

<https://doi.org/10.1038/s41538-025-00458-z>

From phytochemicals to recipes: health indications and culinary uses of herbs and spices

Rishemjit Kaur^{1,2}, Shuchen Zhang³, Bhavika Berwal¹, Sonalika Ray^{1,2}, Ritesh Kumar^{1,2} & Lav R. Varshney³ ✉

Herbs and spices each contain about 3000 phytochemicals on average, and there is much traditional knowledge on their health benefits. However, there is a lack of systematic studies to understand the relationship among herbs and spices, their phytochemical constituents, their potential health benefits, and their usage in regional cuisines. Here, we use a network-based approach to elucidate established relationships and predict novel associations between the phytochemicals present in herbs and spices and health indications. Our top 100 inferred indication-phytochemical relationships rediscover 40% known relationships and 20% that have been inferred via gene-chemical interactions with high confidence. The remaining 40% are hypotheses generated in a principled way for further experimental investigations. We also develop an algorithm to find the minimum set of spices needed to cover a target group of health conditions. Drawing on spice usage patterns in several regional Indian cuisines and a copy-mutate model for regional cuisine evolution, we characterize the spectrum of health conditions covered by existing regional cuisines. The spectrum of health conditions can expand through the nationalization/globalization of culinary practice.

The co-evolution of plants with their pests and pathogens has led to chemical defenses in plants in the form of phytochemicals^{1–3}. These phytochemicals, even in trace quantities, have a range of disease-alleviating properties as antioxidants, anti-inflammatories, and even anticarcinogens. Across the globe, traditional knowledge systems have long recognized these properties of phytochemicals and incorporated them into culinary and therapeutic practices to mitigate foodborne pathogens and promote health. In China and India, extensive pharmacopeias and long-standing ethnobotanical traditions have led to the systematization of spice use in both culinary and formal medical systems, such as traditional Chinese medicine (TCM) and Ayurveda^{4–10}. Additionally, Mediterranean civilizations have used oregano (*Origanum vulgare*), thyme (*Thymus vulgaris*), and rosemary (*Salvia rosmarinus*) for millennia, whereas Mesoamerican cultures developed sophisticated applications of chili peppers (*Capsicum annuum*) and epazote (*Dysphania ambrosioides*) to alleviate intestinal parasites and digestive discomfort^{11,12}. This widespread pattern of spice use for both culinary enhancement and medicinal benefits reflects a cross-cultural understanding of bioactive properties of phytochemicals that has evolved across diverse traditions worldwide¹³.

There has been some past work to understand the evolution of flavor compounds and phytochemicals in culinary practices^{10,14}. Several studies have investigated the health benefits of well-known spices and herbs, such as

turmeric, saffron, fennel, and clove^{15–22}. Although their role in disease alleviation is well known, only 63 spices and herbs are tracked by the United States Department of Agriculture (USDA)²³. As of 2024, FooDB²⁴ has listed a total of 70,926 phytochemicals and only 124 spices and herbs, out of 797 foods. Yet, plants have a high chemical diversity with approximately 3000 phytochemicals or more^{23,24}. Still, 85% of these chemicals, which may play a role in disease prevention, remain untracked by national databases, unexplored through experimental research, and unknown to the public at large²⁵. This necessitates a systematic study of the relationships between spices-herbs, phytochemicals, and health conditions.

Researchers have developed network-based frameworks to study phytochemical-disease relationships, largely focused on a single type of phytochemical or disorder, e.g., understanding the impact of polyphenols on cardiovascular health^{26,27}. Tools like PhyteByte²⁸ and HyperFoods²⁹ employ machine learning to identify cancer-fighting molecules in foods, but focus solely on carcinogenic molecules. While these studies provide valuable insights, they do not comprehensively analyze the relationships between spices-herbs, phytochemicals, and health conditions. Rakhi et al.³⁰ attempt to explore spice-phytochemical-disease relationships, but they conduct only a small-scale study with 188 ingredients, yielding only 8957 spice-disease connections. Further, they do not generate phytochemical-indication association hypotheses that can be tested experimentally. Similarly, Gao

¹CSIR-Central Scientific Instruments Organisation, Chandigarh, India. ²Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, U.P., India. ³University of Illinois Urbana-Champaign, Urbana, IL, USA. ✉ e-mail: varshney@illinois.edu

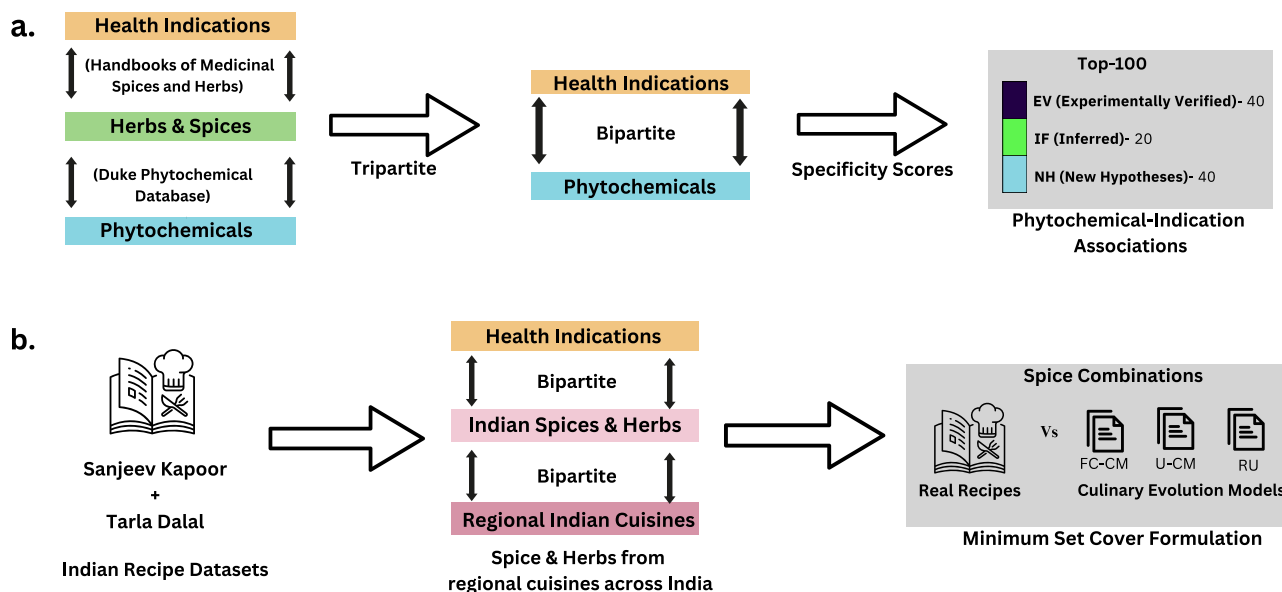


Fig. 1 | Methodology framework for spice and herb phytochemical analysis. **a** The methods used to obtain phytochemical-indication associations. **b** To understand the impact of culinary globalization on health indication coverage.

et al.⁵ used network approaches to study Chinese herb and disease relationships and verified their predictions with real-world patient data. However, the study is limited to Chinese herbs.

The use of ingredients varies across regional cuisines, influenced by factors such as local food availability, climatic conditions, and religious-cultural preferences³¹. While extensive research has been conducted on the health benefits of certain cuisines, such as the Mediterranean diet, these studies have primarily focused on macronutrients and overall dietary composition³². However, the role of phytochemicals, particularly those found in herbs and spices, remains understudied in the context of regional cuisines³³. These bioactive compounds represent a crucial aspect of the “food as medicine” principle and have potential health benefits beyond basic nutrition³⁴. For instance, the Mediterranean diet—rich in olive oil, nuts, and various herbs—contains diverse polyphenols and other phytochemicals that have various health effects³³.

The objectives of our work are twofold. First, our study aims to explore the relationships between spices and herbs, phytochemicals, and health indications using a network-based approach (Fig. 1) and introduce a specificity score to quantify the strength of phytochemical-disease associations. We analyze 1094 herbs and spices to generate 34,113 spice-disease relationships.

With globalization, there has been a notable fusion of cuisines, leading to the adaptation of traditional recipes with new ingredients. The second objective of our work is to investigate whether this culinary evolution offers any health benefits (Fig. 1b). To address this question, we use a dataset of Indian recipes from 18 regions, as spices and herbs are integral to India’s regional cuisines^{35,36}. We estimate the spice and herb usage in different regional cuisines of India and then analyze the spectrum of diseases covered by different regional cuisines based on their spice usage. Further, we generate new spice combinations to simulate culinary globalization using copy-mutate models and compare their effectiveness to traditional Indian recipe spice combinations in terms of broadening health indications. We show that randomly generated spice combinations required fewer ingredients to cover various health indications compared to traditional recipes, suggesting health benefits from culinary fusion.

Results

In this section, we first present results for spice-herb, phytochemical, and indications bipartite networks, systematically analyzing the associations between health indications and phytochemicals. Second, we present results

on spice usage in different Indian regional cuisines using recipe corpora, analyzing the relationships between cuisine and health indications, and comparing the minimum number of spices needed to cover a spectrum of indications using copy-mutate models.

Network analysis of spices & herbs, phytochemicals, and health indications

Two bipartite networks were created: one between spices and phytochemicals, and another between spices and health indications. From these two networks, a third bipartite network was constructed, linking indications and phytochemicals (see Fig. 1a). The first three sub-sections provide a descriptive analysis of the three bipartite networks, and the fourth section makes specific predictions of the indication-phytochemical associations.

