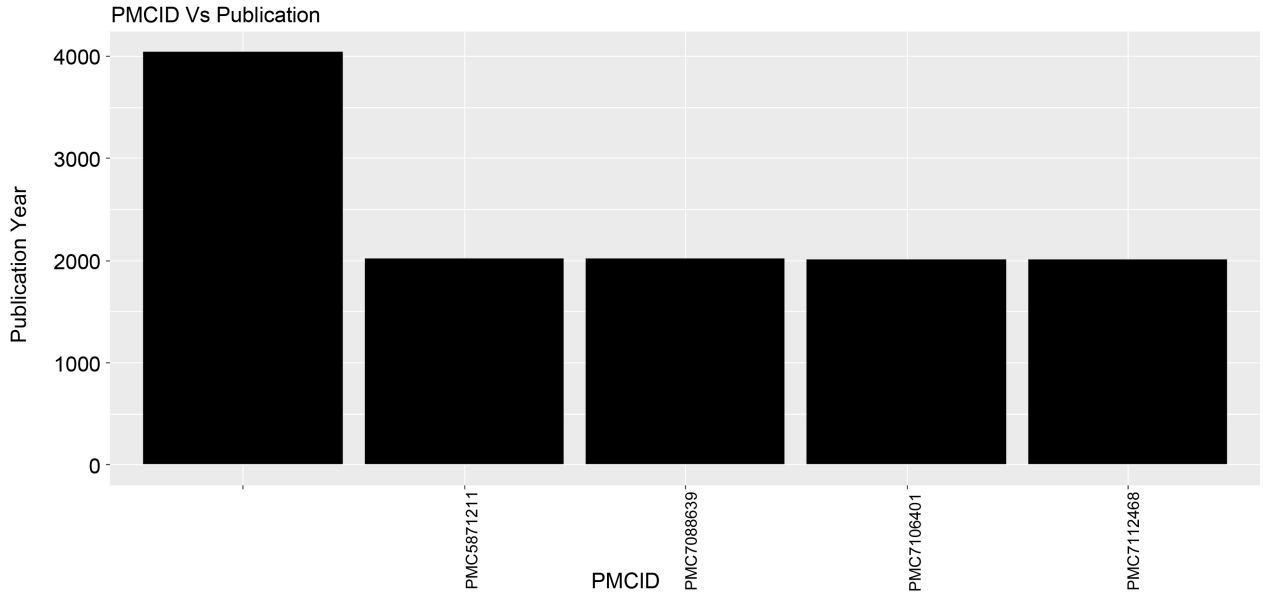


Figure 4. Utilizing molecular sequence data to create a tree illustrating evolutionary relationships among organisms and examining statistical associations between diseases and organisms, exploring potential correlations in disease occurrence across different species.



(A)



(B)

Figure 5. Simple searches and literature record pages continued (A). The literature menu (highlighted in red) contains PMC ID, author name, publication year, etc., and records that show graphs for (B).

Carriers" feature. The Theme feature is ingeniously crafted for database beautification, offering users the ability to personalize their experience by choosing their favorite colors to adorn the interface. This aesthetic customization not only adds a touch of personalization but also contributes to a visually pleasing and engaging research environment. Beyond the visual appeal, the Carrier's information within this menu provides users with insights into the underlying structure and organization of the Database, ensuring transparency and a deeper understanding of the wealth of knowledge at their fingertips. With these thoughtful features, we aim to elevate not only the functionality but also the aesthetic and user-centric aspects of navigating our comprehensive databases.

4. Discussion

In direct comparison to other databases, the Seasonal Database (SD) stands out as a superior platform, excelling in both efficiency and speed. One notable advantage lies in its utilization of a cloud platform during development, contributing to its swift and responsive performance. SD distinguishes itself by providing a wealth of information, encompassing a comprehensive range of research data. From genes and phylogeny to clinical analysis, targets, and blast results, SD covers a vast array of critical information, making it a one-stop resource for researchers and practitioners alike.

One of the critical strengths of SD is its inclusivity, as it successfully incorporates nearly all pertinent research information within its database. This inclusivity extends to the diverse types of updates and curatorial efforts, particularly as new studies emerge, and older data may become obsolete.

Inclusion of New and Emerging Diseases: Although SD currently covers 25 diseases, expanding its scope to include new and emerging seasonal diseases will be crucial for maintaining its utility. Diseases like COVID-19 or avian influenza, which exhibit seasonal variation, highlight the need for databases to adapt quickly to changing global health landscapes. Ensuring that SD can rapidly integrate such diseases will require the development of more dynamic data pipelines and collaborations with public health agencies.

Biases in Data Representation: The database relies on information from established global health institutions like CDC and ECDC, which may introduce biases towards diseases that are more prevalent in high-income countries or those that receive significant research funding. This could limit the representation of diseases prevalent in low- and middle-income regions, where research and data availability might be scarcer. Addressing this bias will be important for ensuring SD's global applicability.

The Advantages of Applications of SD (Seasonal Database)

Seasonal Database (SD) offers a multitude of advantages that position it as a precious resource in the realm of bioinformatics and disease research. Notably, SD excels in comprehensive data integration by seamlessly combining genomic, proteomic, and therapeutic datasets for a wide array of organisms, spanning both pathogenic and nonpathogenic species. The inclusion of approximately 25 diseases from reputable institutions like the CDC and ECDC ensures a diverse and credible dataset, enriching the scope of research possibilities. Drawing information from reliable sources such as GenBank, Epitopes, and PubChem compounds underscores the Database's commitment to accuracy and reliability.

