Article

Cluster Analysis of Medicinal Plants and Targets Based on Multipartite Network

Namgil Lee [1,2], **Hojin Yoo** [2] **and Heejung Yang** [2,3,*]

1   Department of Information Statistics, Kangwon National University, Chuncheon 24341, Korea; namgil.lee@bionsight.com
2   Bionsight, Incorporated, Chuncheon 24341, Korea; hojinyoo@bionsight.com
3   Laboratory of Natural Products Chemistry, Department of Pharmacy, Kangwon National University, Chuncheon 24341, Korea
*   Correspondence: heejyang@kangwon.ac.kr; Tel.: +82-33-250-6919

Abstract: Network-based methods for the analysis of drug-target interactions have gained attention and rely on the paradigm that a single drug can act on multiple targets rather than a single target. In this study, we have presented a novel approach to analyze the interactions between the chemicals in the medicinal plants and multiple targets based on the complex multipartite network of the medicinal plants, multi-chemicals, and multiple targets. The multipartite network was constructed via the conjunction of two relationships: chemicals in plants and the biological actions of those chemicals on the targets. In doing so, we introduced an index of the efficacy of chemicals in a plant on a protein target of interest, called target potency score (TPS). We showed that the analysis can identify specific chemical profiles from each group of plants, which can then be employed for discovering new alternative therapeutic agents. Furthermore, specific clusters of plants and chemicals acting on specific targets were retrieved using TPS that suggested potential drug candidates with high probability of clinical success. We expect that this approach may open a way to predict the biological functions of multi-chemicals and multi-plants on the targets of interest and enable repositioning of the plants and chemicals.

Keywords: medicinal plants; multi-chemicals; multi-targets; multipartite network; network analysis

1. Introduction

Plants have been used as therapeutic agents for a very long time in the human history [1,2]. Single bioactive compounds in plants have been developed as well-known commercial drugs, such as paclitaxel, vinblastine, morphine, and artemisinin [3,4]. Occasionally, single components from natural products have provided chemical scaffolds for developing more potent semi-synthetic drugs, such as rosuvastatin, tramadol, and eribulin [5]. Similar to these single compound drugs that act on a specific target, a mixture of components from the whole plants or their sub-fractions have been widely used as botanical drugs, such as Veregen® and Fulyzaq® [6]. However, the synergistic effects of multi-chemicals in plants on multiple targets and the underlying complex biological pathways involved are less understood.

For the purpose of developing drugs composed of multi-chemicals, the identification of the interactions between the multi-chemicals and multi-targets is important to understand the therapeutic effects as well as the side effects, i.e., the on-target and off-target effects [7–14]. Identifying the synergistic drug combinations is also critical to develop effective therapies against cancers and to improve the therapeutic efficacy [15–18]. However, the complexity of screening all possible candidate combinations of multi-chemicals and multi-targets prevents the use of the classical way of identifying the interactions via biological experiments, using in vitro and in vivo assays. Since a medicinal plant produces a large number of chemical compounds, the target identification of every single chemical is

impossible for the development of botanical drugs, especially to acquire approval in the global market requiring well-controlled clinical studies [19].

In recent decades, computational approaches for the systematic analysis of drug-target interactions have become available, improving efficiency and reducing costs of studying the interactions [20–23]. Among the computational approaches, network-based approaches provide alternative methods to deal with the complex molecular interactions [24,25]. Compared with other computational methods, network-based approaches do not require three-dimensional structures of the targets or negative control samples [26]. These approaches have also been proposed for botanical drug design in search of optimal combinations of medicinal plant components and prediction of multiple target activities [27]. However, studies based on the integrated networks of medicinal plants, chemicals, and multi-targets are still rare. Most studies have focused only on drug-target networks [28] or the molecular networks of traditional Chinese medicinal plants [29].

In this study, we introduced an approach based on the integrated network of medicinal plants, chemical components, and multi-targets to analyze the complex biological roles of medicinal plants in human body. First, we constructed an integrated multipartite network of the biomedical entities including Korean medicinal plants, chemicals, and protein targets from heterogeneous databases. The databases have been curated for biomedical entities, such as genes, proteins, chemicals, and species and are publicly accessible. In particular, we employed the ChEMBL database, which is one of the largest public databases of the activities of small molecules on biological targets [30–33]. It also provides the built-in, ligand-based target prediction methods that we used to generate the links between chemicals with targets in the interaction network. We also used the database on compound ingredients extracted from Korean medicinal plants, which was released in 2007 by the Korean Traditional Knowledge Portal (http://www.koreantk.com, accessed on 1 May 2019) [34].

We then preprocessed the multipartite network and adapted the spectral co-clustering method to identify the subnetworks of closely related biomedical entities within their own structural properties, such as the average bipartite clustering coefficient [35]. The clustering results could be used for the single or combination use of Korean medicinal plants in common target domains. The hierarchical clustering analysis was followed for further identification of the plant clusters based on the similarities in their chemical composition. In the analysis, we introduced a network-based mathematical index of interactions between the medicinal plants and targets, called the target potency score (TPS). The suggested index is used for ranking the plant clusters or protein targets, and for retrieving the best subnetworks of the medicinal plants, multi-chemicals, and multi-targets customized to specific medicinal purposes. Finally, we have demonstrated the utility of the simplified multipartite network with the hierarchical clusters.

2. Materials and Methods

2.1. Raw Data Sets

We employed several heterogeneous data sets to construct a multipartite network of medicinal plants, chemicals, and targets. The data set of the medicinal plants was obtained from the Korean Traditional Knowledge Portal (KTKP, http://www.koreantk.com, accessed on 1 May 2019). The database was built by the Korean Intellectual Property Office, Ministry of Trade, Industry and Energy, Republic of Korea [34]. The KTKP data sets contain a list of 5500 Korean medicinal plants and 10,056 chemicals, and their relations. The data set of chemicals and targets was obtained from the ChEMBL database (ChEMBL25, http://ftp.ebi.ac.uk/pub/databases/chembl/target_predictions, accessed on 15 March 2019). The ChEMBL25 database contains 1,879,206 compounds, 12,482 targets, and 15,504,603 bioactivity entries from 72,271 publications. All the chemicals in the KTKP data sets were identified in the ChEMBL25 database, and the links between the chemicals and single-protein targets were generated using the prediction model reported previously [33,36], and is available on the ChEMBL25 database. Specifically, the probability of

the activation of a compound against a target was estimated by the 10-uM activity cut-off model in the ChEMBL25 database, and we determined the interaction between a pair of a chemical and a target at the probability of >0.9.

2.2. Preprocessing

We excluded non-plants from the raw dataset. For instance, minerals (non-living things) and animals (including corals) were removed, but the fungi were not. After excluding the non-plants, the raw data sets containing 5500 instances of plants and non-plants were reduced to a set containing 3614 plants. Then, the duplicated instances were removed based on their identical "scientific name," "Latin name" and "use target", which led to a remaining list of 2886 plants in the data set.

The multipartite network of the medicinal plants, chemicals, and targets was constructed by combining the two bipartite networks: the plants-chemicals network and chemicals-targets network (Figure 1A). The nodes in the network represented the plants, chemicals, and targets, and the edges in the network represented the relationships between the plants and chemicals, and between the chemicals and targets. From the multipartite network, we removed the nodes of degree zero, i.e., the plants, chemicals, or targets that were not linked to any other entities. After all, 1138 plants, 10,043 chemicals, and 441 targets remained for the combined multipartite network. The basic statistics of the preprocessed network data are summarized in Table 1.

2.3. Methods

We analyzed the multipartite network using the spectral co-clustering algorithm [35]. The algorithm clusters the medicinal plants, chemicals, and targets by solving a network partitioning problem. It partitions the set V of the connected nodes into k subsets, $V_1, V_2, \ldots, V_k$ in such a way that the k subsets approximately minimize the normalized cut criterion:

$$Q(V_1, V_2, \ldots, V_3) = \frac{1}{k} \sum_{i=1}^{k} \frac{\text{cut}(V_i, \overline{V_i})}{\text{vol}(V_i)} \tag{1}$$

where $\overline{V_i} = V \setminus V_i$ is the complement of V_i, the $cut(M, N) = \sum_{m \in M, n \in N} w_{mn}$ of two disjoint sets M and N is defined by the sum of the edge weights, w_{mn}, between two nodes from each of M and N, and $vol(V_i) = \sum_{m \in V_i} d_m = \sum_{m \in V_i} \sum_{n} w_{mn}$ the sum of degrees of all nodes in V_i. The normalized cut criterion represents the level of connectivity between the k subsets, where a smaller value indicates a weaker connectivity between them.