A spice-indication bipartite network on two sets of nodes: (i) 1094 spices and herbs and (ii) 1597 medical indications was first built. Datasets were obtained from the Handbook of Medicinal Herbs³⁷ and the Handbook of Medicinal Spices³⁸, providing extensive information about herbs and spices and their associated medical indications. Next, we created bipartite network projections where two nodes representing spices and herbs are connected if they share at least one indication. The Wakita–Tsurumi algorithm³⁹ was applied to this bipartite projection to detect clusters of spices and herbs. For ease of visualization, we used a backbone extraction method^{40,41} to identify statistically significant edges as shown in Fig. 2a. The central role of garlic is quite evident. To delineate the clusters, we extract bar plots of the prevalence scores of indications (see “Prevalence score” section) in each cluster (see Fig. 2b). Notice that indications belonging to different disease categories, such as respiratory ailments, gastrointestinal disorders, infectious diseases, musculoskeletal conditions, and various forms of cancer, have the highest prevalence scores across all clusters.

In Fig. 2b, cancer is found to be the most prevalent indication in cluster 2, represented by onion (*Allium cepa*) and opium poppy (*Papaver somniferum*), as well as cluster 3, represented by thyme (*Thymus vulgaris*) and green or black tea (*Camellia sinensis*), suggesting they contain phytochemicals that are beneficial in cancer prevention and management. Respiratory diseases, including asthma, mucositis, cough, and bronchitis, are the most prevalent in cluster 7, represented by banana (*Musa spp.*) and peppermint (*Mentha × piperita*), and cluster 8, represented by garlic (*Allium sativum*) and black pepper (*Piper nigrum*). Most of the clusters have a high association with at least one gastrointestinal disease. Roughly 80% of the spices in cluster 6 (refer to Fig. 2b), including basil (*Ocimum*

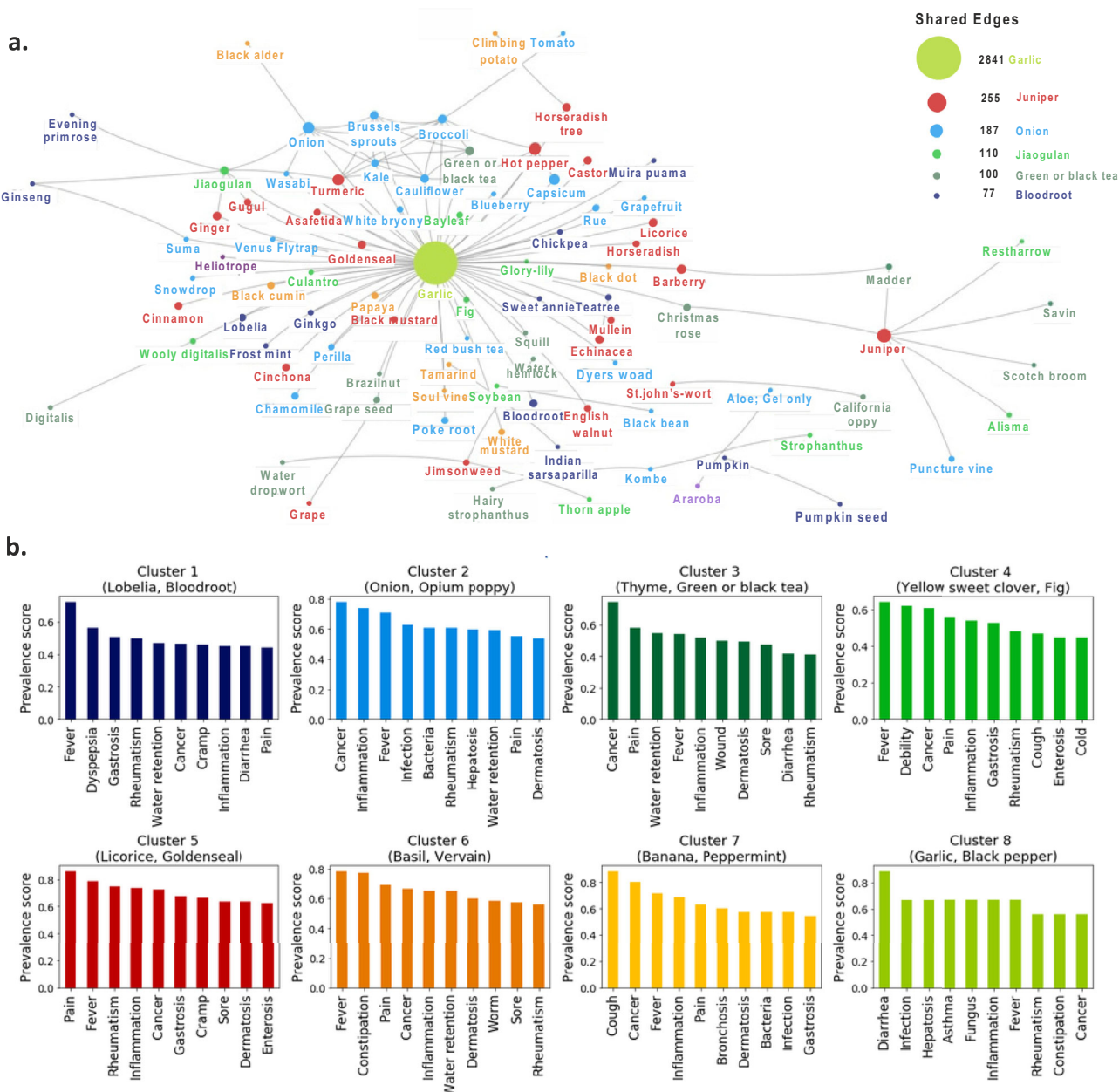


Fig. 2 | Spice and herb association network with indication clustering. **a** Backbone network visualizing connections between various spices and herbs. Each edge represents an association based on a common indication between herbs and spices.

Two spices and herbs are connected if they share a common indication. The node color represents the cluster obtained from the Wakita–Tsurumi algorithm. **b** The bar plots display the prevalent indications associated with each cluster of spices and herbs.

tenuiflorum) and vervain (*Verbena officinalis*), are linked to alleviating constipation. Cluster 4 has a strong association with gastrostis and hepatitis, whereas cluster 8 is strongly associated with hepatitis and constipation. On the level of individual indications, pain, cough, and diarrhea are covered respectively by almost all spices within cluster 5 (represented by licorice (*Glycyrrhiza glabra*) and golden seal (*Hydrastis canadensis*)), cluster 7 (represented by banana and peppermint), and cluster 8 (represented by garlic and black pepper). This structured approach allows us to identify not just individual spices with therapeutic potential but also groups of spices that collectively have a range of health indications.

A second spice-phytochemical bipartite network between (i) 742 spices and herbs and (ii) 2993 bioactive phytochemicals was obtained from the Duke Phytochemical Database⁴² to explore relationships between spices and their constituent phytochemicals. Figure 3 shows the projection graph on phytochemicals clustered using the Wakita–Tsurumi algorithm. The projection graph on spices and herbs is provided as Supplementary Fig. 1 in

Supplementary Section 1. The blue cluster in Fig. 3 on the right primarily consists of terpenes, components of essential oils derived from plants that possess anti-bacterial and anti-inflammatory properties. The light green cluster contains mostly antioxidants, including vanillic acid, quercetin, p-coumaric acid, and caffeic acid. The yellow cluster comprises phytochemicals such as campesterol, stigmasterol, and campesterol, which are plant sterols beneficial for cardiovascular health. The Food and Drug Administration (FDA) has approved that foods containing at least 0.65 g of plant sterol esters per serving, consumed twice daily with meals for a total daily intake of at least 1.3 g, may reduce the risk of heart disease²³. The bottom brown cluster consists of major essential amino acids phenylalanine, methionine, leucine, histidine, and lysine; the absence of these in the diet can lead to decreased immunity, muscle loss, and even mental dysfunction. The dark blue cluster contains a small group of polyunsaturated fatty acids (PUFAs), including linoleic acid, palmitic acid, stearic acid, and oleic acid, commonly found in oils. PUFAs boost immunity in low amounts, but consuming high

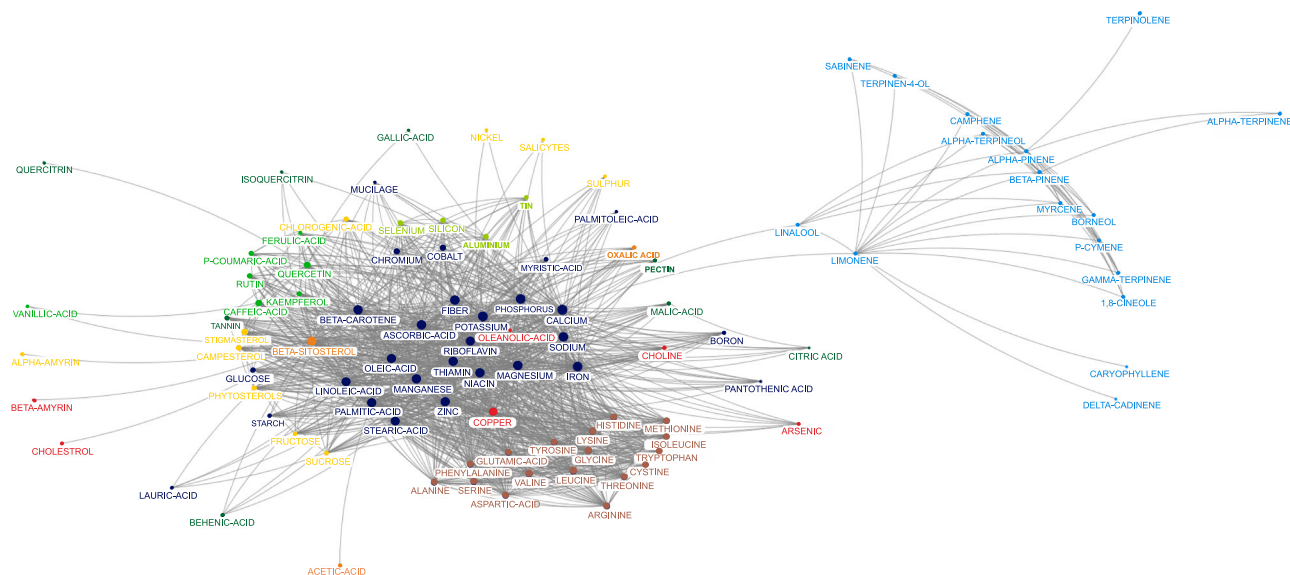


Fig. 3 | Phytochemical association network based on shared spices and herbs. A unipartite backbone network visualizing connections between various phytochemicals. Each edge represents an association based on common spices and herbs

between these phytochemicals. Two phytochemicals are connected if they share a common spice/herb. The node color represents the cluster obtained from the Wakita–Tsurumi algorithm.

amounts of PUFAs with starch can lead to diseases, particularly heart disease and weight gain.