A significant advantage lies in SD's user-friendly interface, complemented by embedded bioinformatics tools, facilitating seamless navigation and experimentation for researchers with varying levels of expertise. The platform's support for specific in-silico experiments empowers researchers to conduct virtual analyses, contributing to hypothesis testing and drug discovery. Mega's incorporation for

illustrating evolutionary relationships provides valuable insights into the genetic dynamics of seasonal diseases, enhancing our understanding of pathogen evolution.

The utilization of a cloud platform during SD's development is a notable factor contributing to its efficiency and speed. This advantage ensures rapid access to data and analyses, enhancing overall performance. Moreover, SD's compatibility across operating systems, including Windows and macOS, promotes widespread adoption, making it accessible to researchers using different platforms.

Beyond these technical advantages, SD holds immense potential for practical applications, particularly in drug discovery and vaccine development. The wealth of information available, coupled with bioinformatics tools, allows researchers to identify potential drug targets, study genetic variations, and explore therapeutic options for various seasonal diseases. In conclusion, the combination of these features positions the Seasonal Database as a robust and versatile tool, empowering researchers to explore and advance knowledge in the dynamic field of seasonal diseases.

5. Conclusion

The Seasonal Database (SD) is a vital resource designed to support researchers focusing on seasonal diseases by providing curated, quality-controlled data and advanced bioinformatics tools. Moving forward, SD aims to expand its data mining capabilities, integrate new data types, and enhance its analytical functionalities to meet the evolving needs of the scientific community. Its potential impact on public health and vaccine development is significant, particularly in areas such as pandemic preparedness, drug discovery, and the rapid identification of vaccine candidates for emerging diseases. By enabling researchers to track pathogen evolution, predict immune responses, and conduct in silico vaccine design, SD plays a crucial role in addressing global health challenges like influenza, COVID-19, and other seasonal diseases. Additionally, SD's commitment to collaboration and feedback from the research community ensures it remains relevant and effective in driving scientific discovery and advancing global disease prevention strategies. Through continuous innovation, SD is positioned to make a substantial contribution to the future of public health and bioinformatics.

Data Availability

All data are incorporated into the article and its online supplementary material.

Author Contributions

Venu Paritala planned the experiments, collected and examined all the data, and authored and edited the manuscript. Every author has reviewed and approved the work.

Ethics Clearance and Participation Consent

It is not applicable because there is no research using human or animal subjects

on this topic.

Conflict of Interest Statement

The authors declare no competing interests.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic Local Alignment Search Tool. *Journal of Molecular Biology, 215,* 403-410. https://doi.org/10.1016/s0022-2836(05)80360-2

Aurrecoechea, C., Barreto, A., Basenko, E. Y., Brestelli, J., Brunk, B. P., Cade, S. et al. (2016). Eupathdb: The Eukaryotic Pathogen Genomics Database Resource. *Nucleic Acids Research, 45,* D581-D591. https://doi.org/10.1093/nar/gkw1105

Bateman, A., Martin, M., Orchard, S., Magrane, M., Ahmad, S., Alpi, E. et al. (2022). UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research, 51,* D523-D531. https://doi.org/10.1093/nar/gkac1052

Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2015). GenBank. *Nucleic Acids Research, 44,* D67-D72. https://doi.org/10.1093/nar/gkv1276

Freedman, S. B., Adler, M., Seshadri, R., & Powell, E. C. (2006). Oral Ondansetron for Gastroenteritis in a Pediatric Emergency Department. *New England Journal of Medicine, 354,* 1698-1705. https://doi.org/10.1056/nejmoa055119

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A. et al. (2011). ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Research, 40,* D1100-D1107. https://doi.org/10.1093/nar/gkr777

Li, J. (2021). The Exploration of the Approach to Data Preparation for Chinese Text Analysis Based on R Language. *OALib, 8,* e7821. https://doi.org/10.4236/oalib.1107821

Paritala, V., Reddy, R. S., & Kalva, S. (2021). Neglected DISDB: A Broad Internet Framework for Gathering and Analysing Data from Neglected Diseases. *Journal of Applied Bioinformatics & Computational Biology, 10,* Article No. 5.

Pratt, L. A., Brody, D. J., & Gu, Q. (2011). Antidepressant Use in Persons Aged 12 and over: United States, 2005-2008. *NCHS Data Brief, 76,* 1-8.

Stajich, J. E., Harris, T., Brunk, B. P., Brestelli, J., Fischer, S., Harb, O. S. et al. (2011). FungiDB: An Integrated Functional Genomics Database for Fungi. *Nucleic Acids Research, 40,* D675-D681. https://doi.org/10.1093/nar/gkr918

Taboureau, O., Nielsen, S. K., Audouze, K., Weinhold, N., Edsgard, D., Roque, F. S. et al. (2010). ChemProt: A Disease Chemical Biology Database. *Nucleic Acids Research, 39,* D367-D372. https://doi.org/10.1093/nar/gkq906

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution, 30,* 2725-2729. https://doi.org/10.1093/molbev/mst197

Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R. et al. (2018). The Immune Epitope Database (IEDB): 2018 Update. *Nucleic Acids Research, 47,* D339-D343. https://doi.org/10.1093/nar/gky1006

Wildeman, S. M. (2016). Consent to Psychiatric Treatment: From Insight (into Illness) to Incite (a Riot). *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3613668

Zarb, P., Coignard, B., Griskeviciene, J., Muller, A., Vankerckhoven, V., Weist, K. et al. (2012). The European Centre for Disease Prevention and Control (ECDC) Pilot Point Prevalence Survey of Healthcare-Associated Infections and Antimicrobial Use. *Eurosurveillance, 17,* Article No. 20316. https://doi.org/10.2807/ese.17.46.20316-en

Supplementary Material

Supplementary material is available online at the Database. Feedback and engagement can be directed to help@seasonaldatabase.org.