9
SpiceRx-An Integrated Resource for the Health Impacts of Culinary Spices and Herbs

9.1 INTRODUCTION

Integration of scientific evidence available from an exponentially growing literature reporting health consequences of culinary spices and herbs will enable drawing inferences for their informed culinary use as well as for generating hypotheses to discover underlying molecular mechanisms. The previous Chapters 7 and Chapter 8 conducted a detailed data-driven analysis of spice-disease associations from a vast array of biomedical literature available in MEDLINE. Our data analysis on the text mined literature relating spices and diseases from published peer-reviewed biomedical research articles revealed that spices have a broad spectrum utility meaning that spices have therapeutic properties with a wide range of disease categories [Rakhi, Tuwani, Mukherjee, et al., 2018]. This goes against the popular belief [Billing and Sherman, 1998] that spices are used in cuisines merely for enhancing the palatability of food or due to their food preserving antimicrobial properties. We further enhanced our analysis by finding the molecular basis of spice-disease associations by linking the phytochemicals implicated in the therapeutic benefit of a spice with a disease by incorporating data available from Comparative Toxicogenomics Database (CTD). We designed SpiceRx with the aim of bridging information associating culinary spices/herbs, their phytochemicals and diseases with the help of evidence compiled from research articles and external resources to provide a platform for open-ended explorations of their tripartite relationships (Figure 9.1).

9.2 DATABASE OVERVIEW

SpiceRx provides a platform for exploring the health impact of spices and herbs used in food preparations through a structured database of tripartite relationships with their phytochemicals and disease associations (Figure 9.1). Starting with an extensive dictionary of 188 culinary spices and herbs, their disease associations were text mined from 28 million MEDLINE abstracts. These data were further combined with evidence of spice-phytochemical and phytochemical-disease associations. SpiceRx presents a compilation of 11750 MEDLINE abstracts containing 8957 disease associations (8172 positive and 783 negative) for 152 spices linked to 848 unique disease-specific Medical Subject Headings (MeSH) IDs [Lipscomb, 2000]. The hierarchical organization of MeSH was used as a basis for the ontological classification of disease terms. Spice names were tagged using a dictionary matching method whereas diseases were labeled using a machine learning-based model, TaggerOne [Leaman and Lu, 2016]. A convolutional neural network-based relation extraction model [Kumar Sahu et al., 2016; Nguyen and Grishman, 2015] was trained on 6712 manually annotated sentences to extract and classify positive, negative and neutral associations.