To understand the therapeutic properties of spices, we aimed to identify the constituent phytochemicals that contribute to their disease associations using a third indication-phytochemical bipartite network (refer to Fig. 4). We defined a specificity score to quantify the uniqueness of phytochemicals and their associations with indications (see “Specificity score” section). Notice in Fig. 4a that in endocrine diseases, high specificity was observed for dianethole and p-anisaldehyde with andropause. These phytochemicals are present in fennel and anise and are effective against endocrine diseases and other types of diseases⁴³. The efficacy of 1,2,6-tri-*o*-galloyl-beta-d-glucose—found in *Cornus officinalis*—against protein glycation has been demonstrated, making it effective for reducing blood pressure⁴⁴. Other molecules in the blood disease category did not show high specificity values. Capsaicin and its precursor, vanillylamine, are useful as analgesics and are used in ointments for musculoskeletal pain management, which is evident in the musculoskeletal diseases specificity plots⁴⁵. Other compounds with high specificity in this category include capsorubin and capsanthin, carotenoids found in red bell peppers that are used for pain management⁴⁶, as well as dihydrocapsaicin, a compound from the same capsaicin family. In the metabolic disease specificity plots in Fig. 4d, high values were observed for vitamin K, glucosamine (found in many plants, including aloe vera and *Cannabis sativa*), daidzein (found in soybean), coumestrol (found in soybean, spinach, Brussels sprouts, and legumes), and imperatorin (found in *Ammi majus* and *Angelica archangelica*). Researchers have found that these molecules are effective against fatty liver, steatosis, and hyperuricemia^{47–49}.

Galanthamine, found in *Galanthus nivalis* and other sources, showed high specificity for myasthenia gravis and Alzheimer’s disease in the neurological diseases category (refer to Fig. 4e)⁵⁰. Similarly, high specificity scores were obtained for tigloidine, periplocymarin, cymarin, cymarol, strophanthidin, and tropine for neurological diseases like Parkinson’s and neurodystonia^{51,52}. The high specificity scores for vanillylamine, capsorubin, and other capsaicin family molecules for cluster headaches and diabetic neuropathy are noteworthy and can be observed in Fig. 4e. In the cardiovascular disease category (Fig. 4f), most molecules were non-specific, with some high-scoring specific molecules such as asarinin found in sesame, nitidine found in *Zanthoxylum americanum*, and periplocymarin found in *Strophanthus hispidus*^{53,54}. Some molecules, such as trans-isoasarone found in *Acorus calamus*, show antifungal properties but are toxic and difficult to use for therapeutic purposes⁵⁵.

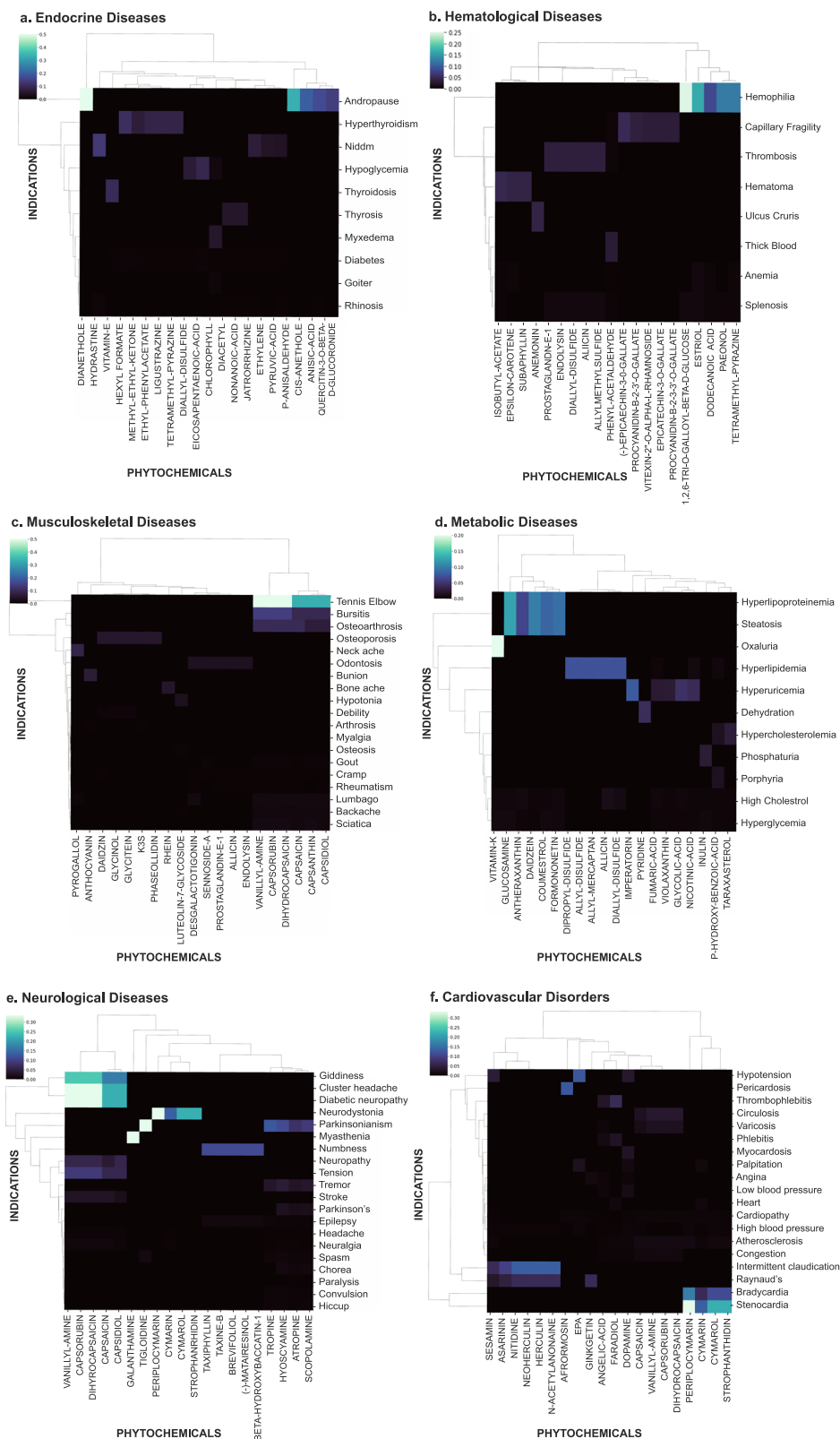
To further assess the capability of the specificity score in discovering new associations and validating known relationships, we conduct a systematic analysis. We focus on the top-100 inferred indication-phytochemical relationships in terms of specificity scores. For each inference, we first compared the results against the indication-chemical relationships provided by the Comparative Toxicogenomics Database (CTD)⁵⁶, a reliable public database containing both curated and inferred relationships. If our inferred relationships were not found in CTD, we manually searched for supporting evidence in other literature using Google Scholar. Out of the 100 top inferences (see Fig. 5), we can validate 60 indication-phytochemical relations through CTD or literature. Among these 60 inferences, 20 could be inferred through gene-chemical interactions and gene-disease associations according to CTD, but have not been experimentally proven yet. The remaining 40 inferences were confirmed through experimental literature. Thus, 20 of our top inferences are new discoveries with high confidence that have also been predicted in CTD through alternative means, 40 are correct predictions backed by experimental scientific evidence, and the remaining 40 are new hypotheses that can be tested with molecular experiments. Indeed, a key use of our specificity score method is to distill novel scientific hypotheses from traditional knowledge of herbs and spices.

Understanding spice usage and their health implications

We use public Indian recipe corpora obtained from Sanjeev Kapoor⁵⁷ and Tarla Dalal⁵⁸ websites, comprising 18 regional cuisines, to understand spice usage patterns across India and their association with disease categories.

To understand the similarity between different regional cuisines of India, we calculated the usage frequencies of different spices in each cuisine (see “Usage and authenticity of spices” section). The principal component analysis (PCA) bi-plot (Fig. 6a) of Indian cuisines and spices usage frequencies, with spices as factors projected on the principal components (PCs), reveals a clear North to South geographical orientation. The plot highlights the significant role of coconut and curry leaves in South Indian cuisines. The dendrogram shows the splitting of South Indian cuisines into distinct regional cuisines, with Andhra and Kerala cuisines exhibiting higher similarity. Gujarati and Jain cuisines, characterized by extensive use of asafoetida (*Ferula assa-foetida*) and the absence of onion and garlic, cluster together and share similarities with Maharashtrian and South Indian cuisines in terms of spice usage.

