Hindawi Publishing Corporation
Evidence-Based Complementary and Alternative Medicine
Volume 2016, Article ID 2106465, 25 pages
http://dx.doi.org/10.1155/2016/2106465

Research Article

Global Mapping of Traditional Chinese Medicine into Bioactivity Space and Pathways Annotation Improves Mechanistic Understanding and Discovers Relationships between Therapeutic Action (Sub)classes

Siti Zuraidah Mohamad Zobir,[1,2] Fazlin Mohd Fauzi,[3] Sonia Liggi,[4] Georgios Drakakis,[5] Xianjun Fu,[6] Tai-Ping Fan,[7] and Andreas Bender[1]

[1]*Centre for Molecular Science Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK*
[2]*Malaysian Institute of Pharmaceuticals and Nutraceuticals (IPharm), Ministry of Science, Technology and Innovation, 11800 Penang, Malaysia*
[3]*Department of Pharmacology and Chemistry, Faculty of Pharmacy, Universiti Teknologi MARA, Puncak Alam Campus, 42300 Bandar Puncak Alam, Selangor, Malaysia*
[4]*Department of Biomedical Sciences, University of Cagliari, Cittadella Universitaria di Monserrato, SP 8, 09042 Monserrato, Italy*
[5]*School of Chemical Engineering, National Technical University of Athens, Zografou Campus, 9 Heroon Polytechniou Street, 15780 Athens, Greece*
[6]*School of Information Management, Shandong University of Traditional Chinese Medicine, Jinan 250355, China*
[7]*Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, UK*

Correspondence should be addressed to Tai-Ping Fan; tpf1000@cam.ac.uk and Andreas Bender; ab454@cam.ac.uk

Received 14 September 2015; Accepted 3 December 2015

Academic Editor: Jae Youl Cho

Traditional Chinese medicine (TCM) still needs more scientific rationale to be proven for it to be accepted further in the West. We are now in the position to propose computational hypotheses for the mode-of-actions (MOAs) of 45 TCM therapeutic action (sub)classes from *in silico* target prediction algorithms, whose target was later annotated with Kyoto Encyclopedia of Genes and Genomes pathway, and to discover the relationship between them by generating a hierarchical clustering. The results of 10,749 TCM compounds showed 183 enriched targets and 99 enriched pathways from Estimation Score ≤ 0 and $\geq 5\%$ of compounds/targets in a (sub)class. The MOA of a (sub)class was established from supporting literature. Overall, the most frequent top three enriched targets/pathways were immune-related targets such as tyrosine-protein phosphatase nonreceptor type 2 (PTPN2) and digestive system such as mineral absorption. We found two major protein families, G-protein coupled receptor (GPCR), and protein kinase family contributed to the diversity of the bioactivity space, while digestive system was consistently annotated pathway motif, which agreed with the important treatment principle of TCM, "the foundation of acquired constitution" that includes spleen and stomach. In short, the TCM (sub)classes, in many cases share similar targets/pathways despite having different indications.

1. Introduction

Traditional Chinese medicine (TCM) has been practiced for thousands of years for the prevention and treatment of diseases using a unique system of theory, diagnosis, and treatment [1, 2]. The philosophical background of TCM is based on Yin and Yang, as well as the Five Elements (agents) theories. The Yin and Yang are the harmony of two opposite energies and the Five Elements describe the five interdependent functional organs, heart, liver, spleen, lung, and kidney, with each organ's own Yin and Yang [3, 4]. When a human body suffers from a disease, the dynamic balance

and the relationship of the Five Elements are disturbed; hence, to rectify the disturbance, TCM applies a holistic approach with the key therapeutic principles being "*Zheng*" (meaning syndrome or pathological patterns seen in patients) and "*Fufang*" or "*Fang Ji*" (meaning compound formulations consisting of materia medica) [5–11]. The Chinese medicines can be organized into several classifications such as therapeutic actions, source of the medicine, and internal organs [12]. In this study, the classification of the Chinese medicines follows the therapeutic actions, of which some also possess subclassifications based on clinical applications recorded by TCM monographs (Table 1) [13]. A combination of two or more Chinese medicine categories makes up a treatment formula, which then contains a considerable number of chemical compounds [14].

The mixture of compounds in the formula works through the therapeutic principle *Jun-Chen-Zuo-Shi*, by maximizing the therapeutic effects and minimizing the side effects [1, 15, 16]. Based on the healing/pharmacological properties and constituents of each medicine, the Jun (emperor) component is the principal phytocomplex targeting the major symptom of the disease. There are only a few varieties of Jun medicinal that are administered as a single formula, usually in large doses. The Chen (minister) components synergize with Jun to strengthen its therapeutic effects and may also treat secondary symptoms. The Zuo (assistant) medicinal reduces or eliminates possible adverse or toxic effects of the Jun and/or Chen components, while also enhancing their effects and sometimes treating secondary symptoms. Finally, the Shi (courier) components facilitate delivery of the principal components to the lesion sites or facilitate the overall action of the other components [17, 18]. Therefore, at the molecular level, a TCM formula, which is a multicomponent and multitarget agent, is assumed to modulate a series of protein targets in an integrative manner to harmonise the body system [19]. In brief, TCM is a well-structured system from diagnosis to healing, whose theories and medicines are rationally connected and interdependent. Note that a TCM *Fufang* is primarily based on medicinal plants but may also contain fungi (e.g., *Ganoderma lucidum*), mineral (realgar), and occasionally animal products (e.g., *Calculus bovis*).

TCM compounds derives from the biodiversity of natural products, of which is a rich resource for discovering new TCM-based drugs. However, to develop a novel TCM-based drug still remains challenging. One of the factors that contributes to making it a challenge is an undefined medicine concoction, where characterising the complex formulation by using methods to isolate the compounds is an exhaustive task that is very time-consuming [20, 21]. Although many compounds have been isolated from Chinese herbals [22, 23], their modes-of-action (MOAs), in many cases, are still not yet understood at the molecular level [21]. Another challenging issue in TCM is measuring the efficacy, pharmacokinetic-pharmacodynamic profiles, and dose-efficacy relationship of multiple compounds simultaneously, both *in vitro* and *in vivo* [20, 21, 24]. However, as shown in this work, particularly related to MOAs, it is now possible to suggest the MOA of TCM compounds using *in silico* target prediction, hence providing testable hypotheses to guide towards finding new molecular entities.