To determine the number of clusters, k, we computed the modularity, defined by

$$\text{Modularity}(k) = \frac{1}{2E} \sum_{i=1}^{k} \sum_{m \in V_i, n \in V_i} \left(w_{mn} - \frac{d_m d_n}{2E} \right) \tag{2}$$

where E is the number of edges and d_m is the degree of a node m [37–39]. The modularity measures the strength of a partition with respect to the distribution of edges because it compares the number of edges, $\sum_{m,n} w_{mn}$ with the expected number of edges placed at random, $\sum_{m,n} d_m d_n / 2E$, within the partition. An optimal number of clusters is determined at the maximum point of the modularity.

(**A**)
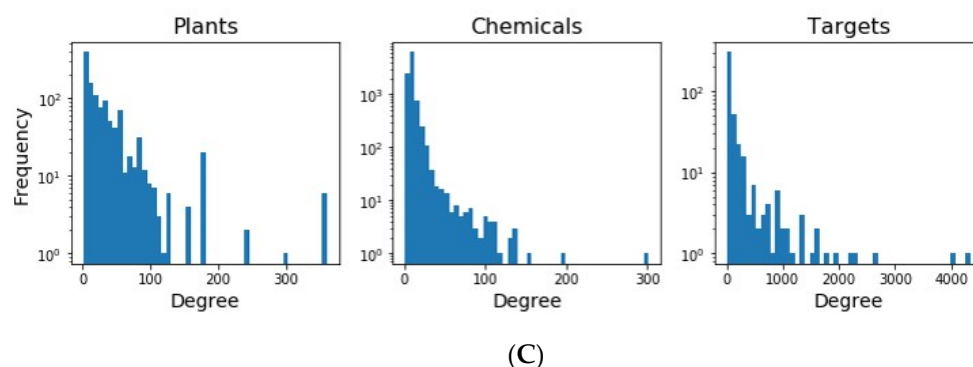

(**B**)

Figure 1. *Cont.*

(**C**)
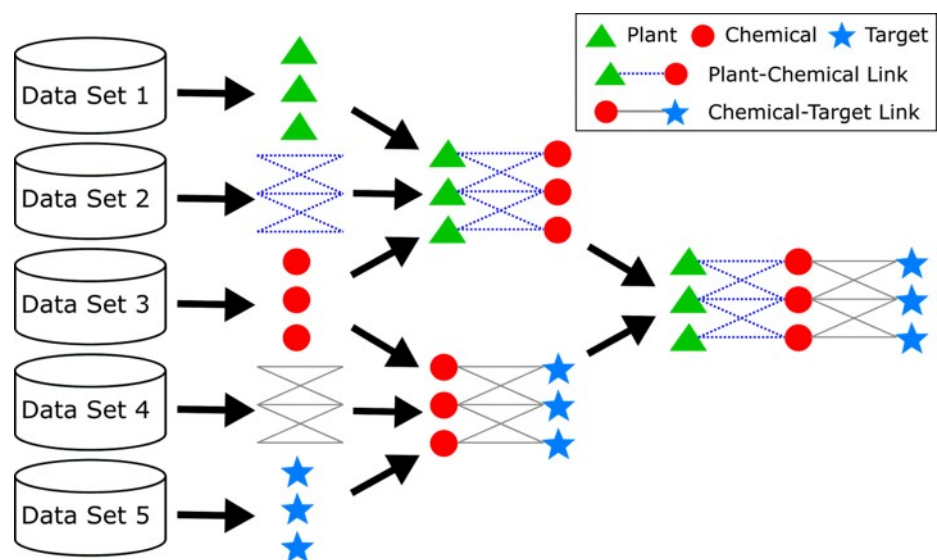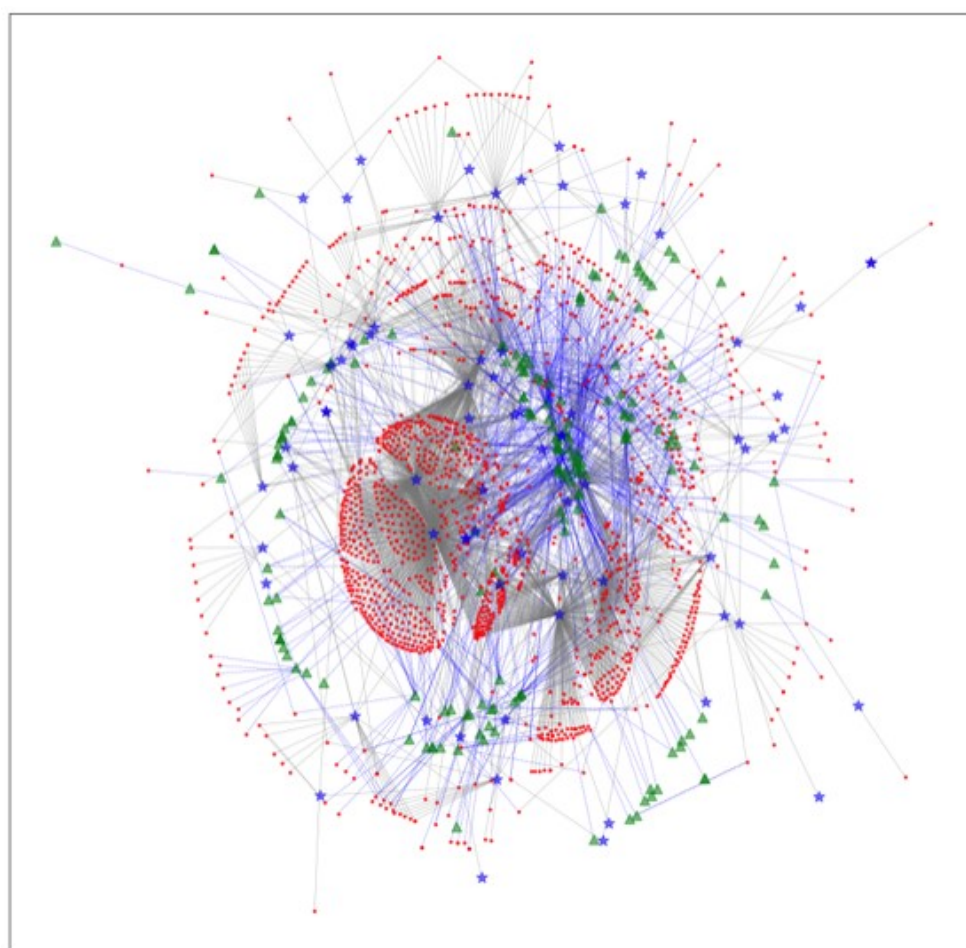
Figure 1. (**A**) The construction and visualization of the multipartite network database. (**B**) The representative multipartite network generated from the nodes which were randomly selected from 10% of the nodes 1138 plants, 10,043 chemicals, and 441 targets in the whole network. (**C**) The distribution of the node degrees for the plants, chemicals, or targets.

Table 1. Basic statistic of the combined multipartite network data.

	Number of Nodes				Number of Edges			Density [1]
	Plant (Np)	Chemical (Nc)	Target (Nt)	Total (N)	Plant-Chemical	Chemical-Target	Total (E)	
Raw Data	5500	10,056	1224	16,780	58,068	100,290	158,358	0.00234
without Non-plants	3614	10,056	1224	14,894	54,585	100,290	154,875	0.00318
without Non-plants or Duplicates	2886	10,056	1224	14,166	34,549	100,290	134,839	0.00326
Only targets with the probability of above 0.9	2886	10,056	441	13,383	34,549	73,112	107,661	0.00322
Pre-processed Data	1138	10,043	441	11,622	34,549	73,112	107,661	0.00679

[1] The density of a multipartite network is computed by $E/(Np \times Nc + Nc \times Nt)$.

The hierarchical clustering analysis led to the identification of the clusters of medicinal plants based on their chemical profiles. The goal of the analysis was to provide insight into the biologically or chemically inter-related plants and search for their potentials medicinal effects in the integrated network. We use the Bray-Curtis distance for the hierarchical clustering, which is defined by, between two plants u and v,

$$\mathrm{d}(u,v) = \frac{\sum_j |u_j - v_j|}{\sum_j |u_j| + |v_j|} \tag{3}$$

where u_j (or v_j) is equal to one if the plant u (or v) is connected to the jth chemical in the network, and zero otherwise. The Bray-Curtis distance takes a value between 0 and 1, and it has two favorable properties for the multipartite network analysis: (i) It is a scale-free distance metric, that is, it depends only on the fraction of the number of chemicals connected to the plants, rather than the total number of chemicals in the network. The scale-freeness is a crucial property for analyzing sparsely connected networks, such as the multipartite network in this study where most of the u_j (or v_j) values are zero; see, for example, Figure S1 in Supplemental Information for a visualization. (ii) It is suitable for the unweighted networks because it only considers the number of edges while ignoring the edge weights.