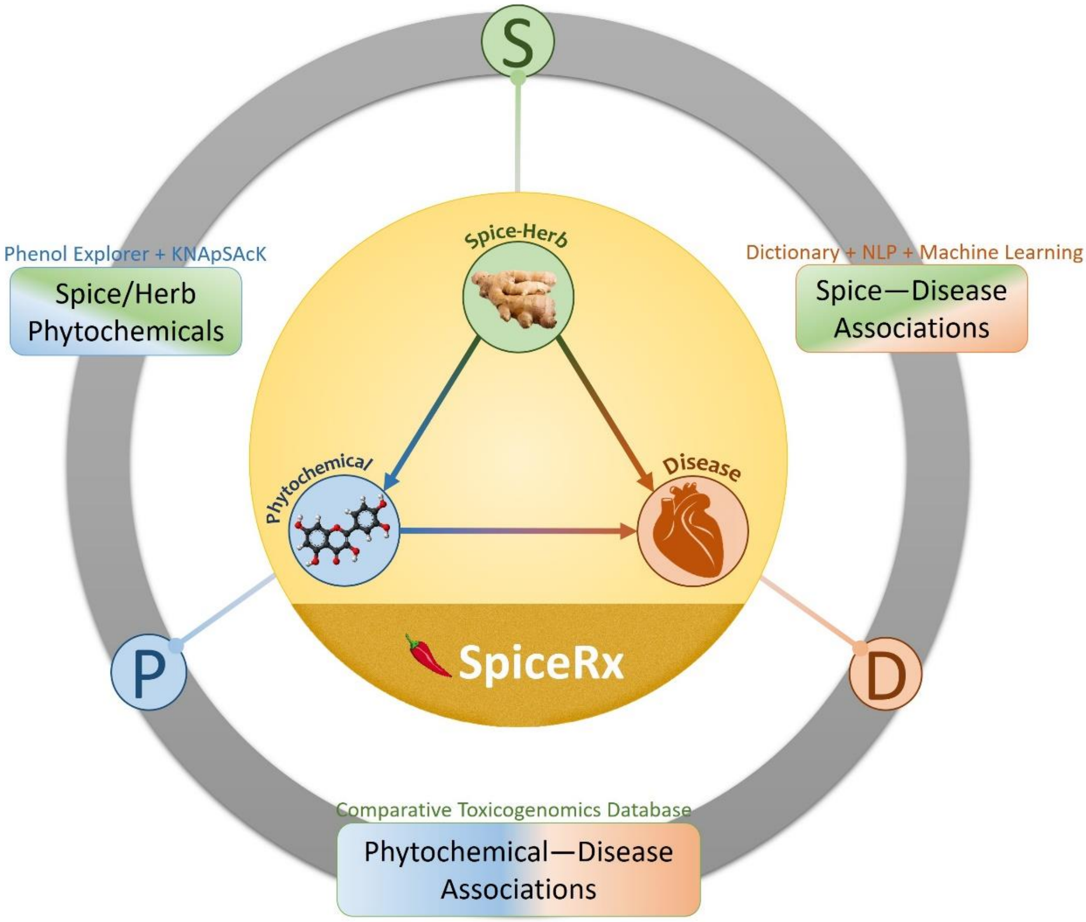

Figure 9.1: Overview of SpiceRx: SpiceRx is an integrated repertoire of evidence-based knowledge pertaining to the health impacts of culinary spices/herbs and their phytochemicals. It seamlessly integrates scientific evidence from biomedical literature and external resources using a text mining protocol.

Information of 866 phytochemicals was obtained for 142 of the spices using KNApSAcK [Afendi et al., 2012] and PhenolExplorer [Rothwell et al., 2013] that comprised of 570 bioactive compounds with a total of 2042 spice-phytochemical associations. These compounds were further linked to diseases with the help of the Comparative Toxicogenomics Database (CTD) [Davis et al., 2017], a public database of literature-curated and inferred chemical-disease associations. The resource unearths literature-supported spice-disease associations for which the spice phytochemical(s) have been independently reported for therapeutic effects. Thus, through interlinked triangular relationships between culinary herbs and spices, their phytochemicals, and diseases, SpiceRx facilitates a seamless exploration of evidence-based knowledge for their disease-specific culinary recommendations as well as enquiry into underlying molecular mechanisms.

The data and the code used for training the model is available at the Complex System Laboratory, IIIT-Delhi's GitHub page: https://github.com/cosylabiiit/spice-disease- associations.

9.3 ARCHITECTURE AND WEB INTERFACE

SpiceRx has been designed to facilitate explorations starting from either a spice/herb, a disease, or a phytochemical query to find its association with the remaining two elements.

9.3.1 Spice Search

Culinary spices/herbs can be searched using their common name, scientific name or NCBI taxonomy ID to obtain their disease associations and constituent phytochemicals (Figure 9.2A). Searching for a spice yields a paginated list of its disease associations (Figure 9.2B). (ranked in descending order of number of publications) and that of its phytochemicals (Figure 9.2D). Research articles reporting the association are listed with link-outs to PubMed (Figure 9.2C). For 20% associations, therapeutic spice phytochemicals involved in a spice-disease association discovered through triangular causal linking are listed. These provide a ground for their pharmaceutical and nutraceutical applications in addition to informed culinary use. Separately, a list of all spice-linked phytochemicals is provided. Phytochemicals could be further explored for their detailed physicochemical features, drug-likeness, and links to associated spices and diseases.

9.3.2 Disease Search

To identify culinary spices and herbs that are reported with either therapeutic or adverse effect against a specific disease, SpiceRx provides disease search integrated with the hierarchical organization of MeSH disease terms. One may search by Disease Name, Disease Category, Disease Sub Category or MeSH ID. The Disease Name could be used to search by common names, such as 'Diabetes', 'Weight Gain', or 'Obesity'. The 'MeSH Disease Category' represents a broad class of diseases such as 'Cardiovascular Diseases', 'Endocrine System Diseases' or 'Neoplasms', whereas 'Disease Sub Category' represents more refined disease terms such as 'Heart Diseases', 'Metabolic Diseases', 'Liver Diseases' etc. A null search performed with no specific disease term is designed to present an exhaustive list of diseases and their spice associations. Details of spices reported with positive or negative effects for the queried disease term along with link-outs to PubMed articles are presented, in addition to any specific phytochemical(s) involved in the therapeutic association, whenever available.