Various methods are available for *in silico* target prediction such as chemical similarity searching [25], analysis of "bioactivity spectra" [26, 27], data mining in annotated chemical database [28, 29], and protein panel docking [30, 31]. Protein panel docking is one of the earliest tools that has been widely used in TCM, where a compound is docked to a wide panel of potential proteins and the proteins are subsequently ranked based on calculated binding affinity scores [32]. Ehrman et al. employed pharmacophore-assisted docking to screen TCM compounds active against inflammation against four protein targets, which were cyclooxygenases (COX) 1 and 2, p38 MAP kinase (p38), c-Jun terminal-NH2 kinase, and type 4 cAMP-specific phosphodiesterase (PDE4) [33]. In a different study, where the *in silico* target prediction was further validated experimentally, Zhang et al. identified putative targets of 19 natural products from two medicinal plants in TCM which were used for the treatment of diabetes and inflammation using a reverse docking approach, TarFisDock server [34]. The natural products showed moderate inhibitory activities against the most frequent target candidate, dipeptidyl peptidase IV (DPP-IV), with IC_{50} values ranging from 14.14 μM to 113.76 μM in an *in vitro* enzyme assay [34]. Although protein panel docking method requires only the chemical structures of the putative active ingredients, it is limited to high quality protein structures and the accuracy of the docking programs used [35].

In this study, the methodology opted for was the ligand-based target prediction based on large bioactivity database available. Data mining in annotated chemical databases was used to predict the protein targets of TCM compounds to suggest their MOAs. This method becomes viable due to the increasing availability of bioactivity databases [36, 37]. The target prediction algorithm consists of small molecule databases annotated with bioactivity data to map "chemical space" onto "biological activity space." Based on the principle of molecular similarity, the method, by generating a statistical model using the available data, measures the likelihood of an orphan compound to modulate a target [38]. The model infers the ligand-target modulation based on the molecular similarity, which can suggest the MOA of an orphan compound of TCM by associating the target with the known phenotypic effects of the compound (Figure 1).

Related methods have been developed before and applied in various different settings. One of the earliest studies published on target prediction is the Prediction of Bioactivity Spectra for Substances (PASS) method [39]. This method, which was developed using structure-activity relationships for more than 300,000 biological active compounds in its training set, allows the prediction of more than 4,000 biological activities, based on the structural formula of a compound, with an average accuracy above 95% in the internal model evaluation. Nidhi et al. used a large chemogenomic dataset to develop the *in silico* target prediction [28]. In this work, compound-target pairs from the WOrld of Molecular BioAcTivity (WOMBAT) database were extracted and models trained on extended-connectivity fingerprints to predict the most likely targets for all compounds from a different

Table 1: The list of therapeutic action classes, subclasses, and their number of respective Chinese medicines and compounds. A total of 10,749 compounds from 46 therapeutic action subclasses were included in the analysis presented in this work.

Chinese medicine class	Chinese medicine subclass	Chinese medicine (sub)class (Chinese names)	Abb.	NoH	NoC
Exterior releasing	Wind cold dispersing	Sàn hán jiě biǎo yào (散寒解表药)	ER-wind cold	21	538
	Wind heat dispersing	Qīng rè jiě biǎo yào (清热解表药)	ER-wind heat	22	413
Heat-clearing medicinal	Heat-clearing and blood cooling	Liáng xuè huó xuè yào (凉血活血药)	HC-blood cool	14	99
	Heat-clearing and dampness drying	Qīng rè lì shī yào (清热利湿药)	HC-damp	11	264
	Deficiency	Qīng xū rè yào (清虚热药)	HC-def	10	186
	Heat-clearing and detoxicating	Qīng rè jiě dú yào (清热解毒药)	HC-detox	54	1029
	Heat-clearing and fire purging	Qīng rè xiè huǒ yào (清热泻火药)	HC-fire purge	18	234
Purgative medicinal	Laxative medicinal	Rùn xià yào (润下药)	Purg-lax	3	27
	Offensive purgative	Gōng xià yào (攻下药)	Purg-off	6	54
	Drastic (purgative) water-expelling	Jùn xià zhú shuǐ yào (峻下逐水药)	Purg-water expel	14	206
Dampness resolving	Water draining and anti-icteric	Lì shī tuì huáng yào (利湿退黄药)	Damp-antiicteric	6	189
	Water draining and strangury resolving	Lì niào tōng lín yào (利尿通淋药)	Damp-stran	15	133
	Water draining and swelling dispersing	Lì shuǐ xiāo zhǒng yào (利水消肿药)	Damp-swell	13	265
Qi regulating		Lǐ qì yào (理气药)	Qi	36	699
Digestant medicinal		Xiāo shí yào (消食药)	Digest	8	146
Hemostatic medicinal	Astringent hemostatic	Shōu liǎn zhǐ xiě yào (收敛止血药)	Hemo-astringent	5	92
	Blood cooling hemostatic	Liáng xuè zhǐ xiě yào (凉血止血药)	Hemo-blood cool	13	198
	Meridian warming hemostatic	Wēn jīng zhǐ xuè yào (温经止血药)	Hemo-meridian	2	146
	Stasis resolving hemostatic	Huà yū zhǐ xiě yào (化瘀止血药)	Hemo-stasis	6	245
Blood activating and stasis resolving	Blood activating analgesic	Huó xuè zhǐ tòng yào (活血止痛药)	BASR-analgesic	7	487
	Blood breaking mass eliminating	Pò xiě xiāo zhēng yào (破血消癥药)	BASR-break	9	177
	Blood activating menstruation resolving	Huó xuè tiáo jīng yào (活血调经药)	BASR-menstrual	15	457
	Blood activating trauma curing	Huó xuè liáo shāng yào (活血疗伤药)	BASR-trauma	12	261
Cough suppressing and panting-calming	Clearing and heat phlegm resolving	Qīng huà rè tán yào (清化热痰药)	CSPC-heat	30	237
	Cold phlegm resolving and warming	Wēn huà hán tán yào (温化寒痰药)	CSPC-cold	19	147
	Cough suppressing and panting-calming	Zhǐ ké píng chuǎn yào (止咳平喘药)	CSPC-panting	16	334

Table 1: Continued.