We used the complete linkage method for hierarchical clustering, that is, the distance between two clusters is defined by the distance between two elements that are farthest from each other. The combined use of the Bray-Curtis distance and complete linkage yields a meaningful interpretation of clustering results. That is, if we combine two clusters one whenever they are at a distance of less than one, then every pair of elements will have a

distance less than one. It implies that every pair of medicinal plants in a cluster will have at least one common chemical because $d(u,v) < 1$ implies that $u_j = v_j = 1$ for some chemical j.

Finally, we proposed an index called TPS, which is a similarity measure between a plant cluster and a target. TPS of a plant cluster P to a target T is defined by the fraction

$$\text{TPS}(P, T) = \frac{|C_P \cap C_T|}{|C_T|} = \frac{\text{number of chemicals connected to both } P \text{ and } T}{\text{number of chemicals connected to } T} \qquad (4)$$

where C_P is the set of all chemicals connected to at least one plant in the plant cluster P, C_T is the set of all chemicals connected to the target T, and $|\cdot|$ denotes the number of elements in a set. TPS can represent potential efficacy of a plant cluster for predicting a protein target. We used the TPS to search for the multipartite network for plant clusters of the highest efficacy or most active targets.

3. Results

3.1. Multipartite Network Analysis

We constructed the multipartite network with the nodes of medicinal plants, chemicals, and targets for the interpretation of the biological roles of these chemicals and plants containing them on multi-targets. The preprocessed network is shown in Figure 1B, where 10% of the nodes are randomly selected for visualization. Note that the plants and targets have connections only to the chemicals. Figure 1C shows the distribution of the node degrees for the plants, chemicals, or targets, where the vertical axes are on a log scale. The distribution of the power-law degree imply that the network is a scale-free network [40–42].

The structural properties of the multipartite network were further analyzed by the network analysis techniques. We divided the network into five subnetworks by applying the spectral co-clustering algorithm. Then, the structural properties of the multipartite network and the five subnetworks were computed and summarized in Table S1 in Supplemental Information. The computed structural properties are a standard set of theoretic graphical measures, such as the diameter, average bipartite clustering coefficient, and degree of assortativity. Figure S3 in Supplemental Information illustrates the multipartite network, modularity, and five subnetworks, where only 10% of the whole nodes were randomly selected for visualization.

3.2. Hierarchical Clustering Analysis of Plants

We performed the hierarchical clustering analysis of the medicinal plants in the multipartite network, and generated dendrograms to visually inspect the clusters of the medicinal plants (Figure 2 and Figure S2 in Supplemental Information). In the two largest clusters of the plants (Figure 2A,C, Tables S2 and S3 in Supplemental Information), the same species with different usage parts have been aggregated at the bottom level (i.e., at the height of zero) in the dendrogram, and different species belonging to the same family are grouped together at low levels. However, some species belonging to the same family are far apart or the ones belonging to a different family are closely located. In the largest cluster, many different species sharing β-sitosterol, which is one of the most common phytosterols and is distributed over a wide range in the plant kingdom, existed [43]; many plants belonging to different families share more chemicals than the ones belonging to the same family (Figure 2A). The chemical diversity of the cluster shows that the specialized distribution of the secondary metabolites in plants is not dependent on their family level. However, in the second largest cluster, the chemical profiles of the plants depending on the family contributed to the classification of the smaller clusters (Figure 2C). Additionally, the chemical diversity of the plants in the second largest cluster (123 types of subclasses and 311 types of direct parent levels) was bigger than the largest cluster (109 types of subclasses and 257 types of direct parent levels) (Figure 2B,D). Among the smaller clusters, two clusters containing the specialized types of chemicals, protoberberine alkaloids (Figure 2E,F, and Table S4 in Supplemental Information) and vitamins (Figure 2G,H, Table S5 in Supplemental Information), were clearly categorized by the family. We expect that the clustering

information by the chemical profiles between the plants can be utilized not only for the chemotaxonomic study of the medicinal plants, but also for alternative use of the medicinal plants.
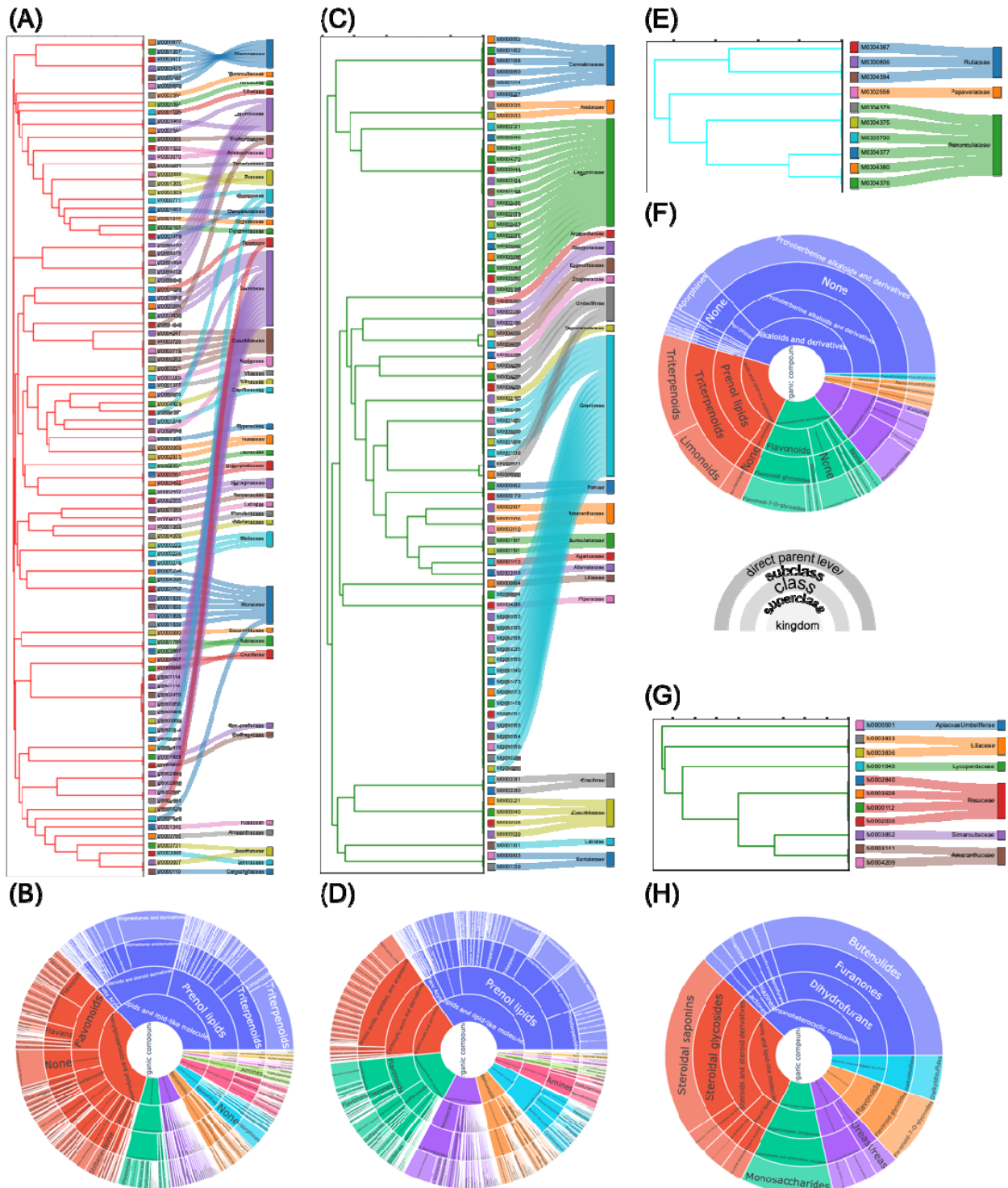

Figure 2. The four sub-clusters in the whole dendrogram (Figure S2). (**A**) The largest sub-cluster in the whole network. (**B**) The classification of chemicals in (**A**). (**C**) The second largest sub-cluster in the whole network. (**D**) The classification of chemicals in (**C**). (**E**) The sub-cluster of the nodes sharing protoberberine-type chemicals in the whole network. (**F**) The classification of chemicals in (**E**). (**G**) The sub-cluster of the nodes sharing vitamins in the whole network. (**H**) The classification of chemicals in (**G**).

### 3.3. Interaction between Plants and Targets Based on TPS

In the multipartite network, the interaction between the plants and targets is mediated by the connectivity of the chemicals in the network. We used the TPS to quantitatively measure the interaction that represents the potential efficacy of the plants for the targets. Because the plants in the same cluster share chemicals in common, the TPS was calculated between a cluster of plants and target.