Fig. 4 | Heatmaps showing the specificity scores obtained for phytochemical-indication pairs for six disease categories. a endocrine diseases **b** hematological diseases **c** musculoskeletal diseases **d** metabolic diseases **e** neurological diseases **f** cardiovascular diseases.



The cluster map (Fig. 6c) provides insights into evolutionary relationships among Indian cuisines based on spice usage. The divergence in spice combinations across Indian cuisines may also be traced to early Vedic traditions and dietary norms³⁹. For example, Brahmin communities often avoided onions and garlic—classified as tamasic foods—resulting in Jain, Gujarati, and some Maharashtrian cuisines embracing asafoetida (*Ferula*

assa-foetida) as a substitute flavoring agent. This is reflected in their clustering in our spice usage analysis (Fig. 6c). Punjabi and Sindhi cuisines demonstrate a lineage to Kashmiri and Mughlai cuisines, as evident from their spice usage patterns. The similarity between Kashmiri and Mughlai cuisines can be attributed to their shared use of saffron (*Crocus sativus*), cardamom (*Amomum subulatum*), and clove (*Syzygium aromaticum*) (Fig.

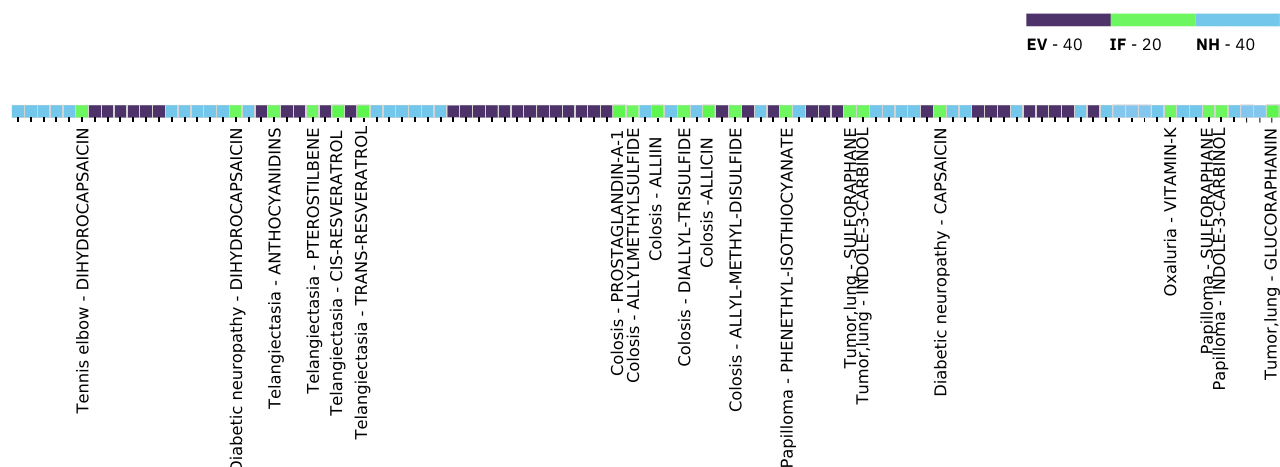


Fig. 5 | Validation of indication-phytochemical relationships. The string map represents the number and nature of relationships extracted from the top-100 indication-phytochemical relationships, categorized as Inferred (IF), Experimentally Verified (EV), and New Hypotheses (NH), and sorted based on specificity scores.

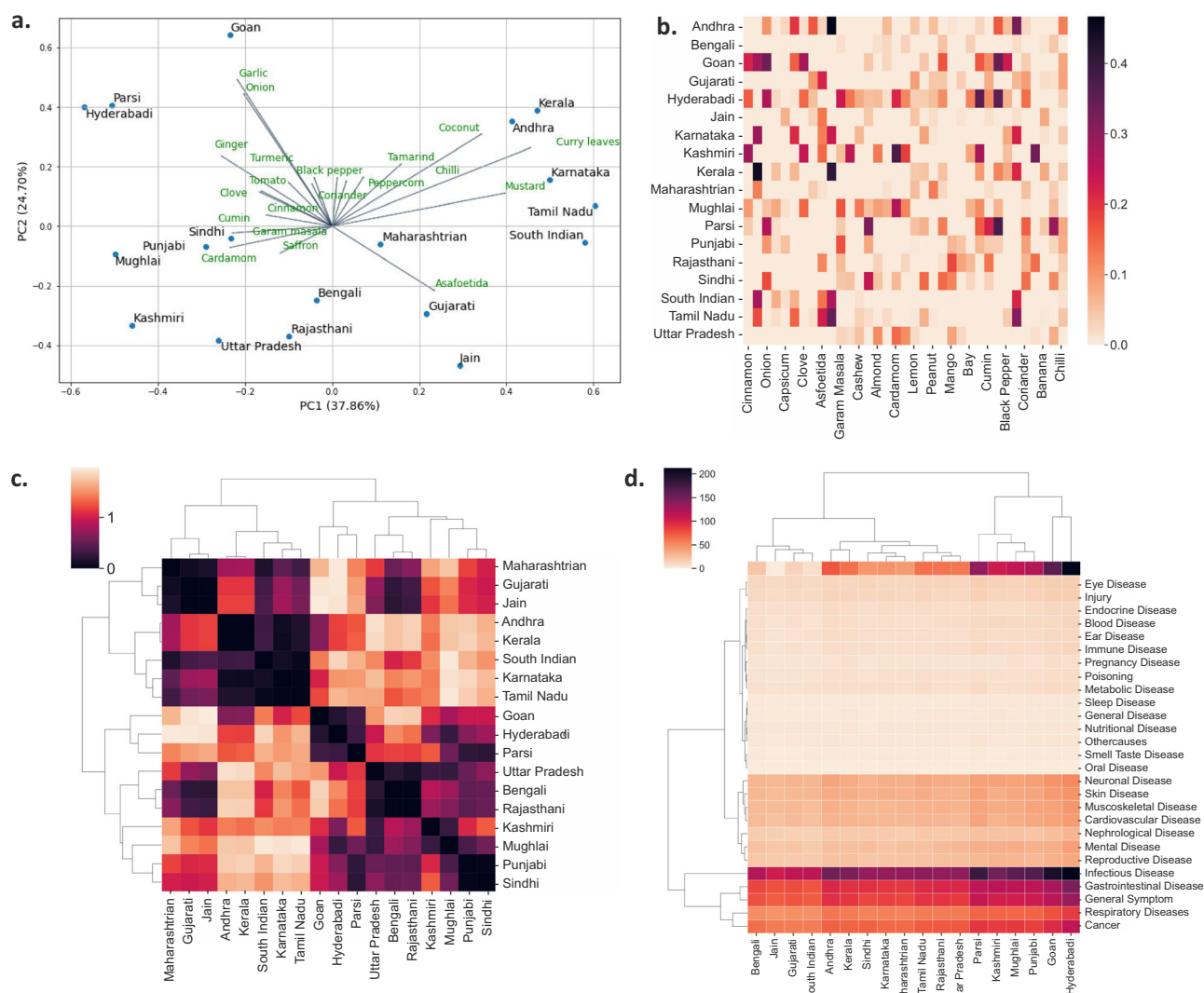


Fig. 6 | Regional cuisine analysis of spice usage and indication coverage. **a** PCA biplot obtained from the frequency of spice usage in different cuisines, representing regions as scores and spices as loadings. PC1 and PC2 account for 37.86% and 24.79% of the total variance, respectively. **b** Heatmap showing the authenticity of spices in Indian cuisines. The darker regions indicate more frequent use of certain

spice pairs. **c** A cluster map was obtained by calculating the cosine similarities between the PCs of different regions. Cuisines that are closer together in the dendrogram have more similar spice usage profiles. **d** Cluster heatmap showing the indication coverage for the different regional cuisines of India.

6a). This lineage may be attributed to the historical spread of Mughal culinary practices across northern India during the 16th–18th centuries. The adoption of saffron, cardamom, and clove in these cuisines mirrors the emphasis on aromatic richness seen in royal Mughal kitchens, as documented in historical manuscripts such as the *Nimatnama* and *Ain-i-Akbari*⁵⁹. Similarly, Hyderabad and Parsi cuisines show resemblance due to their pronounced use of garlic and onion. Coastal cuisines such as Goan, Hyderabad, and Parsi also exhibit culinary patterns shaped by historical trade and colonization. The introduction of ingredients like chili, tomato, and vinegar during the Portuguese colonial era influenced dishes such as vindaloo and xacuti, which later diffused into regional adaptations⁵⁹. The shared use of garlic and onion in Hyderabad and Parsi cuisines further emphasizes these connections. These historical layers have shaped ingredient availability and regional taste preferences, preparation styles, and the symbolic role of spices in culinary identity. However, it is important to note that the similarities observed are based on spice usage data and require further evidence to corroborate the cultural or historical aspects of these connections.