9.3.3 Phytochemical Search

Spice phytochemicals form the basis for molecular mechanisms involved in therapeutic effects against specific diseases. Apart from querying SpiceRx for a spice or a disease, one may also explore the resource to search compounds by their structure, common name, IUPAC name, PubChem ID, molecular weight, Hydrogen bond donors/acceptors or molecular hydrophobicity (AlogP) (Figure 2e). Each phytochemical could further be explored for its chemical profile (MESH ID, PubChem ID, Common name, IUPAC name, Molecular Formula, Canonical and Isomeric SMILES; Phytochemical and ADMET properties), disease associations as well as a list of spices in which it is reported to be found in. Apart from 2D and 3D visualizations and download options (Mol, 2D Image, and SDF), lookup for structurally similar spice compounds within the database as well as those commercially available from external sources is also provided (Figure. 9.2f). The 'null search' (with all fields empty) is designed to yield a list of all spice phytochemicals.

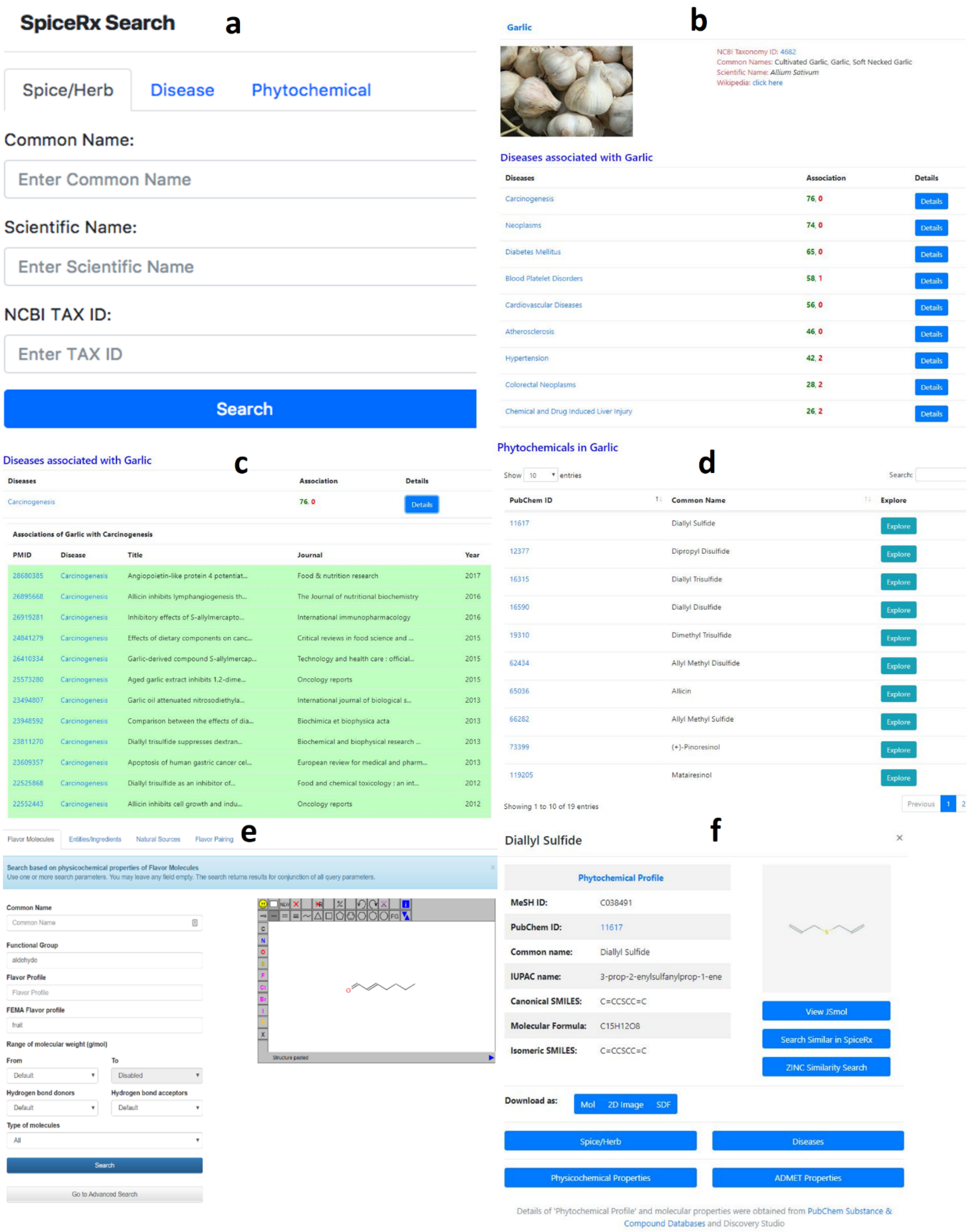

Figure 9.2: The user interface of SpiceRx: The user interface of SpiceRx has been designed to enable discovery of interesting relationships through seamless exploration of triangular associations SpiceRx facilitates exploration of tripartite associations between spices/herbs, their phytochemicals and disease starting with any one of the elements: (a) Spice, Disease and Phytochemical. Search panels are designed to enable experts and non-experts alike and are assisted with multiple features such as autocomplete, spice synonyms, and such. (b) SpiceRx results page for spice search displaying its positive and negative disease associations. (c) For each disease association, its corresponding empirical evidence in the form of citation of research article reporting the association is listed. (d) The list of phytochemicals contained in a spice. The detailed property for the phytochemical by expanding the more info tab. (e) Phytochemical search is facilitated by various search options including searching using molecular drawing (JSME Molecular Editor). (f) The phytochemical

information panel displaying the detailed properties of the queried phytochemical along with download tab, physicochemical properties, ADMET properties and diseases associated with it.

9.4 WEBSERVER TECH STACK