The TPS was applied in two ways. First, the targets with the highest TPS values were selected for each of the four clusters of plants that were mentioned in the previous section and were manually investigated to validate our approaches. The subnetworks of the selected targets, chemicals connected to the targets, and corresponding cluster of plants connected to the chemicals are illustrated in Figure 3A–D, for the inspection of the effects of the mixture of chemicals or the plants in the cluster. Specifically speaking, the targets with the TPS values greater than 0.05~0.5 were selected while the number of chemicals connected to the targets was kept sufficiently large enough (at least 20) to produce reliable TPS values.
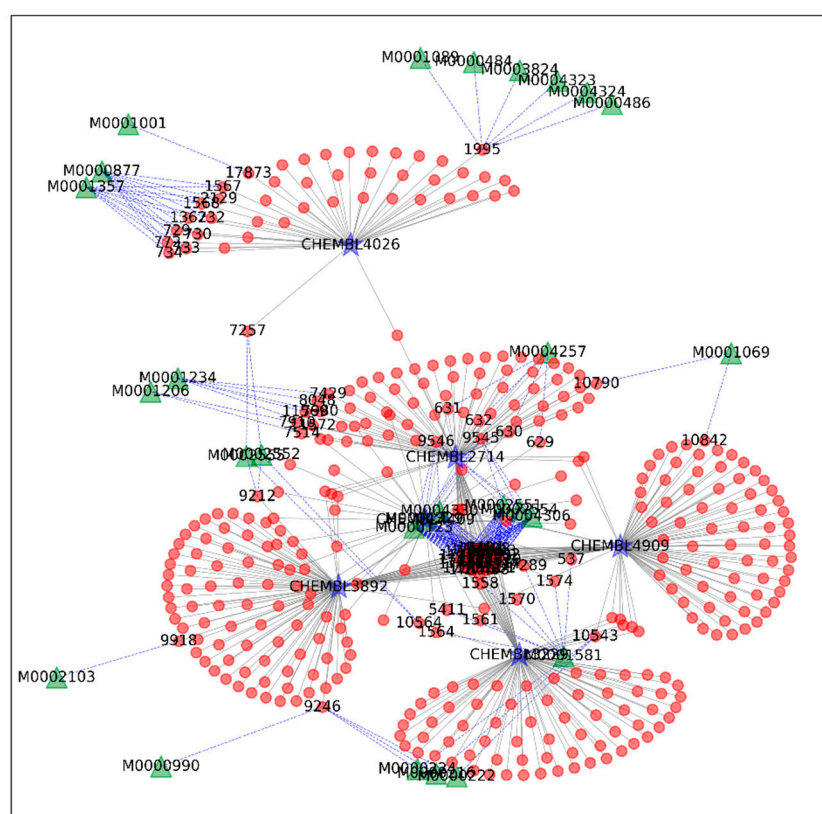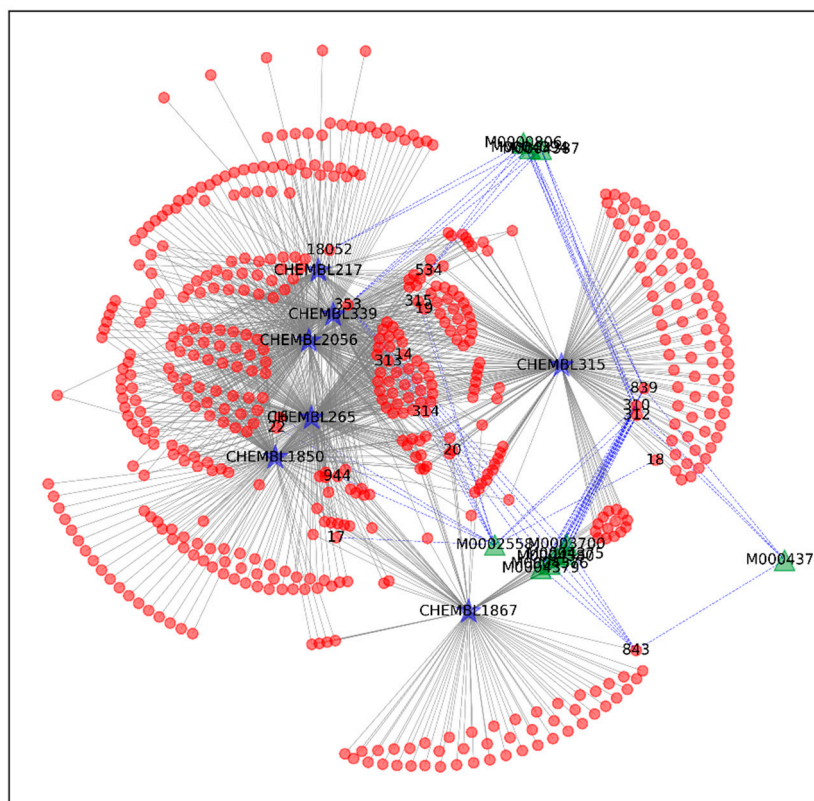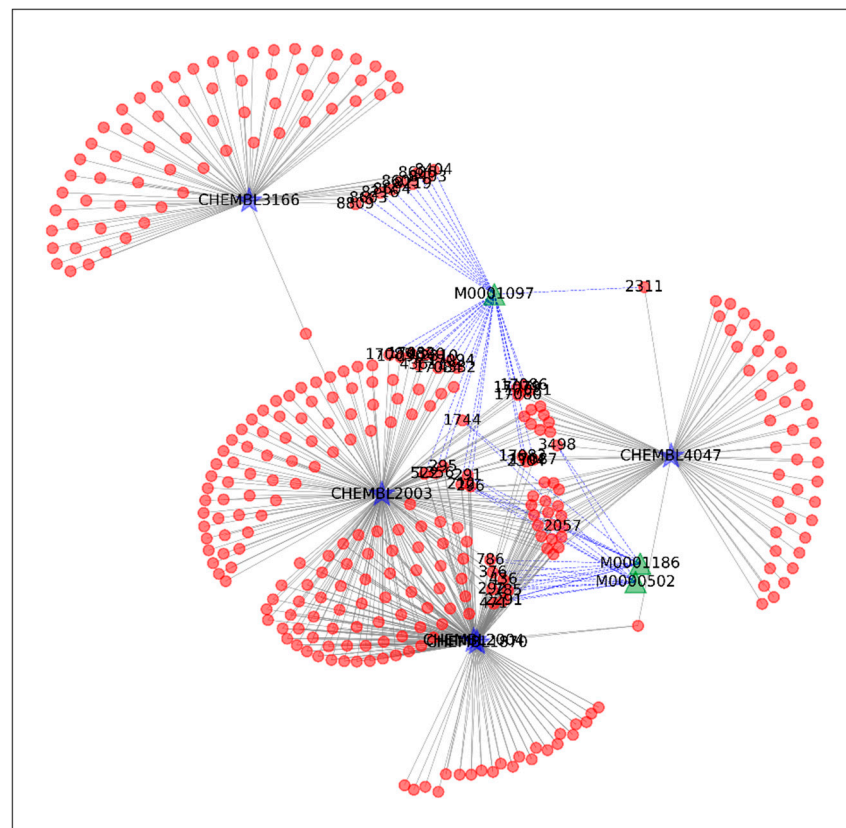

(**A**)

Figure 3. *Cont.*
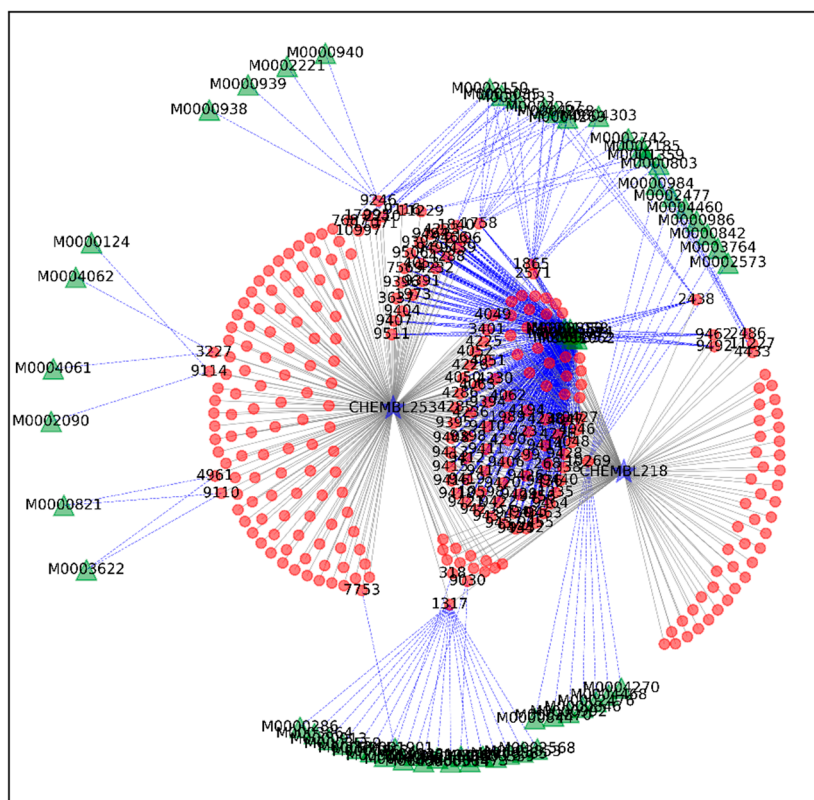
(**B**)