Chinese medicine class	Chinese medicine subclass	Chinese medicine (sub)class (Chinese names)	Abb.	NoH	NoC
Tranquilizing	Heat nourishing tranquilizing	Yǎng xīn ān shén yào (养心安神药)	Tranquil-heat	1	145
	Settling tranquilizing	Zhòng zhèn ān shén yào (重镇安神药)	Tranquil-settle	6	1
Orifice opening		Kāi qiào yào (开窍药)	Orifice	7	68
Liver-pacifying and wind extinguishing	Extinguishing wind to resolve convulsion	Xí fēng zhǐ jìng yào (息风止痉药)	LPWE-convulsion	8	85
	Liver yang calming	Píng yì gān yáng yào (平抑肝阳药)	LPWE-liver	7	22
Tonifying and replenishing	Blood tonifying	Bǔ xiě yào (补血药)	TR-blood	7	388
	Qi tonifying	Bǔ qì yào (补气药)	TR-qi	15	474
	Yang tonifying	Bǔ yáng yào (补阳药)	TR-yang	23	559
	Yin tonifying	Bǔ yīn yào (补阴药)	TR-yin	17	259
Astringent	Anhidrotic	Gù biǎo zhǐ hàn yào (固表止汗药)	Ast-anhidro	3	17
	Lung-intestine astringent	Liǎn fèi sè cháng yào (敛肺涩肠药)	Ast-lung	8	145
	Securing essence, reducing urination, and checking vaginal discharge	Gù jīng suō niào zhǐ dài yào (固精缩尿止带药)	Ast-secure	6	125
Wind-dampness dispelling	Bone (sinew) strengthening	Qū fēng shī qiáng jīn gǔ yào (祛风湿强筋骨药)	WD-bone	5	44
	Heat-clearing	Qū fēng shī rè yào (祛风湿热药)	WD-heat	8	175
	Cold dispersing	Qū fēng hán shī yào (祛风寒湿药)	WD-cold	13	309
Interior warming		Wēn lǐ yào (温里药)	Warm	13	457
Worm expelling medicinal		Qū chóng yào (驱虫药)	Worm	9	93
Emetic medicinal		Yǒng tǔ yào (涌吐药)	Emetic	3	9
Parasite destroying, dampness eliminating, and itchiness relieving		Gōng dú shā chóng zhǐ yǎng yào (攻毒杀虫止痒药)	Parasite	8	81
Antimalarial medicinal		Kàng nüè yào (抗疟药)	Malarial	4	30
			Total compounds		10,749

Abb.: abbreviation.
NoH: number of Chinese medicines.
NoC: number of compounds.

database, the MDL Drug Database Report (MDDR). The results showed that, on average for 10 test cases, the top three targets were predicted correctly for 77% of the compounds. In the Similarity Ensemble Approach (SEA), the target prediction was developed by grouping and relating the protein targets based on chemical similarity among their ligands, which were quantitatively measured using an algorithm adapted from BLAST [40]. The study discovered that the chemical structure of methadone showed close structural similarity to ligands for the M3 muscarinic receptor, and the prediction was validated experimentally. In TCM, this method has also been applied to study either TCM compounds in general

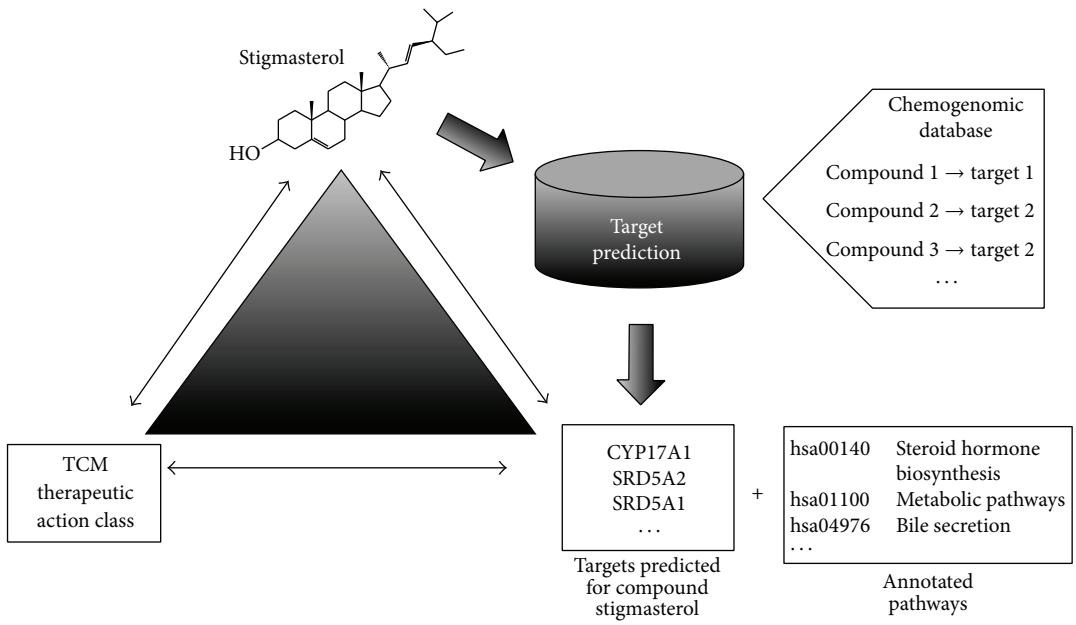


Figure 1: Visualization of the link between orphan compounds taken from TCM databases, predicted targets, and TCM medicinal subclasses. The mode-of-action of compounds in TCM can be hypothesised *via* an *in silico* target prediction algorithm. A predicted target can furthermore be annotated with pathways, which could provide a better insight into the compound's MOA.