Note that regional variation in spice usage across Indian cuisines may align with underlying genetic differences in taste perception. For instance, South Indian cuisines such as those from Kerala, Tamil Nadu, and Andhra Pradesh make extensive use of bitter-tasting ingredients like mustard seeds (*Sinapis alba*) and curry leaves (*Murraya koenigii*) (see Fig. 6a). This may be related to population-level variation in the CA6 gene, which affects bitter taste sensitivity through the rs2274333 polymorphism. The ancestral A allele is associated with higher gustatory sensitivity (supertasters), while the derived G allele is linked to reduced bitter perception (non-tasters). According to Prakriti et al.⁶⁰, the A allele is more prevalent in western Indian populations, whereas northern and northeastern populations exhibit higher frequencies of the G allele. Although allele frequency data for southern India remains sparse, the strong presence of bitter ingredients in southern cuisines suggests a possible role for chemosensory adaptation. Similarly, cuisines such as Hyderabad and Parsi, which are rich in garlic and onion, may reflect reduced sensitivity to pungent sulfur compounds mediated by TRPV1 polymorphisms. Mughlai, Kashmiri, and Punjabi cuisines—characterized by aromatic spices like saffron, cardamom, and clove—may correlate with population-level variation in olfactory receptor genes such as OR7D4. A similar genetic influence is evident in cilantro preference, where variants in the OR6A2 olfactory receptor gene (e.g., rs72921001) are associated with heightened perception of a soapy flavor in coriander leaf, contributing to population-level differences in its acceptance. These findings support a tentative link between chemosensory genotypes and regional spice practices. However, cuisine evolution is complex and shaped by historical, ecological, and cultural factors beyond genetics^{61–63}.

Figure 6b presents a heatmap of the authentic spices, defined by their unique use in each regional Indian cuisine. While there is a substantial overlap in spice usage across cuisines, the analysis reveals several interesting observations, some well-known and others less so. The presence of asafoetida in Jain and Gujarati cuisines is well-established, as Jains and many Gujaratis exclude onion and garlic from their diet for religious reasons, but asafoetida contains di-allyl sulfur, the same pungent phytochemical as in garlic and onion, making it an ideal substitute⁶⁴. As noted earlier, curry leaves and coconut are integral to South Indian cuisines. Mughlai and Hyderabad cuisines also heavily feature cardamom and clove, while Kashmiri cuisine is uniquely characterized by the presence of saffron and fennel. A lesser-known fact is the use of peanuts as an authentic spice/herb in Maharashtrian cuisine, which is not widely recognized. These findings highlight the diversity and complexity of Indian cuisines, showcasing the interplay of regional preferences, religious influences, and unique spices that define the authentic flavors of each culinary tradition. The three most frequently used spices and the three most authentic spices across Indian cuisines, used to generate the culinary mappings in Fig. 6a, b are detailed in Supplementary Table 1 in Supplementary Information. Notably, chili

(*Capsicum annuum*) is the most frequently used spice across all regional cuisines, reflecting its central role in Indian culinary practices.

Figure 6d presents a heatmap with hierarchical clustering of Indian cuisines based on their indication coverage. It shows that regional cuisines have better coverage for five disease categories: cancer, respiratory diseases, general symptoms, gastrointestinal diseases, and infectious diseases. Hyderabad, Goan, Parsi, Punjabi, and Mughlai cuisines show a broader and stronger coverage of the indication spectrum than the other cuisines. The analysis reveals that Hyderabad and Goan cuisines exhibit the highest scores for alleviating infectious diseases, followed closely by Parsi cuisine.

Each cuisine has a unique profile of herbs and spices, with some having combinations with greater disease mitigation, as observed in Mughlai and Hyderabad cuisine. New fusion cuisines have emerged with increasing globalization as ingredients from different cultures are blended to create new recipes. Here, we study how well combinations of spices from culinary practice cover a spectrum of diseases, using a minimum set-cover algorithm to find the minimum set of spices required to cover a range of indications for each disease category. We then compare the disease coverage capability of recipes generated under four different settings: real settings (using recipes from Tarla Dalal and Sanjeev Kapoor) and three random settings. The random settings simulating culinary globalization include the uniform copy-mutate (U-CM) model, the frequency-conserved copy-mutate (FC-CM) model^{14,36}, and the random uniform (RU) model (see “Random recipe generation” section). To ensure a fair comparison against the real recipes, each recipe in the random settings contained six spice ingredients, which corresponds to the median number of spices per recipe in the real recipe dataset. We generated 50 sets of 5636 recipes for each random model and used the mean size of the minimum recipe sets for comparison. Figure 7 compares the mean size of the minimum set of spices needed to cover health indications in both the random and the actual settings (obtained from the recipe datasets) for 12 different disease categories. For example, to comprehensively address gastrointestinal indications, traditional recipes frequently include cumin, ginger, and fennel, spices known to aid digestion. Conversely, under randomly generated conditions (e.g., the RU model), fewer spices, typically dominated by garlic and turmeric, achieve similar health coverage due to their multi-functional medicinal properties. Similarly, infectious diseases are typically covered by extensive spice combinations such as turmeric, garlic, and black pepper in Mughlai and Hyderabad cuisines, highlighting their comprehensive therapeutic applications within culinary traditions. Note that the size of the minimum set of spices under FC-CM is close to that of the original recipe datasets, which can be attributed to the fact that it conserves the frequency of spices used. For most disease categories, fewer spices are needed to cover the spectrum of indications for U-CM and RU models. These randomly generated recipes require fewer spices to cover infectious, gastrointestinal, and cancer indications than traditional Indian cuisines. This efficiency may be due to spices like garlic that address multiple health concerns and are used extensively in various Indian regional cuisines. However, actual spice usage in cuisines is also influenced by flavor, ingredient interactions, availability, and cultural factors, not just health benefits. Going forward, it is of interest to explore mixing while preserving or enhancing the flavor of the recipes⁶⁵.

Discussion

We took a network-driven approach to discover knowledge regarding spices and herbs, their constituent phytochemicals, health indications, and their culinary usage. While spices and herbs cover many disease categories, general symptoms, respiratory disease, gastrointestinal disease, infectious disease, musculoskeletal disease, and cancer are the most widely covered. From the spice-indication and spice-phytochemical bipartite networks, a third bipartite network between phytochemicals and indications was constructed. We then probed deeper into the indication-phytochemical bipartite network by defining specificity scores, which indicated the degree of association between the indications and phytochemicals.

Various indication-phytochemical associations emerged out of our analysis, such as dianethole, p-anisaldehyde found in fennel, and anise,

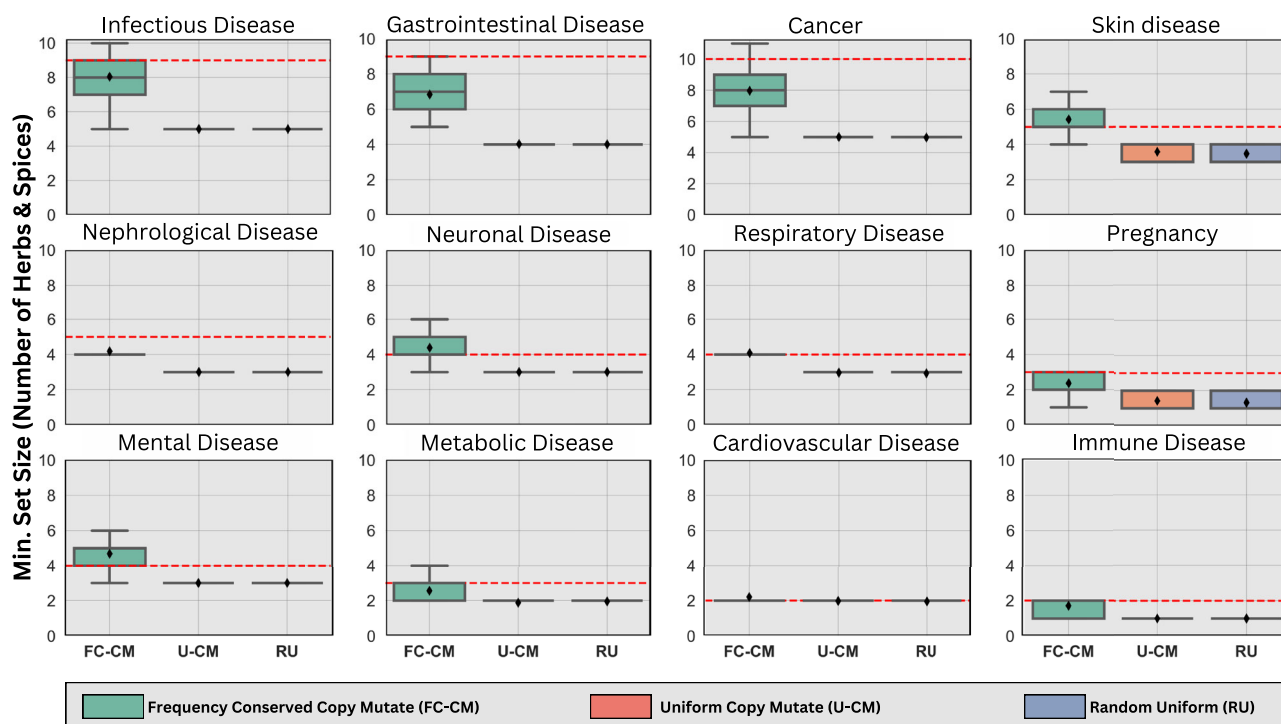


Fig. 7 | Minimum spice set distribution across disease categories via different generative models. Each plot represents the distribution of the minimum number of spices obtained using the minimum set-cover algorithm for the three copy-mutate

models (FC-CM, U-CM, RU) for 12 disease categories. The red line represents the size of the minimum set of spices obtained for the recipe dataset.

which are useful for treating endocrine diseases. For better understanding, we performed a systematic analysis of indication-phytochemical associations by ranking them based on specificity scores. The analysis of the top-100 inferences showed that we achieved a high confidence score within 60 associations, and some of the inferred relationships are yet unknown. This may prove to be a new method for the identification of new phytochemical-disease associations. The known relationships were verified against common gene-chemical and gene-disease associations and thus are biologically relevant. The remaining 40 associations of the top inferences labeled 'NH' are the new predicted associations that could be tested experimentally. Compared to SpiceRx³⁰, our study provides deeper insights by analyzing 1094 herbs and spices to generate 34,113 spice-disease relationships. While the past work does use existing databases like PhenolExplorer⁶⁶ and CTD⁶⁶ to find phytochemical-diseases, we introduced a normalized specificity score to quantify the strength of phytochemical-disease associations, providing higher granularity in assessing unique spice-phytochemical-diseases relationships. Our methods thus address these limitations, providing a more comprehensive and quantitative analysis of spice-disease interactions.