(**C**)

Figure 3. *Cont.*

(**D**)

Figure 3. (**A**) The network diagram for the plants cluster in Figure 2A and the 6 targets retrieved based on the TPS index greater than or equal to 0.1. The chemicals effective to the targets are depicted in red circles. (**B**) The network diagram for the plants cluster in Figure 2C and the 2 targets retrieved based on the TPS index greater than or equal to 0.5. The chemicals effective to the targets are depicted in red circles. (**C**) The network diagram for the plants cluster in Figure 2E and the 7 targets retrieved based on the TPS index greater than or equal to 0.05. The chemicals effective to the targets are depicted in red circles. (**D**) The network diagram for the plants cluster in Figure 2G and the 5 targets retrieved based on the TPS index greater than or equal to 0.05. The chemicals effective to the targets are depicted in red circles.

In each network, a single target had multiple edges with chemicals, but only a few chemicals contributed to the connection between the plants and targets. In the largest cluster, six targets, signal transducer and activator of transcription 3, phosphodiesterase 10A, anandamide amidohydrolase, prostanoid EP2 receptor, sphingosine 1-phosphate receptor Edg-3, and pepsin A, are displayed (Figure 3A). For example, in case of phosphodiesterase 10A, which has been studied as a therapeutic target for colon cancer and neurological symptoms [44], 12 chemicals, mainly carboline-type alkaloids, are displayed with linkages to the target (Table S6 in Supplemental Information). The well-known targets, such as cannabinoid CB1 and CB2 receptors, in the network from the second largest cluster (Figures 2C and 3B) shared the monoterpenoids and benzopyran-type chemicals including cannabinoids from the genus Cannabis [45]. In the third network, protoberberine-type alkaloids from three families, Papaveraceae, Ranunculaceae, and Rutaceae, were linked with the targets, dopamine receptors, and alpha adrenergic receptors, which are known as the main target of these chemicals [46,47]. Tetraterpenoids, such as lycopene, lutein, and β-carotene, which are the well-known chemicals targeting the retinoid signaling pathway, had multiple linkages with retinoic acid receptor gamma (RAR-γ) and retinoid X receptor beta (RXR-β) [48–50]. These results show that the target information predicted by the

chemicals in the multipartite network is reliable enough to investigate alternative targets of the chemicals, for which biological roles have not been defined in previous studies or they existed as mixed forms with other chemicals in a plant.

Second, the proposed TPS was applied for searching the closest related plants, which are predictive of a given target based on the highest TPS value. The results provide clues to understand the complex biological roles of the combination of medicinal plants or the chemicals produced from the plants. In this scenario, a cluster of plants is recommended from among the large-scale network as the candidate medicinal plant for a single target of interest. For example, we used "tyrosinase" as the model target to validate the suggested method and to investigate the medicinal effects of the retrieved cluster of plants (Figure 4A). In the multipartite network retrieved for the target "tyrosinase," 165 chemicals from 31 different species were linked with the target node. Among them, 50 chemicals from Morus alba were found in the middle of the network (Table S7 in Supplemental Information). Additionally, *Carthamus tinctorius*, *Cudrania tricuspidata*, *Glycyrrhiza inflata*, and *Glycyrrhiza uralensis* had more than 10 chemicals, which were predicted to act on tyrosinases (Table S8 in Supplemental Information). These five plants have been known as the most popular and successful sources for the whitening products used in the cosmetic industry [51,52].
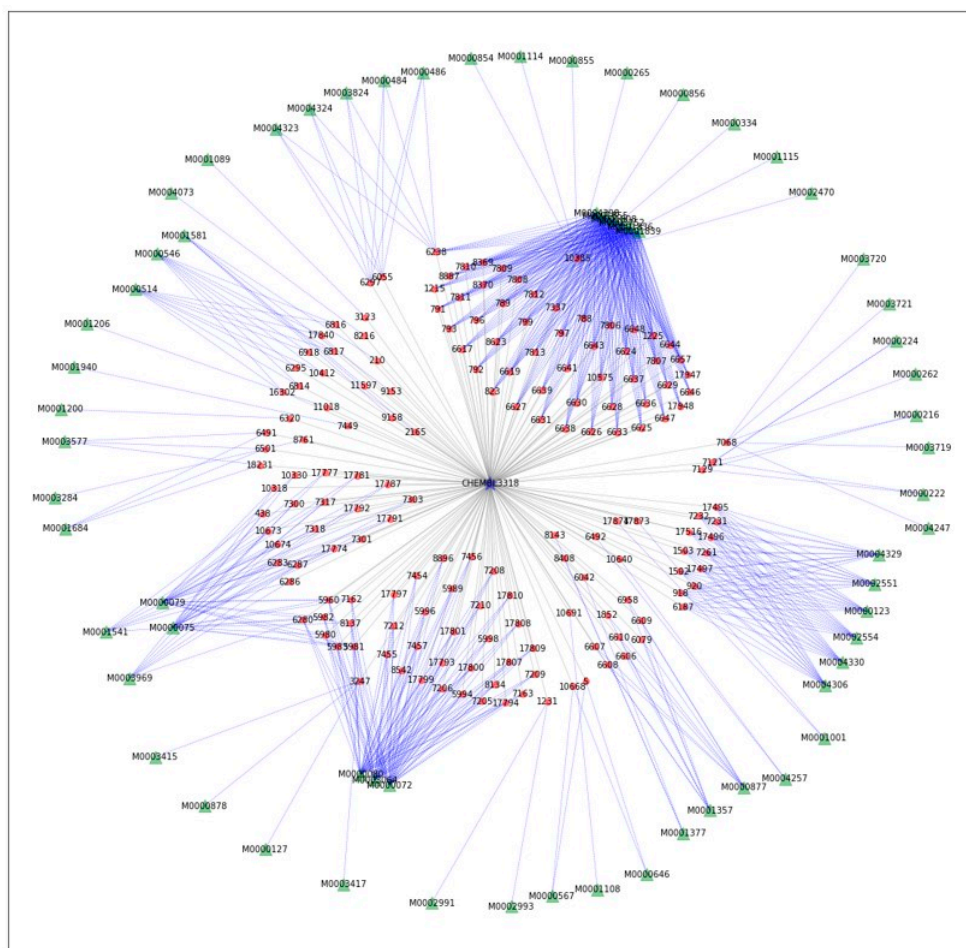
(A)

Figure 4. *Cont.*
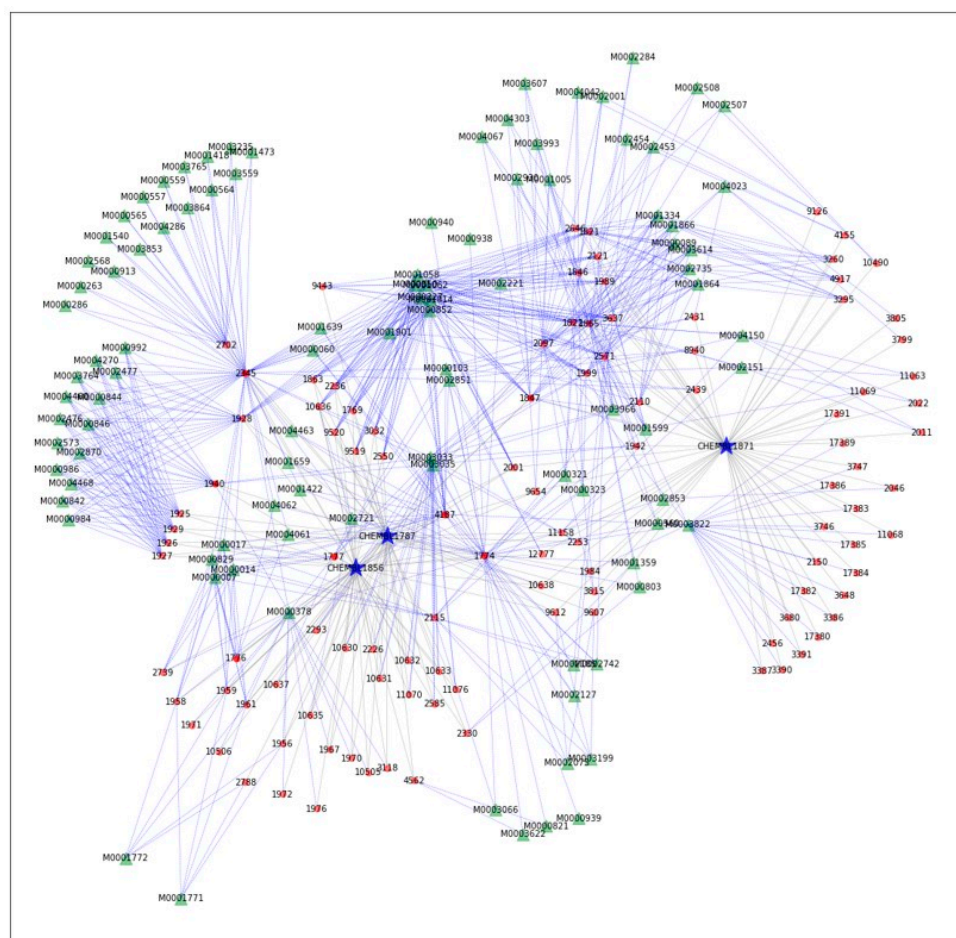
(B)