or, specifically, compounds from TCM formulations. *In silico* target prediction was developed by integrating chemical, genomic, and pharmacological approaches using Random Forest and Support Vector Machine with sensitivity of 81.33% and specificity of 93.62% for internal validation [41]. This model was applied by Li et al. to study the MOA of Chinese herbs for the treatment of cardiovascular disease, with the predicted targets being further validated by molecular docking approach [42]. The model was also applied to predict proteins for TCM formulations, Xiao-Chaihu-Decoction and Da-Chaihu-Decoction, with their MOAs being verified by building network pharmacology [43].

The motivation of the current study is prompted by the success to link TCM and Ayurveda compounds to the predicted targets that were relevant for the indications of various therapeutic action classes in TCM and cancer in Ayurveda [44]. In the study, an *in silico* target prediction method using the Naïve Bayes Classifier was able to associate TCM compounds from the "tonifying and replenishing" therapeutic action class, for example, with their known phenotype, hypoglycaemic effect, through the predicted targets such as sodium glucose cotransporter 1 (SGLT1), SGLT2, and protein tyrosine phosphatase to explain their MOA. In addition, in few cases, the predicted targets led to novel MOA and side effects being G-protein bile acid receptor 1 (GPBAR1), predicted from Ayurveda compounds, which contributes to cancer pathogenesis through the apoptosis and cell proliferation signalling pathway. The aim of the study was to extend the analysis of the targets prediction compounds of TCM therapeutic action classes including their subclasses by first generating hierarchical clustering. The first part of the study was to understand the MOAs of the subclasses from the predicted targets as well as from annotated KEGG pathways because target alone is insufficient to provide a full biological profile towards the effect of the ligand on a biological system [45, 46]. In the second part, bioactivity spaces of all therapeutic action (sub)classes were compared in order to understand relationships between the clusters. Hence, the global mapping of TCM compounds explored in this work, based on the therapeutic action (sub)classes, not only does provide better insight of the MOAs of the TCM compounds but also describes for the first time the relationship between the therapeutic action (sub)classes.

2. Materials and Methods

2.1. Dataset and Dataset Preparation. TCM compounds were obtained from TCM Database@Taiwan [13] in SD format. A total of 13,091 compounds from 46 different therapeutic action (sub)classes were imported into MOE [47]. To prepare structures for further analysis, covalently bound Group I metals were disconnected into ionic representation, while keeping only the largest molecular fragments, neutralising the compound, and adjusting the hydrogen and partial charges using MMFF94 (modified) partial charges. The duplicates from each therapeutic action (sub)class were removed, resulting in a total number of 10,749 distinct compounds. The list of classes with the final number of compounds for each (sub)class is provided in Table 1.

2.2. Target Prediction. The processed molecular data was then subjected to target prediction that was modelled using the Laplacian-modified Naïve Bayes Classifier; the details of the target prediction algorithm can be found in [48]. Briefly,

this model contained 189,147 ligand-protein pairs extracted from CHEMBL [49] v.14.0 across 477 human targets which was used as the training set. The training set contained active compounds with reported activities ($K_i/K_d/IC_{50}/EC_{50}$) of at least 10 μM with a confidence score of 8 or 9 and at least 20 compounds were available to associate the chemical features with a target class.

The molecular descriptors of the compounds were represented by Molprint2D circular fingerprints which have been shown to relate chemical structural space and bioactivity space well with each other [50]. The targets predicted for the new compound use the Naïve Bayes Classifier as a method for classification as follows [51]:

$$P(\omega_\alpha \mid \varkappa) = \frac{P(\varkappa \mid \omega_\alpha)\, P(\omega_\alpha)}{P(\varkappa)}. \quad (1)$$

Here, the probability of a new compound belonging to a target class, ω_α, with a given vector molecular feature, $\varkappa$, is calculated. The prior target class probability, $P(\omega_\alpha)$, is assumed to be equal in proportion to the molecules, $N\omega$, which modulate that class from the total number of molecules, N, in the training set. Hence,

$$P(C = \omega) = \frac{N\omega_\alpha}{N}. \quad (2)$$

The denominator, $P(\varkappa)$, calculates the sum of the fractions of molecules from each class from $\beta = 1$ to L in the training set multiplied by the probability of the vector of molecular feature given the target class:

$$\sum_{\beta=1}^{L} \frac{N\omega_\beta}{N} P(\varkappa \mid \omega_\beta). \quad (3)$$

The posterior probability, $P(\varkappa \mid \omega_\alpha)$, is the likelihood of the feature, $\varkappa$, given the class, ω_α.

The internal validation of the *in silico* target prediction was measured using the 5-fold cross validation with a recall of correct targets larger than 80% in the top 1% of predictions. In the external validation, the algorithm showed a recall of 63.6% in the top 1% of predictions using dataset extracted from WOrld of Molecular BioAcTivity version 2011.1 (WOMBAT) [52].

Only target scores above a confidence score, which are defined individually for each target class, were taken as the output. The confidence score for each class was calculated by the optimal balanced accuracy (precision and recall tradeoff) on a per-target class basis and was used to retain protein targets likely to interact with the compounds in the dataset [53].

2.3. Pathway Annotation. Each predicted target was annotated with its full set of pathways from KEGG biological pathways (release 58.1) [54]. It was possible to annotate 405 out of 477 targets with KEGG biological pathways [55, 56].