The analysis of Indian cuisines provides insights into the spice usage and geography of the country. There is a trend in the spice usage from North to South. We also found authentic spices in the cuisines and found that spices like asafoetida, curry leaves, cardamom, and clove are important ingredients in the Jain, South Indian, Hyderabadi, and Mughlai cuisines. We further looked into the spice usage in Indian cuisines and their disease associations. One important observation coming out of this analysis is that the spices/herbs used do not cover all disease categories, and it would be beneficial to add other kinds of spices/herbs to the diet. This analysis also indicated that some cuisines, like Hyderabadi, Goan, Punjabi, and Mughlai, provide better coverage of indications than the other cuisines. It points out that mixing some cuisines could further help in indicating coverage, as the minimum number of spices needed to cover the different kinds of diseases is quite small for randomly generated recipes, as compared to the real recipes. This highlights that the fusion of cuisines may allow for a diverse diet and enhance the health benefits of our diet. However, the cultural significance of spices and herbs cannot be understated. Globalization may break traditions,

but it might have positive impacts too. For example, AI-generated recipes can be quite compelling.

Our analysis does not consider the impact of cooking on the stability or transformation of phytochemicals. Recent studies suggest that heat treatment can both degrade and enhance phytochemicals. For example, antioxidant activity has been shown to increase in some spices like cardamom and clove after cooking, possibly due to the release of bound phenolic compounds or formation of new antioxidant molecules during heating⁶⁷. Moreover, compounds in aromatic spices like thyme and rosemary have shown enhanced bioavailability post-cooking and digestion, suggesting that processing can sometimes improve the health benefits of herbs and spices rather than diminish them⁶⁸. Since spices are typically consumed in combinations, their bioactivity may be modulated by synergistic or antagonistic interactions. Future work could explore these interactions more explicitly, potentially modeling combinatorial effects within the spice-indication bipartite network to assess whether certain spice pairings amplify or inhibit specific health benefits. Additionally, note that our generative models for recipe generation are randomized and may not adequately mimic food globalization.

To summarize, the contributions of this paper are twofold. Firstly, we used a network-based approach to develop a method for the generation of scientific phytochemical-health hypotheses by distilling the traditional knowledge of herbs and spices. Then we systematically analyzed the health impacts of culinary traditions and how the global evolution of food may cover more health indications.

Methods

Data, preprocessing, and categorization

We collected the medicinal indications of herbs and spices from the Handbook of Medicinal Spices³⁸ and Handbook of Medicinal Herbs³⁷. These handbooks provide medicinal information for a large collection of essential spices/herbs, along with their cultivation and chemistry. Each spice/herb is associated with a list of indications. There are 1094 spices/herbs and 1597 indications associated with them. We have also collected 2993 constituent phytochemicals of these herbs and spices from the Duke

Phytochemical Database⁴². The indications were labeled into 24 categories using the International Classification of Diseases (ICD-11)⁶⁹ and Malacards⁷⁰. These categories include gastrointestinal, respiratory, cancer, infectious, mental, reproductive, and cardiovascular diseases, among others. In cases where an indication was not found in either database, we searched for its synonyms using web searches and repeated the categorization process. Note that some indications may belong to multiple categories.

To investigate the role of spices in Indian cuisine, we collected recipe data from two popular culinary websites in India by Tarla Dalal and Sanjeev Kapoor. These websites feature a diverse collection of recipes from various Indian and international cuisines. Our combined dataset comprises 13,212 recipes in total, with 3876 recipes sourced from Tarla Dalal⁵⁸ and 9336 recipes from Sanjeev Kapoor⁵⁷. Each recipe entry includes details such as the ingredients used, their quantities, preparation methods, and the associated cuisine or geographical region.

For our analysis, we focused specifically on recipes belonging to Indian cuisines, such as Punjabi, Bengali, and others. We filtered the dataset to include only Indian cuisine recipes containing spices and herbs, resulting in a final dataset of 5636 recipes, with 3595 recipes from Tarla Dalal and 2041 recipes from Sanjeev Kapoor.

For systematic analysis to understand the phytochemical-indication association, we used the Comparative Toxicogenomics Database (CTD)⁵⁶ and published literature. CTD is a reliable public database containing both curated and inferred relationships. The curated relationships are extracted from published literature by CTD curators, while inferred relationships are established through CTD-curated chemical-gene interactions. In the latter case, chemical A is “inferred” to be associated with disease C via gene B if chemical A directly interacts with gene B, and gene B is associated with disease C. We classified the results as *Inferred* if we found a chemical-gene interaction in CTD. If not, then we looked into the literature. Unless we found an associated phytochemical-disease relationship, we classified it as *Experimentally Verified*. If it was not found in either, we classified it as a *New Hypotheses*.

Network construction and visualization

We created three bipartite networks to understand the relationships between spices/herbs, indications, and phytochemicals: spice-indication, spice-phytochemical, and indication-phytochemical. For each bipartite network $G = (U, V, E)$ on two disjoint sets of nodes U and V , projection graphs on both U and V are generated. In a projection graph on set U , two nodes are connected if they are both connected to at least one node in V . The bipartite networks can also be represented by a bi-adjacency matrix B , where the rows represent the nodes of U , and the columns represent the vertices of V . In this matrix, $B_{ij} = 1$ if there is an edge between vertex i of U and vertex j of V , and 0 otherwise. We can project this bi-adjacency matrix to a unipartite projection by calculating BB^T .

The spice-indication and spice-phytochemical bipartite networks were derived directly from the Handbooks of Medicinal Spices and Herbs^{37,38} and the Duke Phytochemical Database⁴², respectively. Indication-phytochemical associations were inferred through an integration of the spice-phytochemical and spice-indication bipartite networks. The indication-phytochemical bipartite network was constructed by linking each constituent phytochemical of a spice directly to all the indications the spice is associated with, and weak links were filtered out.

For visualization purposes, we used a backbone extraction algorithm to abstract the network while preserving the statistically significant links and nodes^{40,41}. The weight of each link was calculated as the total number of shared nodes, and the size of each node was proportional to the number of its connections within the backbone graph. The Wakita-Tsurumi algorithm³⁹ was applied to the full unipartite network to identify groups of nodes that are close in the network topology. NodeXL was used for the visualization of networks.

Prevalence score

The Wakita-Tsurumi algorithm is applied to the spice-indication and spice-phytochemical bipartite networks to identify spice clusters with similar

therapeutic properties and chemical compositions, respectively. To characterize each cluster, we introduce the prevalence score, which quantifies the importance of each indication or phytochemical within a cluster. For a spice cluster K_p , the prevalence score of an indication or phytochemical is calculated by counting the number of spice links associated with it within the cluster in the respective bipartite network and then normalizing the count by dividing it by the total number of spices in the cluster ($|K_p|$). This normalization ensures fair comparisons across clusters of different sizes.

A higher prevalence score indicates a stronger association between an indication or phytochemical and the spices in the cluster, suggesting its importance within that group. Examining prevalence scores within each cluster helps identify key therapeutic properties or chemical compounds characterizing the spices.

Specificity score

To quantify the uniqueness of phytochemicals in their association with a certain disease, we measured a specificity score for each indication-phytochemical pair (u_i, v_j) . The specificity score is computed as the number of links between them, normalized by the product of the count of u_i 's spice associations and the count of v_j 's spice associations:

$$S_{ij} = \frac{e_{ij}}{e_i \cdot e_j} \quad (1)$$

where e_{ij} is the number of edges between nodes u_i and v_j , and e_i and e_j are the total number of edges connected with nodes u_i and v_j , respectively.

Usage and authenticity of spices

The authenticity score¹⁰ measures how unique or representative a spice or herb is to a specific cuisine compared to its usage in other cuisines.

The usage frequency F_i^c captures the proportion of recipes within a cuisine c that include the spice or herb i , and it is defined as:

$$F_i^c = \frac{n_i^c}{N^c} \quad (2)$$

where n_i^c represents the number of recipes in cuisine c that contain spice or herb i , and N^c represents the total number of recipes in cuisine c . A higher value of F_i^c indicates that the spice or herb i is more commonly used in cuisine c . Next, we calculate the average usage frequency of spice or herb i across all other cuisines except c :

$$\langle F_i^c \rangle_{c' \neq c} = \frac{\sum_{c' \neq c} F_i^{c'}}{|C| - 1} \quad (3)$$

where c' represents all cuisines other than c , and $|C|$ represents the total number of cuisines. The term $\langle F_i^c \rangle_{c' \neq c}$ represents the average usage frequency of spice or herb i in all cuisines other than c . It provides a baseline for comparing the usage of spice or herb i in cuisine c to its usage in other cuisines. Finally, the authenticity score of spice or herb i in cuisine c is defined as:

$$A_i^c = F_i^c - \langle F_i^c \rangle_{c' \neq c} \quad (4)$$

The authenticity score A_i^c measures the relative usage frequency of spice or herb i in cuisine c compared to its average usage frequency in all other cuisines. A positive value of A_i^c indicates that spice or herb i is used more frequently in cuisine c than in other cuisines, suggesting that it is more authentic or representative of cuisine c . Conversely, a negative value of A_i^c indicates that spice or herb i is used less frequently in cuisine c compared to other cuisines, suggesting that it is less authentic or representative of cuisine c .