SpiceRx is implemented with the Python web development framework Django (https://www.djangoproject.com). and PostgreSQL (https://www.postgresql.org/). Django has a built-in ORM (Object Relational Mapper) for querying the database, thus performing complex queries apart from reducing the development period. The frontend was built using HTML, CSS, JavaScript, AJAX, jQuery, JSME Molecular Editor [Bienfait and Ertl, 2013], Bootstrap, Jmol, DataTables and Google Charts. An Apache HTTP Server has been used to route requests to the Django application and to enable data compression for faster page load times. The site is best viewed in the latest versions of Google Chrome, Firefox, Opera, Internet Explorer, and Microsoft Edge. Google charts library was used to generate the visualizations for statistics for SpiceRx. Details of the SpiceRx statistics is provided in Annexure D.1

9.5 CASE STUDIES

SpiceRx could be used for generating a variety of queries intended to find disease associations of spices and phytochemicals, to find spices linked to specific disease terms or for probing triangular associations between the three elements. Such queries could be aimed at finding ways for informed culinary use of spices and herbs against specific disease(s), discovering novel therapeutic spice-disease associations unreported hitherto, finding new ways of repurposing spices or finding potential spice phytochemicals against a specific disease, among others. Here we present a few case studies.

9.5.1 Searching for Drug-Like Compounds in SpiceRx

One may search for spice phytochemicals that have drug-like properties using the phytochemical search tab in SpiceRx by using criteria for fulfilling Lipinski’s rule (Figure 9.3). Lipinski's rule specifies conditions to evaluate drug-likeness of a compound, i.e., the suitability of a chemical compound to have pharmacological or biological activity making it a likely candidate for orally active drug. According to Lipinski’s rule, an orally active drug has no more than one violation of the following criteria: No more than 5 hydrogen bond donors; No more than 10 hydrogen bond acceptors; Molecular mass less than 500 Daltons; and an octanol-water partition coefficient (log P) not greater than 5.

By specifying these conditions in the relevant search tabs, all compounds satisfying the above conditions can be obtained. This search in SpiceRx yields 639 (74 %) of all the phytochemicals in SpiceRx (866), suggesting the potential therapeutic value of these compounds.