2.4. Enrichment Calculations. To normalise the classification results from target prediction/pathway annotation, enrichment calculations were performed by normalising frequencies of the target prediction/pathway annotation of each therapeutic action subclass of compound to a background of 10,000 compounds that were selected randomly from PubChem [57] and ZINC [58], which consists of 194,849 compounds in total. Two methods were used to perform enrichment calculation in this step, namely, Estimation Score and Average Ratio. The calculation of the scores was performed as follows [55, 56].

2.4.1. Estimation Score. The Estimation Score is based on the frequency on the number of predicted targets/pathways in the random dataset larger than or equal to the frequency on the number of predicted targets/pathways in the test dataset. The absolute frequency (C in (4)) was divided by the total number of random datasets, giving a value between 0 (enriched) and 1 (random). Hence,

$$\text{Estimation Score} = \frac{C}{1000}. \quad (4)$$

2.4.2. Average Ratio. The Average Ratio is calculated by the ratio of the frequency (F) of predicted target/annotated pathway in each random dataset with the frequency (F) of predicted targets in the test dataset:

$$\text{Average Ratio} = \frac{F(\text{random set1})/F(\text{test set}) + F(\text{random set2})/F(\text{test set}) + \cdots + F(\text{random set10000})/F(\text{test set})}{10000}. \quad (5)$$

In this study, enriched targets/pathways were considered if they showed an Estimation Score ≤ 0.01, and descending Average Ratio was used to further discriminate important targets in agreement with previous work by Liggi et al. [55, 56]. The relative cutoff of both target frequency (TF)/pathway frequency (PF) that was ≥5% of the highest predicted target/pathway frequency was used after ranking the targets using the Estimation Score and Average Ratio methods to determine which targets were considered to be enriched in particular therapeutic action (sub)classes.

2.5. Hierarchical Clustering Based on the Bioactivity Space of the Therapeutic Action (Sub)classes. The frequencies of compounds across 477 targets for each therapeutic action (sub)class were subjected to agglomerative hierarchical clustering [59, 60]. The clustering method involved two steps as follows.

2.5.1. Selecting Measures of Dissimilarity or Similarity. The dissimilarity distance between two therapeutic action

(sub)classes was calculated using the "dist" function in *R* [61] using the "Euclidean" method after scaling the frequencies.

2.5.2. Clustering. Clustering was performed using the "hclust" function in *R* [61] based on the previously calculated Euclidean distance and Ward's clustering method [62]. In this method, two clusters were merged if the sum square Euclidean distance was minimal.

A cutoff dissimilarity distance of approximately 20 was applied in order to obtain a manageable number of clusters, defining 14 groups of therapeutic action (sub)classes, namely, clusters I to XIV.

2.6. Targets and Pathways Analysis. The top three enriched targets/pathways were inspected with regard to their ability to explain the MOA of the compounds classified in the therapeutic action (sub)classes. To improve the mechanistic understanding of the MOA, the top three enriched targets/pathways were linked to the phenotypes of the (sub)classes with supporting evidence from the literature and supporting *in vitro* or *in vivo* studies of the Chinese medicines' extracts or isolated compounds. However, supporting *in vitro* and *in vivo* studies were excluded in the pathway analysis because, in many (sub)classes, no information was found. The 14 clusters were grouped based on the number of (sub)classes in a cluster, which ranged from ten (sub)classes to only one (sub)class that derived from different classes. Three clusters were analysed in detail in the next section. Cluster VII was the only cluster composed of (sub)classes with the same TCM vital substance of its meridian system, blood. Cluster X was picked as a representative of a cluster of different classes while cluster XII was selected as a representative of a cluster with only one type of (sub)class. The top three enriched targets/pathways per selected (sub)classes were summarized in Tables 3 and 4 and the top three enriched targets/pathways of all (sub)classes can be found in the Supplementary Material available online at http://dx.doi.org/10.1155/2016/2106465.

To compare the bioactivity spaces among the clusters, all enriched targets in a cluster were classified according to their protein family as derived from UniProt [63]. The enriched pathways for all clusters were also classified according to KEGG ortholog groups, which are derived by comparing sequence similarity of individual genes and defining the functional group from the list of genes in the respective group [45]. The KEGG ortholog group will be called a pathway motif onwards. All the enriched targets were annotated from 59 protein families and the enriched pathways were annotated from 33 pathways motifs. A major protein family/pathway motif for a cluster was defined if the number of enriched targets/pathways was at least 5% of the total number of enriched targets/pathways in the respective cluster and it was present in at least eight clusters. Only five major protein families and eight major pathway motifs were identified. The frequencies of enriched targets/pathways per cluster were normalised before constructing two different heatmaps using heatmap.2 function of gplots package in *R* [61], in order to visualize whether the major protein families/pathway motifs were equally important across clusters.

3. Results

3.1. Target Prediction and Pathway Annotation. The *in silico* target prediction of 10,749 TCM compounds yielded 409 unique targets, of which 183 were enriched targets. In the pathway annotation, the total number of unique pathways was 171, of which 99 were enriched pathways. The results discussed from here onwards cover 45 of the 46 therapeutic action (sub)classes only. One therapeutic action subclass is not included because no target was retained from the "Tranquilizing-Settling" (Tranquil-settle) subclass, which contained only one compound. This subclass was therefore omitted from hierarchical clustering and all subsequent analyses.

3.2. Hierarchical Clustering. In Figure 2, a dendrogram shows the hierarchical clustering of 45 TCM therapeutic action (sub)classes based on their bioactivity fingerprints. The cluster tree generates diverse spread of the 45 therapeutic action (sub)classes, which is defined into 14 clusters. In many instances, branches of the dendrogram paired up from different types of classes or subclasses. Based on the molecular similarity principle [64], this observation indicates that many similar compounds are present in both (sub)classes despite having different therapeutic actions. The link of the top three enriched targets of the therapeutic action (sub)classes is discussed in the next sections as well as their top three enriched pathways.