Indication coverage in cuisines

The spice-indication matrix, M_{SI} , capturing the association between spices and their indications, is defined as:

$$M_{SI}[s, i] = \begin{cases} 1 & \text{if spice } s \text{ is associated with indication } i \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

The cuisine-spice matrix M_{CS} represents the usage of spices in different cuisines as defined in equation (2). The cuisine-indication matrix M_{CI} representing the indication coverage of cuisines is calculated as:

$$M_{CI} = M_{CS} \times M_{SI} \quad (6)$$

Each element of M_{CI} represents the strength of association between cuisine c and indication i .

Minimum set-cover algorithm

To find the minimum subset of spices that can cover a specific set of indications, we use the minimum set-cover problem on a bipartite network. This approach helps in understanding the efficiency and optimization of spice usage in regional cuisines for potential health benefits. Given a bipartite network $G = (U, V, E)$, where U and V are two disjoint sets of nodes representing spices and diseases, respectively, and E is the set of edges connecting them, our goal is to identify the smallest subset of nodes in U that covers all the nodes in a specified subset of V .

$$\text{Minimize } \mathbf{c}^T \mathbf{X} \quad (7)$$

$$\begin{aligned} \text{Subject to : } & 0 \leq \mathbf{X}_j \leq 1 \quad \text{for } j = 1, 2, \dots, |V| \\ & \sum_j \mathbf{B}_{ij} \mathbf{X}_j \geq 1 \quad \text{for } i = 1, 2, \dots, |U| \end{aligned}$$

We approach this problem by formulating it as an integer linear programming problem. Let $|U|$ and $|V|$ denote the number of nodes in sets U and V , respectively. The problem was formulated with the constraints expressed as:

$$\mathbf{GX} \leq \mathbf{h} \quad (8)$$

where,

$$\mathbf{G} = [\mathbf{B}, \mathbf{I}_{|V|}, -\mathbf{I}_{|V|}]^T \quad (9)$$

$$\mathbf{h} = [-\mathbf{1}_{|U|}, \mathbf{1}_{|V|}, \mathbf{0}_{|V|}]^T \quad (10)$$

\mathbf{B} is the adjacency matrix of the bipartite network with dimension $|V| \times |U|$, where $|U|$ and $|V|$ are the total number of nodes in sets U and V , respectively. Each entry of \mathbf{B} is either 1 or 0, indicating whether or not there exists at least one link between each node pair between U and V . Here, \mathbf{X} is a vector of size $|V|$ to be solved, and each entry of \mathbf{X} is either 1 or 0, indicating whether or not the node from set U should be included in the minimum set cover. We obtain a fractional solution for \mathbf{X} , which is then rounded to obtain a feasible solution to the original minimum set-cover problem. The rounded solution represents the minimum subset of spices (nodes from U) that covers all the diseases (nodes in the specified subset of V) in the bipartite network G .

Random recipe generation

We constructed three random recipe datasets using different models: frequency-conserved copy-mutate (FC-CM), uniform copy-mutate (U-CM), and random uniform (RU)^{14,36}. These models aim to simulate the process of recipe creation and evolution over time. The FC-CM algorithm begins by creating an initial random pool I_0 of 10 spice ingredients and a seed pool R_0 of 20 recipes. Each recipe in R_0 is generated by randomly selecting $S = 6$ spice ingredients from I_0 . The value of S is chosen based on the median number of spice ingredients per recipe in the real recipe dataset.

Each spice ingredient in I_0 is assigned a fitness value based on its empirical frequency in the real data, reflecting how frequently it is used across different cuisines. Spice ingredients with higher frequencies in real-world recipes are considered more fit and versatile, and thus have a higher chance of being selected for a recipe.

At each time step, a mother recipe is randomly selected from the recipe pool R_0 . A copy of this mother recipe is made, and the copy undergoes a mutation process. During mutation, an ingredient with fitness f_i is randomly chosen from the copied recipe and compared with another ingredient with fitness f_j , which is randomly selected from the ingredient pool I_0 . If $f_j > f_i$, the old ingredient i is replaced with the new ingredient j . This mutation process is repeated $M = 6$ times, after which the mutated copy recipe is added back to the recipe pool R_0 , becoming a potential candidate for selection as a mother recipe in the next time step. To maintain a diverse ingredient pool, the ratio r between the size of the ingredient pool I_0 and the size of the recipe pool R_0 is checked at the beginning of each time step. If r falls below a threshold of 0.2, new spice ingredients are introduced to I_0 by randomly selecting from the list of all available spice ingredients. These new ingredients are added to the existing pool, expanding the variety of spices available for recipe creation. The FC-CM process continues until the desired number of recipes is reached, which in this case is 5636, matching the number of recipes in the real recipe dataset.

The U-CM model follows a similar copy-mutate process as the FC-CM model, with the key difference being that no fitness value is assigned to each spice ingredient. In the U-CM model, each spice has an equal probability of being selected for a recipe, and the mutation process involves replacing a randomly chosen ingredient with another randomly selected ingredient from the pool.

In the RU model, recipes are constructed by randomly choosing ingredients with uniform probability, without considering any fitness values or mutation processes.

Data availability

The data used in this study was obtained from public sources and has been referenced in the manuscript. Extended results that further support the findings of the study are provided as Supplementary Information. The top-100 indication-phytochemical relationship results are also given in Supplementary Data 1.

Code availability

Code is publicly available in the GitHub repository: https://github.com/rishemjit/Spices_Herbs_ML.

Received: 14 November 2024; Accepted: 17 May 2025;

Published online: 29 May 2025

References

- Lewis, K. & Ausubel, F. M. Prospects for plant-derived antibacterials. *Nat. Biotechnol.* **24**, 1504–1507 (2006).
- Pino-Otín, M. R. et al. Antibiotic properties of *Satureja Montana* L. hydrolase in bacteria and fungus of clinical interest and its impact in non-target environmental microorganisms. *Sci. Rep.* **12**, 18460 (2022).
- Ullah, F. et al. Potential role of plant extracts and phytochemicals against foodborne pathogens. *Appl. Sci.* **10**, 4597 (2020).
- Laldingliani, T. B. C. et al. Ethnomedicinal study of medicinal plants used by Mizo tribes in Champhai district of Mizoram, India. *J. Ethnobiol. Ethnomed.* **18**, 22 (2022).
- Gan, X. et al. Network medicine framework reveals generic herb-symptom effectiveness of traditional Chinese medicine. *Sci. Adv.* **9**, eadh0215 (2023).
- Gonellimali, F. D. et al. Antimicrobial properties and mechanism of action of some plant extracts against food pathogens and spoilage microorganisms. *Front. Microbiol.* **9**, 01639 (2018).
- Lai, P. K. & Roy, J. Antimicrobial and chemopreventive properties of herbs and spices. *Curr. Med. Chem.* **11**, 1451–1460 (2004).