Figure 4. (**A**) The network diagram for the plants cluster having the largest TPS index to the target "Tyrosinase" (CHEMBL3318). (**B**) The combined network generated from three network diagrams for the plants cluster having the largest TPS index to the three targets, "steroid 5α-reductase 1" (S5R1, ChEMBL1787), "steroid 5α-reductase 2" (S5R2, ChEMBL1856) and "androgen receptor" (AR, ChEMBL1871). The chemical profiles sharing from chemicals from S5R1 and S5R2, from S5R1 and AR and from S5R2 and AR.

Next, we combined three subnetworks-networks from three targets, steroid 5α-reductase 1 (S5R1), steroid 5α-reductase 2 (S5R2), and androgen receptor (AR), which are involved in the androgen metabolism, for the interaction of plants and chemicals on multi-targets (Figure 4B). In total, 44 among 52 and 54 chemicals from S5R1 and S5R2, respectively, shared with each other, but 10 of 52 from S5R1 and 7 of 54 from S5R2 only shared those with 58 chemicals from AR (Tables S9–S11 in Supplemental Information). The types of chemicals between the pairs of targets showed significantly different patterns. The chemicals co-targeting S5R1 and S5R2 were mainly the steroid derivatives, such as cholestanes, ergostanes, and stigmastanes, which were synthesized via the triterpenoid pathways (Figure 4); however, sesquiterpenoids and diterpenoids were the main components sharing the pairs of S5R1 and AR, or S5R2 and AR (Figure 4). We found that only 6 chemicals simultaneously acted on three targets and plants containing them (Table 2).

Table 2. List of chemicals simultaneously acting on three targets, S5R1, S5R2 and AR.

Chemical Name	CAS Number	Chemicals Classification				
		Kingdom	Superclass	Class	Subclass	Direct Parent
Kusunol	20489-45-6	Organic compounds	Lipids and lipid-like molecules	Prenol lipids	Sesquiterpenoids	Eremophilane, 8,9-secoeremophilane and furoeremophilane sesquiterpenoids
Stigmasta-4,6-dien-3-one	29374-98-9	Organic compounds	Lipids and lipid-like molecules	Steroids and steroid derivatives	Stigmastanes and derivatives	Stigmastanes and derivatives
Stigmastan-3-one (5-alpha)	102734-69-0	Organic compounds	Lipids and lipid-like molecules	Steroids and steroid derivatives	Stigmastanes and derivatives	Stigmastanes and derivatives
Dehydroabietinal	13601-88-2	Organic compounds	Lipids and lipid-like molecules	Prenol lipids	Diterpenoids	Diterpenoids
(-)-alpha-Copaene	3856-25-5	Organic compounds	Lipids and lipid-like molecules	Prenol lipids	Sesquiterpenoids	Sesquiterpenoids
(-)-Ylangene	14912-44-8	Organic compounds	Lipids and lipid-like molecules	Prenol lipids	Sesquiterpenoids	Sesquiterpenoids

4. Discussion

In this study, we tried to unravel the complex interaction between multi-chemicals and multiple targets using the computational target prediction model and multipartite networks comprising the relationship between the plants, chemicals, and targets. In hierarchical cluster analysis, plants producing similar specialized metabolites are clustered only within a taxon. In particular, in chemotaxonomy, secondary metabolites can be used as specialized markers to distinguish among plants [53]. However, not all secondary metabolites are produced in the same taxon. For example, flavonoids are widely distributed in plants and even in fungi, fulfilling many functions, such as ultraviolet filtration, symbiotic nitrogen fixation, and floral pigmentation [54]. The chemical profile of each plant has a minor effect on the classification of the plant.

To unravel the clusters of multipartite networks from hierarchical clustering analysis, we introduced an index of potential efficacy of plant clusters toward a target, called TPS. TPS indicated the proximity of subnetworks of the metabolic relationships between plant clusters and targets. We utilized it to search for and suggest candidate plants and chemicals acting on specific targets (Figure 3). It was validated that the plants with more than 10 chemicals had enough experimental evidence to explain the biological roles on specific targets retrieved in the network. TPS was also used for the extraction of a list of plants and chemicals acting on a specific target of interest (Figure 4).

We also obtained potent plant and chemical candidates for the specific target, "tyrosinase", confirming the validity of the approach. The node information of the plants in the network is expected to guide toward new alternative sources developed by plant combinations.

The approach is based on the multipartite network of plants and chemicals for a single target using the TPS and could be expanded for multi-targets that closely interact with each other. The plants and chemicals, which are expected to act on multiple targets, including S5R1, S5R2, and AR, could be suggested to control the levels of testosterone (T) and dihydrotestosterone (DHT), involved in the androgen metabolism. It is known that the blockade of AR can decrease benign prostate hyperplasia (BPH) symptoms and further relieve the lower urinary tract symptoms [55]. Since the alterations in AR are associated with prostate cancer, the inhibition of AR by drugs is the prevention target for prostate cancer growth [56]. The 5α-reductase inhibitors, which suppress the conversion of T into DHT, contribute to the reduction in BPH or preventing prostate cancer [55,57]. In the present study, the AR-targeted chemicals were sesquiterpenoid and diterpenoid-type, while the steroid derivatives were suggested as the chemicals targeting 5α-reductases

(S5R1 and S5R2). We expected that the plants containing the chemicals acting on all three targets, S5R1, S5R2, and AR, can be the new candidates for controlling the symptoms or diseases involving androgen metabolism, such as BPH and prostate cancer. The well-organized multipartite network is expected to explain the synergistic effects between the multi-chemicals in single or multiple medicinal plants.

We have constructed a multipartite network by integrating the databases of the medicinal plants, chemicals, and targets for the prediction of the biological roles of the plants and chemicals on specific targets. We have presented a novel computational approach based on the multipartite network with which the plant species can be distinguished based on their chemical profiles. In the clustered bipartite network of plants and chemicals, the chemical profiles of the plants in identical or similar species were clustered closely at the bottom level of the dendrogram. However, the plants in the same family were found far apart in the dendrogram in many cases, which implies that the chemical profiles of the plants do not significantly contribute to the classification of the species.

5. Conclusions

We have demonstrated that the network analysis helps to understand the biological roles of multi-chemicals on multi-targets. With the target prediction models, it also provides a way to discover alternative therapeutic agents from natural sources by extrapolating from the known bioactivities of chemicals on targets. The predicted results about the biological roles of multi-chemicals on multi-targets allow to explain the therapeutic activity of the medicinal plant extracts via subnetworks of molecular mechanisms and suggest the combinatory preparation of several medicinal plants on multi-targets to be involved in the symptoms or diseases of interest. In our future study, we will study the candidates, especially the combinations of plants or chemicals, predicted from the multipartite network in in vitro and in vivo models to validate our approaches and further develop novel therapeutic botanical drugs.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/biom11040546/s1, Figure S1: Visualization of non-zero values of the data matrix for the plants and chemicals, where the rows correspond to the plants and the columns correspond to the chemicals, Figure S2: The whole dendrogram for clustering of plants obtained by the hierarchical clustering analysis, Figure S3: Visualization of the results of the spectral co-clustering. (A) Spectral co-clustering of the whole network. (B) The modularity values for the determination of the number of clusters for the spectral co-clustering algorithm. (C) The subnetworks induced by each of the five subsets of nodes obtained by the spectral co-clustering. The 10% of the whole nodes were randomly selected for visualization, Table S1: The structural properties of the whole network and the five subnetworks obtained by the spectral co-clustering algorithm, Table S2: List of plants illustrated in Figure 2A,B, Table S3: List of plants illustrated in Figure 2C,D, Table S4: List of plants illustrated in Figure 2E,F, Table S5: List of plants illustrated in Figure 2G,H, Table S6: List of chemicals which are targeted on phosphodiesterase 10A in Figure 3A, Table S7: List of chemicals linked with the plant, *Morus alba* and the target, tyrosinase, Table S8: List of chemicals linked with the plants, *Carthamus tinctorius*, *Cudrania tricuspidata*, *Glycyrrhiza inflata* and *Glycyrrhiza uralensis* and the target, tyrosinase, Table S9: List of chemicals which are co-targeted on S5R1 and S5R2, Table S10: List of chemicals which are co-targeted on S5R1 and AR, Table S11. List of chemicals which are co-targeted on S5R2 and AR.