3.3. Target Analysis

3.3.1. Clusters with Four Therapeutic Action (Sub)classes (Cluster X). Only cluster X represents this group. The subclasses in cluster X are "wind-dampness dispelling, bone (sinew) strengthening" (WD-bone), "tonifying and replenishing, qi tonifying" (TR-qi), "cough suppressing and panting-calming, clearing and heat phlegm resolving" (CSPC-heat), and "tranquilizing, heat nourishing tranquilizing" (Tranquil-heart). The top three enriched targets from cluster X are mainly implicated in immunomodulation, namely, steryl-sulfatase (STS) [65], tyrosine-protein phosphatase nonreceptor type 2 (PTPN2) [66], and peptidyl-prolyl cis-trans isomerase FKBP1A (FKB1A) [67], glucose homeostasis, such as sodium glucose transporter 1 (SGLT1) and sodium glucose transporter 2 (SGLT2) [68], cancer, such as DNA topoisomerase 1 (TOPO1) [69], reproductive system, such as testosterone 17-beta-dehydrogenase 3 (17-beta-HSD 3) [70], and central nervous system (CNS), such as glutamate carboxypeptidase 2 (CGPII) [71].

Dissecting the *in silico* target prediction per subclass, starting from the "wind-dampness dispelling, bone strengthening" (WD-bone), the top three enriched targets are TOPO1 (Estimation Score (ES) = 0, Average Ratio (AR) = 0.0144), SGLT1 (ES = 0, AR = 0.0342), and STS (ES = 0, AR = 0.0370). From the *in silico* target prediction, acankoreosides A–C derived from *Acanthopanax gracilistylus* were suggested to modulate TOPO1 and SGLT1 in the subclass, while compounds from *Homalomena occulta* such as asperpenoid, bullatantriol, and homalomenol were suggested to modulate STS, as well as compound such as quercetin of *Taxillus*

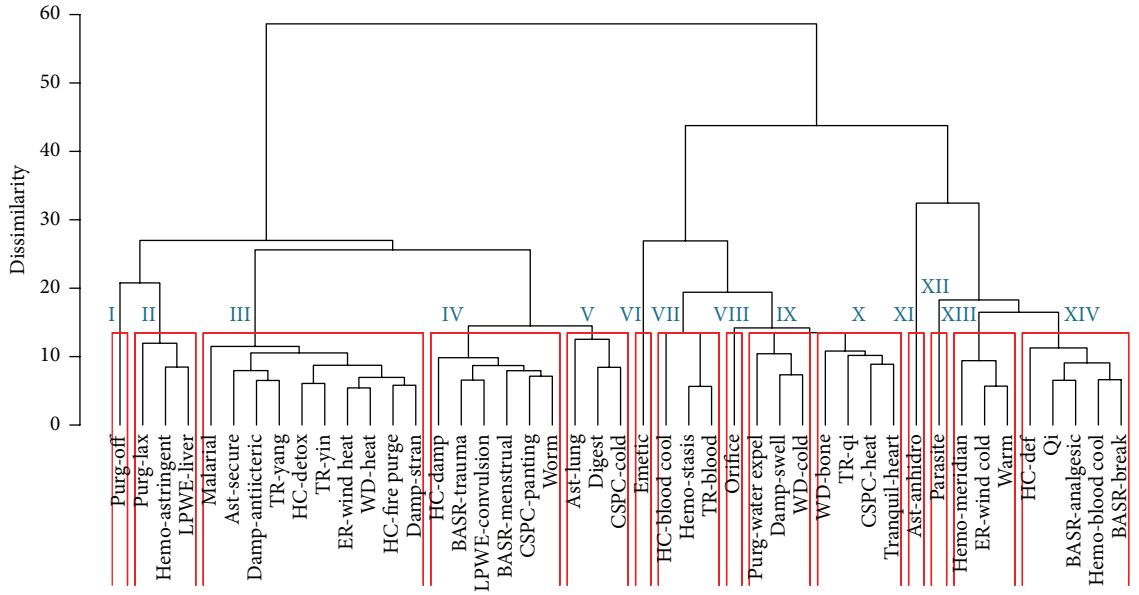


Figure 2: Hierarchical clustering analysis of TCM classes and subclasses is based on the similarity of the bioactivity fingerprint of each class. The "Tranquil-settle" subclass was not included here (and in the further analysis) since it only contained a single compound for which no reliable targets could be predicted. The (sub)classes were defined into 14 clusters, where clusters VII, X, and XII were selected for further analysis based on the top three enriched targets/pathways.

chinensis. The target prediction of compounds from *Acanthopanax gracilistylus* is supported by an *in vitro* study where the herb's extract inhibits cell proliferation of several types of cancer cells [72]. It is also reported that the extract of *Taxillus chinensis* exhibits significant anti-inflammatory activity *in vitro* [73]. The herbs from this class are used to relieve pain, relax muscle and tendons, open channels and collaterals, and strengthen tendons and bones [12, 74]. The actions of the herbs are related to the disturbance of muscular function in diabetes [75] and cancer cachexia which affects protein and lipid metabolism in skeletal muscle [76].

The top three enriched targets for "tonifying and replenishing, qi tonifying" (TR-qi) subclass are PTPN2 (ES = 0, AR = 0.0174), SGLT2 (ES = 0, AR = 0.0282), and SGLT1 (ES = 0, AR = 0.0321). Many compounds from *Glycyrrhiza glabra*, *Glycyrrhiza uralensis*, *Dolichos lablab*, *Panax ginseng*, and *Astragalus membranaceus* were predicted to modulate PTPN2 and SGLT1. In particular to PTPN2 [77], a compound, glycyrrhizin from *Glycyrrhiza glabra* and *Glycyrrhiza uralensis*, is supported by an *in vivo* study that suggests the compound ameliorates all established chronic histopathologic changes of lung tissue in the mouse model of asthma [78]. TCM describes the notion that the medicines from this subclass act on the spleen and lung, in which the deficiency of lung qi is characterised by shortness of breath like in asthma [12].