8. Fraenkel, G. S. The Raison d'Être of secondary plant substances. *Science* **129**, 1466–1470 (1959).
9. Guldiken, B. et al. Phytochemicals of herbs and spices: health versus toxicological effects. *Food Chem. Toxicol.* **119**, 37–49 (2018).
10. Ahn, Y.-Y., Ahnert, S. E., Bagrow, J. P. & Barabási, A.-L. Flavor network and the principles of food pairing. *Sci. Rep.* **1**, 196 (2011).
11. Geck, M. S. et al. Traditional herbal medicine in Mesoamerica: toward its evidence base for improving universal health coverage. *Front. Pharmacol.* **11**, 1160 (2020).
12. de Jesús Dzul-Beh, A. et al. Antimicrobial potential of the Mayan medicine plant *Matayba oppositifolia* (A. Rich.) Britton against antibiotic-resistant priority pathogens. *J. Ethnopharmacol.* **300**, 115738 (2023).
13. Sherman, P. W. & Billing, J. Darwinian gastronomy: why we use spices: Spices taste good because they are good for us. *BioScience* **49**, 453–463 (1999).
14. Kinouchi, O., Diez-Garcia, R. W., Holanda, A. J., Zambianchi, P. & Roque, A. C. The non-equilibrium nature of culinary evolution. *N. J. Phys.* **10**, 073020 (2008).
15. Mafra, D. et al. Food as medicine: targeting the uraemic phenotype in chronic kidney disease. *Nat. Rev. Nephrol.* **17**, 153–171 (2021).
16. Li, H. et al. Curcumin, the golden spice in treating cardiovascular diseases. *Biotechnol. Adv.* **38**, 107343 (2020).
17. Lopresti, A. L. & Drummond, P. D. Saffron (*Crocus sativus*) for depression: a systematic review of clinical studies and examination of underlying antidepressant mechanisms of action. *Hum. Psychopharmacol.* **29**, 517–527 (2014).
18. Diao, W.-R., Hu, Q.-P., Zhang, H. & Xu, J.-G. Chemical composition, antibacterial activity, and mechanism of action of essential oil from seeds of fennel (*Foeniculum vulgare* mill.). *Food Control* **35**, 109–116 (2014).
19. Otunola, G. A. Culinary spices in food and medicine: an overview of *Syzygium aromaticum* (L.) Merr and L. M. Perry [Myrtaceae]. *Front. Pharmacol.* **12**, 793200 (2022).
20. Rowles, J. L. & Erdman, J. W. Carotenoids and their role in cancer prevention. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1865**, 158613 (2020).
21. Kumar, A. et al. Major phytochemicals: recent advances in health benefits and extraction method. *Molecules* **28**, 887 (2023).
22. Singletary, K. Turmeric: potential health benefits. *Nutr. Today* **55**, 45–56 (2020).
23. US Food and Drug Administration. CFR - Code of Federal Regulations Title 21. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=101.83> (2020).
24. Harrington, R. A., Adhikari, V., Rayner, M. & Scarborough, P. Nutrient composition databases in the age of big data: foodDB, a comprehensive, real-time database infrastructure. *BMJ Open* **9**, e026652 (2019).
25. Barabási, A.-L., Menichetti, G. & Loscalzo, J. The unmapped chemical complexity of our diet. *Nat. Food* **1**, 33–37 (2020).
26. Do Valle, I. F. et al. Network medicine framework shows that proximity of polyphenol targets and disease proteins predicts therapeutic effects of polyphenols. *Nat. Food* **2**, 143–155 (2021).
27. Mozaffarian, D. Nutrition's dark matter of polyphenols and health. *Nat. Food* **2**, 139–140 (2021).
28. Westerman, K. E., Harrington, S., Ordovas, J. M. & Parnell, L. D. Phytebyte: identification of foods containing compounds with specific pharmacological properties. *BMC Bioinform* **21**, 238 (2020).
29. Veselkov, K. et al. Hyperfoods: Machine intelligent mapping of cancer-beating molecules in foods. *Sci. Rep.* **9**, 9237 (2019).
30. Rakhi, N. K., Tuwani, R., Mukherjee, J. & Bagler, G. Data-driven analysis of biomedical literature suggests broad-spectrum benefits of culinary herbs and spices. *PLoS ONE* **13**, e0198030 (2018).
31. Harrington, R. J. Defining gastronomic identity. *J. Culinary Sci. Technol.* **4**, 129–152 (2005).
32. Sofi, F., Abbate, R., Gensini, G. F. & Casini, A. Accruing evidence on benefits of adherence to the Mediterranean diet on health: an updated systematic review and meta-analysis. *Am. J. Clin. Nutr.* **92**, 1189–1196 (2010).
33. Tosti, V., Bertozzi, B. & Fontana, L. Health benefits of the Mediterranean diet: metabolic and molecular mechanisms. *J. Gerontol. A Biol. Sci. Med. Sci.* **73**, 318–326 (2017).
34. Huang, Y., Edirisinghe, I., Burton-Freeman, B. M. & Sandhu, A. K. Characterization and pharmacokinetic profile of herbs and spices' phytochemicals over 24 h after consumption in overweight/obese adults. *Mol. Nutr. Food Res.* **67**, 2200785 (2023).
35. Jain, A., Rakhi, N. K. & Bagler, G. Analysis of food pairing in regional cuisines of India. *PLoS ONE* **10**, e0139539 (2015).
36. Jain, A. & Bagler, G. Culinary evolution models for Indian cuisines. *Physica A* **503**, 170–176 (2018).
37. Duke, J. A. *Handbook of Medicinal Herbs* (CRC Press, Boca Raton, 2002).
38. Duke, J. A. *Handbook of Medicinal Spices* (CRC Press, Boca Raton, 2002).
39. Wakita, K. & Tsurumi, T. Finding community structure in mega-scale social networks. In *Proceedings of the 16th International Conference on World Wide Web* (2007).
40. Ghalmane, Z., Cherifi, C., Cherifi, H. & El Hassouni, M. Extracting backbones in weighted modular complex networks. *Sci. Rep.* **10**, 15539 (2020).
41. Yassin, A., Haidar, A., Cherifi, H., Seba, H. & Togni, O. An evaluation tool for backbone extraction techniques in weighted complex networks. *Sci. Rep.* **13**, 17000 (2023).
42. Duke, J. & Bogenschutz, M. J. *Dr. Duke's Phytochemical and Ethnobotanical Databases* (CRC Press, Boca Raton, 1994).
43. Badgujar, S. B., Patel, V. V. & Bandivdekar, A. H. *Foeniculum Vulgare* mill: a review of its botany, phytochemistry, pharmacology, contemporary application, and toxicology. *BioMed Res. Int.* **2014**, 42674 (2014).
44. Lee, J. et al. Galloyl glucoses from the seeds of *Cornus officinalis* with inhibitory activity against protein glycation, aldose reductase, and cataractogenesis ex vivo. *Biol. Pharm. Bull.* **34**, 443–446 (2011).
45. Fattori, V., Hohmann, M. S., Rossaneis, A. C., Pinho-Ribeiro, F. A. & Verri, W. A. Capsaicin: current understanding of its mechanisms and therapy of pain and other pre-clinical and clinical uses. *Molecules* **21**, 844 (2016).
46. Matsufuji, H., Nakamura, H., Chino, M. & Takeda, M. Antioxidant activity of capsanthin and the fatty acid esters in paprika (*Capsicum annum*). *J. Agric. Food Chem.* **46**, 3468–3472 (1998).
47. Ferramosca, A., Di Giacomo, M. & Zara, V. Antioxidant dietary approach in treatment of fatty liver: new insights and updates. *World J. Gastroenterol.* **23**, 4146–4157 (2017).
48. Nasser, M. et al. Effects of imperatorin in the cardiovascular system and cancer. *Biomed. Pharmacother.* **120**, 109401 (2019).
49. Wroblewski Lissin, L. & Cooke, J. P. Phytoestrogens and cardiovascular health. *J. Am. Coll. Cardiol.* **35**, 1403–1410 (2000).
50. Pohanka, M. Inhibitors of acetylcholinesterase and butyrylcholinesterase meet immunity. *Int. J. Mol. Sci.* **15**, 9809–9825 (2014).
51. Sanghvi, I., Bindler, E. & Gershon, S. Pharmacology of a potential anti-Parkinson agent: Tigloidine. *Eur. J. Pharmacol.* **4**, 246–253 (1968).
52. Trautner, E. M. & Noack, C. H. Tigloidine as a substitute for atropine in the treatment of Parkinsonism. *Med. J. Australia* **1**, 751–754 (1951).
53. Lee, W., Ku, S.-K., Kim, J. A., Lee, T. & Bae, J.-S. Inhibitory effects of epi-sesamin on HMGB1-induced vascular barrier disruptive responses in vitro and in vivo. *Toxicol. Appl. Pharmacol.* **267**, 201–208 (2013).
54. Martey, O. N. et al. Periplocymarin is a potential natural compound for drug development: highly permeable with absence of p-glycoprotein efflux and cytochrome P450 inhibitions. *Biopharm. Drug Dispos.* **35**, 195–206 (2014).
55. Perrett, S. & Whitfield, P. J. Anthelmintic and pesticidal activity of *Acorus gramineus* (araceae) is associated with phenylpropanoid asarones. *Phytother. Res.* **9**, 405–409 (1995).

56. Mattingly, C. J., Colby, G. T., Forrest, J. N. & Boyer, J. L. The comparative toxicogenomics database (CTD). *Environ. Health Perspect.* **111**, 793–795 (2003).
57. Kapoor, S. <https://www.sanjeevkapoor.com/>.
58. Dalal, T. <https://www.tarladalal.com/>.
59. Antani, V. & Mahapatra, S. Evolution of Indian cuisine: a socio-historical review. *J. Ethn. Foods* **9**, 15 (2022).
60. Prakrithi, P. et al. Landscape of variability in chemosensory genes associated with dietary preferences in Indian population: analysis of 1029 Indian genomes. *Front. Genet.* **13**, 878134 (2022).
61. Deshaware, S. & Singhal, R. Genetic variation in bitter taste receptor gene *tas2r38*, prop taster status and their association with body mass index and food preferences in Indian population. *Gene* **627**, 363–368 (2017).
62. Eriksson, N. et al. A genetic variant near olfactory receptor genes influences cilantro preference. *Flavour* **1**, 22 (2012).
63. Spence, C. Coriander (cilantro): A most divisive herb. *Int. J. Gastron. Food Sci.* **33**, 100779 (2023).
64. El-Saber Batiha, G. et al. Chemical constituents and pharmacological activities of garlic (*Allium sativum* L.): a review. *Nutrients* **12**, 872 (2020).
65. Varshney, L. R. et al. A big data approach to computational creativity: the curious case of Chef Watson. *IBM J. Res. Dev.* **63**, 7:1–7:18 (2019).
66. Rothwell, J. A. et al. Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content. *Database* **2013**, bat070 (2013).
67. Chiang Chan, E. W. et al. Effects of different cooking methods on the bioactivities of some spices. *Emir. J. Food Agric.* **27**, 610–616 (2015).
68. Opara, E. I. & Chohan, M. Culinary herbs and spices: their bioactive properties, the contribution of polyphenols and the challenges in deducing their true health benefits. *Int. J. Mol. Sci.* **15**, 19183–19202 (2014).
69. World Health Organization. International classification of diseases eleventh revision (icd-11). <https://icd.who.int/en>.
70. Rappaport, N. et al. Malacards: an integrated compendium for diseases and their annotation. *Database* **2013**, bat018 (2013).

Acknowledgements

The authors would like to acknowledge the funding support from CSIR ATLAS mission HCP0031.

Author contributions

These authors contributed equally: R. Kaur, S.Z. S.Z. was responsible for software development, methodology design, formal analysis, data curation, and wrote the original draft. R. Kaur contributed to methodology and

software development, conducted formal analysis, and participated in both drafting and editing the manuscript. R. Kumar was involved in data curation and formal analysis and contributed to writing the original draft and its review. B.B. supported the project through formal analysis, manuscript review, and editing. S.R. contributed to the visualization and was involved in the review and editing of the manuscript. L.R.V. conceptualized the study, developed analysis methods, provided supervision, and contributed to writing through review and editing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41538-025-00458-z>.

Correspondence and requests for materials should be addressed to Lav R. Varshney.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025