Author Contributions: Conceptualization, H.Y. (Heejung Yang); methodology, N.L. and H.Y. (Heejung Yang); validation, N.L., H.Y. (Hojin Yoo) and H.Y. (Heejung Yang); data curation, N.L., H.Y. (Hojin Yoo) and H.Y. (Heejung Yang); writing—original draft preparation, N.L. and H.Y. (Hojin Yoo); writing—review and editing, H.Y. (Heejung Yang); visualization, N.L.; supervision, H.Y. (Heejung Yang); project administration, H.Y. (Heejung Yang); funding acquisition, H.Y. (Heejung Yang) All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Petrovska, B. Historical Review of Medicinal Plants' Usage. *Pharmacogn. Rev.* **2012**, *6*, 1. [CrossRef]
2. Porras, G.; Chassagne, F.; Lyles, J.T.; Marquez, L.; Dettweiler, M.; Salam, A.M.; Samarakoon, T.; Shabih, S.; Farrokhi, D.R.; Quave, C.L. Ethnobotany and the Role of Plant Natural Products in Antibiotic Drug Discovery. *Chem. Rev.* **2021**, *121*, 3495–3560. [CrossRef]
3. Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The Re-Emergence of Natural Products for Drug Discovery in the Genomics Era. *Nat. Rev. Drug Discov.* **2015**, *14*, 111–129. [CrossRef] [PubMed]
4. Koehn, F.E.; Carter, G.T. The Evolving Role of Natural Products in Drug Discovery. *Nat. Rev. Drug Discov.* **2005**, *4*, 206–220. [CrossRef]
5. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on Natural Products for Drug Design. *Nat. Chem.* **2016**, *8*, 531–541. [CrossRef] [PubMed]
6. Galanakis, C.M. *Nutraceuticals and Natural Product Pharmaceuticals*; Academic Press: Cambridge, MA, USA, 2019; ISBN 978-0-12-816450-1.
7. Anighoro, A.; Bajorath, J.; Rastelli, G. Polypharmacology: Challenges and Opportunities in Drug Discovery: Miniperspective. *J. Med. Chem.* **2014**, *57*, 7874–7887. [CrossRef]
8. Besnard, J.; Ruda, G.F.; Setola, V.; Abecassis, K.; Rodriguiz, R.M.; Huang, X.-P.; Norval, S.; Sassano, M.F.; Shin, A.I.; Webster, L.A.; et al. Automated Design of Ligands to Polypharmacological Profiles. *Nature* **2012**, *492*, 215–220. [CrossRef]
9. Keiser, M.J.; Setola, V.; Irwin, J.J.; Laggner, C.; Abbas, A.I.; Hufeisen, S.J.; Jensen, N.H.; Kuijer, M.B.; Matos, R.C.; Tran, T.B.; et al. Predicting New Molecular Targets for Known Drugs. *Nature* **2009**, *462*, 175–181. [CrossRef]
10. Lounkine, E.; Keiser, M.J.; Whitebread, S.; Mikhailov, D.; Hamon, J.; Jenkins, J.L.; Lavan, P.; Weber, E.; Doak, A.K.; Côté, S.; et al. Large-Scale Prediction and Testing of Drug Activity on Side-Effect Targets. *Nature* **2012**, *486*, 361–367. [CrossRef]
11. Ramsay, R.R.; Popovic-Nikolic, M.R.; Nikolic, K.; Uliassi, E.; Bolognesi, M.L. A Perspective on Multi-Target Drug Discovery and Design for Complex Diseases. *Clin. Trans. Med.* **2018**, *7*, 3. [CrossRef]
12. Roth, B.L.; Sheffler, D.J.; Kroeze, W.K. Magic Shotguns versus Magic Bullets: Selectively Non-Selective Drugs for Mood Disorders and Schizophrenia. *Nat. Rev. Drug Discov.* **2004**, *3*, 353–359. [CrossRef] [PubMed]
13. Zhang, C.; Hong, H.; Mendrick, D.L.; Tang, Y.; Cheng, F. Biomarker-Based Drug Safety Assessment in the Age of Systems Pharmacology: From Foundational to Regulatory Science. *Biomark. Med.* **2015**, *9*, 1241–1252. [CrossRef] [PubMed]
14. Zhang, Y.; Wang, X.; Li, X.; Peng, S.; Wang, S.; Huang, C.Z.; Huang, C.Z.; Zhang, Q.; Li, D.; Jiang, J.; et al. Identification of a Specific Agonist of Human TAS2R14 from Radix Bupleuri through Virtual Screening, Functional Evaluation and Binding Studies. *Sci. Rep.* **2017**, *7*, 12174. [CrossRef] [PubMed]
15. Bansal, M.; Yang, J.; Karan, C.; Menden, M.P.; Costello, J.C.; Tang, H.; Xiao, G.; Li, Y.; Allen, J.; Zhong, R.; et al. A Community Computational Challenge to Predict the Activity of Pairs of Compounds. *Nat. Biotechnol.* **2014**, *32*, 1213–1222. [CrossRef]
16. Han, K.; Jeng, E.E.; Hess, G.T.; Morgens, D.W.; Li, A.; Bassik, M.C. Synergistic Drug Combinations for Cancer Identified in a CRISPR Screen for Pairwise Genetic Interactions. *Nat. Biotechnol.* **2017**, *35*, 463–474. [CrossRef] [PubMed]
17. Jia, J.; Zhu, F.; Ma, X.; Cao, Z.W.; Li, Y.X.; Chen, Y.Z. Mechanisms of Drug Combinations: Interaction and Network Perspectives. *Nat. Rev. Drug Discov.* **2009**, *8*, 111–128. [CrossRef]
18. Sun, Y.; Sheng, Z.; Ma, C.; Tang, K.; Zhu, R.; Wu, Z.; Shen, R.; Feng, J.; Wu, D.; Huang, D.; et al. Combining Genomic and Network Characteristics for Extended Capability in Predicting Synergistic Drugs for Cancer. *Nat. Commun.* **2015**, *6*, 8481. [CrossRef]
19. Ahn, K. The Worldwide Trend of Using Botanical Drugs and Strategies for Developing Global Drugs. *BMB Rep.* **2017**, *50*, 111–116. [CrossRef]
20. Ezzat, A.; Wu, M.; Li, X.; Kwoh, C.-K. Computational Prediction of Drug-Target Interactions via Ensemble Learning. In *Computational Methods for Drug Repurposing*; Vanhaelen, Q., Ed.; Springer: New York, NY, USA, 2019; Volume 1903, pp. 239–254, ISBN 978-1-4939-8954-6.
21. Forouzesh, A.; Samadi Foroushani, S.; Forouzesh, F.; Zand, E. Reliable Target Prediction of Bioactive Molecules Based on Chemical Similarity Without Employing Statistical Methods. *Front. Pharmacol.* **2019**, *10*, 835. [CrossRef]
22. Lavecchia, A.; Cerchia, C. In Silico Methods to Address Polypharmacology: Current Status, Applications and Future Perspectives. *Drug Discov. Today* **2016**, *21*, 288–298. [CrossRef]
23. Luo, Y.; Zhao, X.; Zhou, J.; Yang, J.; Zhang, Y.; Kuang, W.; Peng, J.; Chen, L.; Zeng, J. A Network Integration Approach for Drug-Target Interaction Prediction and Computational Drug Repositioning from Heterogeneous Information. *Nat. Commun.* **2017**, *8*, 573. [CrossRef]
24. Hopkins, A.L. Network Pharmacology: The next Paradigm in Drug Discovery. *Nat. Chem. Biol.* **2008**, *4*, 682–690. [CrossRef] [PubMed]
25. Monteiro, N.R.C.; Ribeiro, B.; Arrais, J. Drug-Target Interaction Prediction: End-to-End Deep Learning Approach. *IEEE/ACM Trans. Comput Biol. Bioinform.* **2020**. [CrossRef]