A subcluster consists of two therapeutic action subclasses seen to have highly similar bioactivity space. The first subclass is "cough suppressing and panting-calming, clearing and heat phlegm resolving" (CSPC-heat) with the top three enriched targets being PTPN2 (ES = 0, AR = 0.0112), TOPO1 (ES = 0, AR = 0.0236), and 17-beta-HSD 3 (ES = 0, AR = 0.0341). Compounds from *Platycodon grandiflorum* such as platycosides A–M were suggested to be modulated by all the top three enriched targets and compounds from *Bambusa tuldoides*, *Peucedanum decursivum*, and *Trichosanthes kirilowii* were mostly predicted to modulate 17-beta-HSD 3. Although *Platycodon grandiflorum* [79] and *Trichosanthes kirilowii* [80] are reported to exhibit anticancer properties, the reports do not support the link between the top three enriched targets and the indication of therapeutic action subclass to rationalise the MOA of the compounds.

The second subclass, "tranquilizing, heart nourishing tranquilizing" (Tranquil-heart), lists FKB1A (ES = 0, AR = 0.0077), PTPN2 (ES = 0, AR = 0.0167), and CGPII (ES = 0, AR = 0.0209) as its top three enriched targets. Compounds of three different herbs were predicted to modulate FKB1A and PTPN2 and only compounds of *Ganoderma lucidum*, such as ganoderic acid and lucidenic acid derivatives, were predicted to modulate CGPII. The prediction of CGPII from triterpenoids of *Ganoderma lucidum* such as ganoderic acids is supported by Zhang et al., where triterpenoids exhibit nerve growth factor or brain-derived neurotrophic factor activities *in vitro*, which has the therapeutic potential in neurodegenerative diseases [81]. The therapeutic actions of the herbs from the subclass are described to have effects on central nervous system [12]. In conclusion, the MOAs of the compounds for three subclasses in cluster X can be suggested from their top three enriched targets.

3.3.2. Clusters with Three Therapeutic Action (Sub)classes (Cluster VII). Cluster VII is one of the clusters that consists of three therapeutic action subclasses, in which two subclasses, "hemostatic, stasis resolving" (Hemo-stasis) and "tonifying and replenishing, blood" (TR-blood), show highly similar bioactivity space. The third subclass is "heat-clearing, blood cooling" (HC-blood cool). Overall the top three enriched

targets in the cluster can be classified into immunomodulation, which are PTPN2 [66], protein kinase C beta type (PKC-β) [82], protein kinase C eta type (PKC-η) [83], and protein kinase C gamma type (PKC-γ) [84], cancer, namely, TOPO1 [69], and glucose homeostasis, such as SGLT2 [68]. The top three enriched targets in "hemostatic, stasis resolving" (Hemo-stasis) and "tonifying and replenishing, blood" (TR-blood) are PTPN2 (ES = 0, AR = 0.0089), PKC-η (ES = 0, AR = 0.0182), PKC-γ (ES = 0, AR = 0.0209), and PTPN2 (ES = 0, AR = 0.0140), PKC-β (ES = 0, AR = 0.0230), and PKC-ε (ES = 0, AR = 0.0240), in which all of them are implicated in immunomodulation. In the "hemostatic, stasis resolving" (Hemo-stasis) subclass, compounds such as ginsenosides and notoginsenosides of *Panax notoginseng* were predicted to modulate all top three enriched targets. The target prediction also showed that anthraquinone compounds from *Rubia cordifolia* such as purpurin, ruberythric acid, and soranjidiol modulate PKC-β and PKC-ε. In support of the target prediction of anthraquinone compounds of the second herb, a study reported that the herb's ethanol extract shows wound healing activities in mice, which from histological evaluations indicate marked infiltration of the inflammatory cells, increased blood vessel formation, and enhanced cells proliferation [85]. This finding agrees with the description of the subclass to stop bleeding [12].

In the "tonifying and replenishing, blood" (TR-blood) subclass, compounds from *Paeonia lactiflora* such as albiflorin, gallotannin, and casuarictin were predicted to modulate PKC-β and PKC-η. PTPN2 and PKC-η were both frequently predicted to be modulated by compounds from *Panax notoginseng* such as notoginsenoside and ginsenoside. It is found that ginsenoside Rg1 of *Panax notoginseng* ameliorates liver damage and suppresses proinflammatory cytokines secretion in concanavalin A-induced hepatitis in mice [86]. This subclass is described to have pharmacological effects on the liver, heart, and spleen and prevent failures of the organs [12]. The top three enriched targets for "heat-clearing, blood cooling" (HC-blood cool) subclass are PKC-β (ES = 0, AR = 0.0100), TOPO1 (ES = 0, AR = 0.0123), and SGLT2 (ES = 0, AR = 0.0137). In this subclass, compounds from *Paeonia lactiflora* such as albiflorin, isopaeoniflorin, and benzoylpaeoniflorin were predicted to modulate both PKC-β and TOPO1, while SGLT2 was predicted to be modulated by compounds from *Rehmannia glutinosa* such as rehmanniosides A and B. We found that the target prediction of SGLT2 is supported by a study on stachyose extract from *Rehmannia glutinosa*, which shows a significant hypoglycaemic effect in diabetic mice [87]. TCM views the notion that the action of the subclass is to promote the generation of body fluids from excessive heat [12, 74] where the consumption of body fluid is one of the symptoms in diabetes [88] and cancer patients [89]. Altogether, in many instances, the MOAs of the compounds can be explained from the enriched targets and can also be linked to the indications of the (sub)classes.