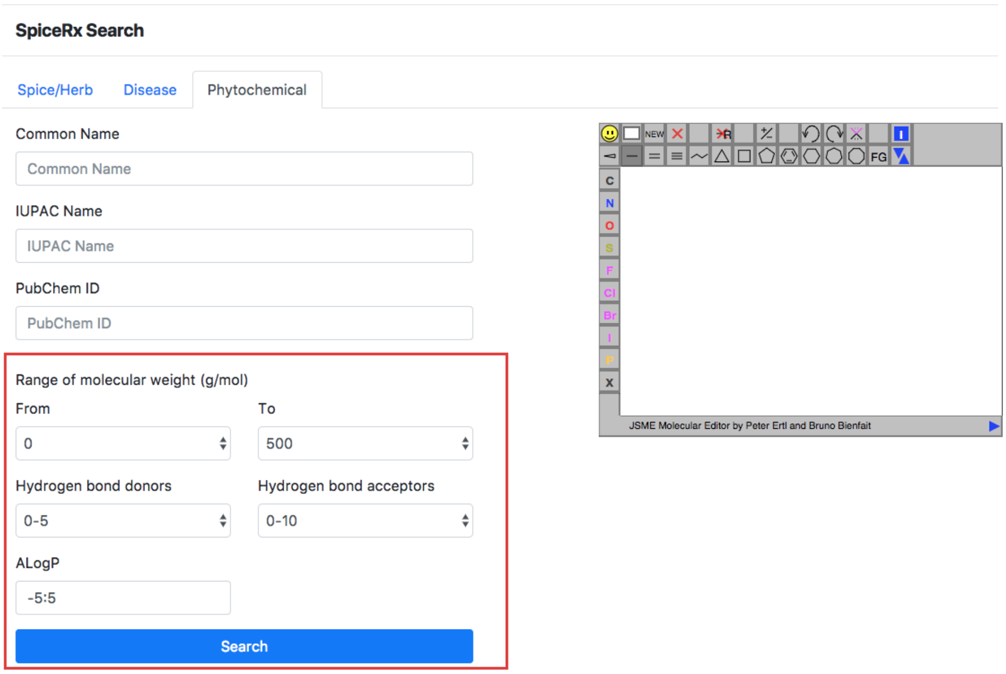

Figure 9.3: Searching for drug-like phytochemicals in SpiceRx fulfilling Lipinski's rule.

9.5.2 Investigating Molecular Mechanisms behind Therapeutic Effects of Spices and Herbs

One of the key attributes of spices and herbs is that they are antimicrobial agents. Billing and Sherman [Billing and Sherman, 1998; Sherman and Billing, 1999a] suggested that one of the reasons for the widespread use of spices in recipes is primarily due to their antimicrobial properties. One can search for spices and their phytochemicals which are active against microbial infections. Searching the broad category "Bacterial Infections and Mycosis" reveals that a wide range of culinary spices and herbs have positive associations against this category of diseases. Fennel and cinnamon have been reported with antimicrobial properties against salmonella infections. Common culinary spices and herbs such as garden thyme, oregano, clove, garlic, cinnamon, rosemary and turmeric have positive associations against "Pneumonia Staphylococcal". Similarly, garlic, turmeric, oregano and clove are among spices which are reported to be beneficial for "*Escherichia coli* Infections". Garlic clove and thyme are among spices which are reported to be beneficial for "Mycosis". Interestingly, the molecular mechanism behind some of these antimicrobial effects of spices are yet to be uncovered. SpiceRx provides a fertile ground for further investigations into such cases where the molecular mechanisms for their actions are not yet evident.

Similarly, a large number of daily culinary spices and herbs have beneficial associations with Diabetes Mellitus, a common Nutritional and Metabolic Diseases. Among them, fenugreek, cinnamon, garlic, turmeric, ginger, and black cumin have been reported most frequently in literature. Beyond the three spice phytochemicals ('gallic acid', 'berberine', 'sophoraflavonoloside') which have been linked to diabetes mellitus, further investigations into the effects of phytochemical content in these spices can reveal their synergistic actions in disease regression and control.

9.5.3 Hypothesis Generation and Knowledge Discovery through SpiceRx

SpiceRx can be used as a platform for knowledge discovery and hypothesis generation. For specific diseases, SpiceRx provides phytochemicals that are therapeutically associated with them. By finding out the spices in which these phytochemicals occur, one can generate and test hypotheses confirming whether these spices have beneficial effects for these diseases. A few examples can further exemplify this point. 'Glycyrrhizin' is a phytochemical therapeutically associated with 'Hepatitis A'. The phytochemical 'Glycyrrhizin' is reported to be found in the herb liquorice. Similarly, '(-)-Epigallocatechin gallate' is another phytochemical associated with the disease 'Herpes Simplex'. This chemical is present in various spices such as peppermint, avocado leaf, bell pepper and german chamomile. These spices and herbs are not directly associated with Herpes Simplex, which gives room for testing the hypothesis. Likewise, 'Apigen' is a phytochemical in many spices and herbs including thyme, coriander, hyssop and fennel among others, found to be effective for Adenoviridae Infections. No spices so far have been reported to be beneficial for Adenoviridae Infections. Further studies can be undertaken to confirm the efficacy of spices/herbs containing apigen against Adenoviridae Infections.

9.6 CONCLUSION

SpiceRx integrates scientific evidence from biomedical literature and external resources to seamlessly collate tripartite associations between culinary spices/herbs, diseases, and phytochemicals. By blending scattered and disorganized evidence, it provides a platform for investigation of spices for their health effects, chemical mechanisms behind their action and paves the way for developing nutraceuticals and drugs.

The effect of spice/herb on a disease may sensitively depend on its quantity. Due to the unavailability of uniform data, SpiceRx does not include this information. Further, disease associations can vary according to various factors such as age, pre-existing conditions and gender among others, which are not represented at present. There is scope to enhance SpiceRx to include such additional features.

…