26.   Cheng, F.; Kovács, I.A.; Barabási, A.-L. Network-Based Prediction of Drug Combinations. *Nat. Commun.* **2019**, *10*, 1197. [CrossRef]
27.   Leung, E.L.; Cao, Z.-W.; Jiang, Z.-H.; Zhou, H.; Liu, L. Network-Based Drug Discovery by Integrating Systems Biology and Computational Technologies. *Brief. Bioinform.* **2013**, *14*, 491–505. [CrossRef] [PubMed]
28.   Yıldırım, M.A.; Goh, K.-I.; Cusick, M.E.; Barabási, A.-L.; Vidal, M. Drug—Target Network. *Nat. Biotechnol.* **2007**, *25*, 1119–1126. [CrossRef] [PubMed]
29.   Zhao, J.; Jiang, P.; Zhang, W. Molecular Networks for the Study of TCM Pharmacology. *Brief. Bioinform.* **2010**, *11*, 417–430. [CrossRef] [PubMed]
30.   Bosc, N.; Atkinson, F.; Felix, E.; Gaulton, A.; Hersey, A.; Leach, A.R. Large Scale Comparison of QSAR and Conformal Prediction Methods and Their Applications in Drug Discovery. *J. Cheminform.* **2019**, *11*. [CrossRef]
31.   Davies, M.; Nowotka, M.; Papadatos, G.; Dedman, N.; Gaulton, A.; Atkinson, F.; Bellis, L.; Overington, J.P. ChEMBL Web Services: Streamlining Access to Drug Discovery Data and Utilities. *Nucleic Acids Res.* **2015**, *43*, W612–W620. [CrossRef]
32.   Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A.P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L.J.; Cibrián-Uhalte, E.; et al. The ChEMBL Database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945–D954. [CrossRef]
33.   Koutsoukas, A.; Simms, B.; Kirchmair, J.; Bond, P.J.; Whitmore, A.V.; Zimmer, S.; Young, M.P.; Jenkins, J.L.; Glick, M.; Glen, R.C.; et al. From in Silico Target Prediction to Multi-Target Drug Design: Current Databases, Methods and Applications. *J. Proteom.* **2011**, *74*, 2554–2574. [CrossRef]
34.   Shin, J.-S.; Lee, Y.-S.; Lee, M.-S. Protection and Utilization of Traditional Knowledge Resources through Korean Traditional Knowledge Portal(KTKP). *J. Korea Contents Assoc.* **2010**, *10*, 422–426. [CrossRef]
35.   Dhillon, I.S. Co-Clustering Documents and Words Using Bipartite Spectral Graph Partitioning. In Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD '01, San Francisco, CA, USA, 26–29 August 2001; ACM Press: San Francisco, CA, USA, 2001; pp. 269–274.
36.   Martínez-Jiménez, F.; Papadatos, G.; Yang, L.; Wallace, I.M.; Kumar, V.; Pieper, U.; Sali, A.; Brown, J.R.; Overington, J.P.; Marti-Renom, M.A. Target Prediction for an Open Access Set of Compounds Active against Mycobacterium Tuberculosis. *PLoS Comput. Biol.* **2013**, *9*, e1003253. [CrossRef] [PubMed]
37.   Newman, M.E.J. Modularity and Community Structure in Networks. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8577–8582. [CrossRef]
38.   Newman, M.E.J. Fast Algorithm for Detecting Community Structure in Networks. *Phys. Rev. E* **2004**, *69*, 066133. [CrossRef]
39.   Newman, M.E.J.; Girvan, M. Finding and Evaluating Community Structure in Networks. *Phys. Rev. E* **2004**, *69*, 026113. [CrossRef] [PubMed]
40.   Goh, K.-I.; Oh, E.; Jeong, H.; Kahng, B.; Kim, D. Classification of Scale-Free Networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12583–12588. [CrossRef]
41.   Lima-Mendez, G.; van Helden, J. The Powerful Law of the Power Law and Other Myths in Network Biology. *Mol. BioSyst.* **2009**, *5*, 1482. [CrossRef] [PubMed]
42.   Newman, M. Power Laws, Pareto Distributions and Zipf's Law. *Contemp. Phys.* **2005**, *46*, 323–351. [CrossRef]
43.   Tao, C.; Shkumatov, A.A.; Alexander, S.T.; Ason, B.L.; Zhou, M. Stigmasterol Accumulation Causes Cardiac Injury and Promotes Mortality. *Commun. Biol.* **2019**, *2*, 20. [CrossRef] [PubMed]
44.   Li, N.; Lee, K.; Xi, Y.; Zhu, B.; Gary, B.D.; Ramírez-Alcántara, V.; Gurpinar, E.; Canzoneri, J.C.; Fajardo, A.; Sigler, S.; et al. Phosphodiesterase 10A: A Novel Target for Selective Inhibition of Colon Tumor Cell Growth and β-Catenin-Dependent TCF Transcriptional Activity. *Oncogene* **2015**, *34*, 1499–1509. [CrossRef]
45.   Wu, J. Cannabis, Cannabinoid Receptors, and Endocannabinoid System: Yesterday, Today, and Tomorrow. *Acta Pharmacol. Sin.* **2019**, *40*, 297–299. [CrossRef] [PubMed]
46.   Fan, J.; Zhang, K.; Jin, Y.; Li, B.; Gao, S.; Zhu, J.; Cui, R. Pharmacological Effects of Berberine on Mood Disorders. *J. Cell. Mol. Med.* **2019**, *23*, 21–28. [CrossRef] [PubMed]
47.   Kawano, M.; Takagi, R.; Kaneko, A.; Matsushita, S. Berberine Is a Dopamine D1- and D2-like Receptor Antagonist and Ameliorates Experimentally Induced Colitis by Suppressing Innate and Adaptive Immune Responses. *J. Neuroimmunol.* **2015**, *289*, 43–55. [CrossRef] [PubMed]
48.   Aydemir, G.; Kasiri, Y.; Birta, E.; Béke, G.; Garcia, A.L.; Bartók, E.-M.; Rühl, R. Lycopene-Derived Bioactive Retinoic Acid Receptors/Retinoid-X Receptors-Activating Metabolites May Be Relevant for Lycopene's Anti-Cancer Potential. *Mol. Nutr. Food Res.* **2013**, *57*, 739–747. [CrossRef] [PubMed]
49.   Aydemir, G.; Carlsen, H.; Blomhoff, R.; Rühl, R. Lycopene Induces Retinoic Acid Receptor Transcriptional Activation in Mice. *Mol. Nutr. Food Res.* **2012**, *56*, 702–712. [CrossRef] [PubMed]
50.   Huang, X.; Gao, Y.; Zhi, X.; Ta, N.; Jiang, H.; Zheng, J. Association between Vitamin A, Retinol and Carotenoid Intake and Pancreatic Cancer Risk: Evidence from Epidemiologic Studies. *Sci. Rep.* **2016**, *6*, 38936. [CrossRef]
51.   Smit, N.; Vicanova, J.; Pavel, S. The Hunt for Natural Skin Whitening Agents. *Int. J. Mol. Sci.* **2009**, *10*, 5326–5349. [CrossRef]
52.   Xin, L.-T.; Yue, S.-J.; Fan, Y.-C.; Wu, J.-S.; Yan, D.; Guan, H.-S.; Wang, C.-Y. Cudrania Tricuspidata: An Updated Review on Ethnomedicine, Phytochemistry and Pharmacology. *RSC Adv.* **2017**, *7*, 31807–31832. [CrossRef]
53.   Wink, M. Evolution of Secondary Metabolites from an Ecological and Molecular Phylogenetic Perspective. *Phytochemistry* **2003**, *64*, 3–19. [CrossRef]
54.   Kumar, S.; Pandey, A.K. Chemistry and Biological Activities of Flavonoids: An Overview. *Sci. World J.* **2013**, *2013*, 1–16. [CrossRef] [PubMed]

55. Izumi, K.; Mizokami, A.; Lin, W.-J.; Lai, K.-P.; Chang, C. Androgen Receptor Roles in the Development of Benign Prostate Hyperplasia. *Am. J. Pathol.* **2013**, *182*, 1942–1949. [CrossRef]
56. Tan, M.E.; Li, J.; Xu, H.E.; Melcher, K.; Yong, E. Androgen Receptor: Structure, Role in Prostate Cancer and Drug Discovery. *Acta Pharmacol. Sin.* **2015**, *36*, 3–23. [CrossRef] [PubMed]
57. Liss, M.A.; Thompson, I.M. Prostate Cancer Prevention with 5-Alpha Reductase Inhibitors: Concepts and Controversies. *Curr. Opin. Urol.* **2018**, *28*, 42–45. [CrossRef] [PubMed]