3.3.3. Cluster with One Therapeutic Action (Sub)class (Cluster XII). One out of five clusters that has only one therapeutic action class is cluster XII, "parasite destroying, dampness eliminating, and itchiness relieving" (Parasite). The top three enriched targets are dihydrofolate reductase (DHFR), which plays a role in bacterial infection and cancer [90] (ES = 0, AR = 0.0532), DNA-dependent protein kinase catalytic subunit (DNA-PKCs), which is implicated in cancer [91] (ES = 0, AR = 0.0644), and tumour necrosis factor (TNF), which is found to exert activities in cancer, inflammation, and bacterial infection [92, 93] (ES = 0, AR = 0.0687). From the target prediction, compounds from *Allium sativum* were predicted to modulate DHFR, such as allicin, allithiamine, and allyl disulphide. This prediction agrees with a study by Adetumbi et al., from which the extract of *Allium sativum* is found to inhibit the synthesis of proteins, nucleic acids, and lipids in *Candida albicans* where the major component of the herb was allicin [94]. The finding relates to the phenotypes of the subclass, which is to kill and expel parasites and subsequently relieve pain [12, 74]. In brief, the MOAs of the class can be linked to the enriched targets and the indications of the subclass. Table 2 summarises the top three enriched targets of the therapeutic action (sub)classes.

The above analysis can be summarized into two different views, namely, biological space and chemical space. From the biological space's view, many of the top three enriched targets, regardless of therapeutic action (sub)classes in any of the clusters, are implicated in immunomodulation such as PTPN2, PKCs, FKBP1A, and STS, in which PTPN2 and PKCs were frequently predicted. Both PKC family and PTPN2 are implicated in immunomodulation. The frequency of immune-related targets can be related when TCM balances the immune system regulation by either promoting or suppressing the immune factors [95]. The PKC isoenzymes act as important mediators in immune cellular signalling in T- and B-lymphocytes in acquired immune system [96]. PTPN2 also plays a major role in the transmission of immune cell signalling events [66]. From the chemical space's view, triterpenoid is the most frequent phytochemical (Table 3), which was predicted to modulate few of the top three enriched targets, such as TOPO1, PTPN2, and PKCs. The compounds are found in different herbs such as *Acanthopanax gracilistylus*, *Glycyrrhiza glabra*, *Platycodon grandiflorum*, *Ganoderma lucidum*, and *Panax notoginseng*. To validate the result from the chemogenomic principle that similar targets share similar compounds, a compound, CHEMBL 1986122, from CHEMBL database [49] is most similar to acankoreoside A, with Tanimoto coefficient (TC) value of 0.91 (Table 4). The GI_{50} value of CHEMBL 1986122 was 43.05 nM in tumour cell line growth inhibition assay. Hence, the chemogenomics principle appears to stand (at least for acankoreoside A) because both compounds are implicated in cancer and are structurally similar.

3.4. Pathway Analysis

3.4.1. Clusters with Four Therapeutic Action (Sub)classes (Cluster X). The top three enriched pathways in cluster X are mainly implicated in digestive system (carbohydrate digestion and absorption, bile secretion, and mineral absorption). In this cluster, only one pathway, from the top three enriched pathways, is not classified in digestive system, which is terpenoid backbone biosynthesis.

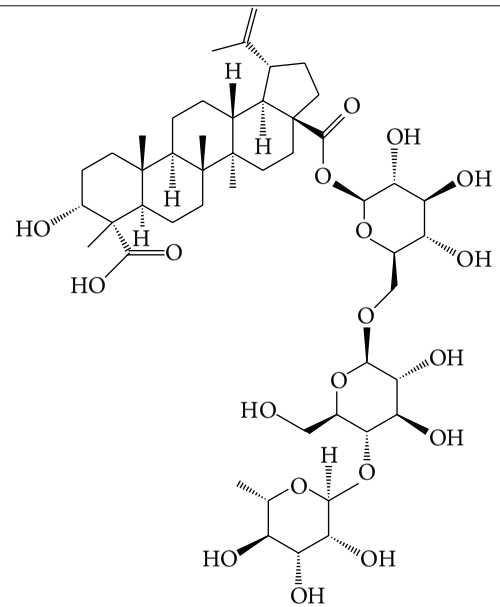
Table 2: The top three enriched targets in clusters VII, X, and XII. It can be seen that, in many cases, the top three enriched targets are implicated in immunomodulation. Estimation Score = 0, for all top three enriched targets.

	TCM therapeutic action class	TCM therapeutic action subclass	Top three enriched targets	Target function reported by literatures	Average Score
Cluster X	Wind-dampness dispelling	Bone (sinew) strengthening	DNA topoisomerase 1	Cancer	0.0144
			Sodium/glucose cotransporter 1	Glucose homeostasis	0.0342
	Tonifying and replenishing	Qi tonifying	Steryl-sulfatase	Immunomodulation	0.0370
			Tyrosine-protein phosphatase nonreceptor type 2	Immunomodulation	0.0174
			Sodium/glucose cotransporter 2	Glucose homeostasis	0.0282
			Sodium/glucose cotransporter 1	Glucose homeostasis	0.0321
	Cough suppressing and panting-calming	Clearing and heat phlegm resolving	Tyrosine-protein phosphatase nonreceptor type 2	Immunomodulation	0.0112
			DNA topoisomerase 1	Cancer	0.0236
	Tranquilizing	Heat nourishing tranquilizing	Testosterone 17-beta-dehydrogenase 3	Reproduction system	0.0341
			Peptidyl-prolyl cis-trans isomerase FKBP1A	Immunomodulation	0.0077
			Tyrosine-protein phosphatase nonreceptor type 2	Immunomodulation	0.0167
			Glutamate carboxypeptidase 2	CNS	0.0209
Cluster VII	Heat-clearing medicinal	Heat-clearing and blood cooling	Protein kinase C beta type	Immunomodulation	0.0100
			DNA topoisomerase 1	Cancer	0.0123
			Sodium/glucose cotransporter 2	Glucose homeostasis	0.0137
	Hemostatic medicinal	Stasis resolving hemostatic	Tyrosine-protein phosphatase nonreceptor type 2	Immunomodulation	0.0089
			Protein kinase C eta type	Immunomodulation	0.0182
			Protein kinase C gamma type	Immunomodulation	0.0209
	Tonifying and replenishing	Blood tonifying	Tyrosine-protein phosphatase nonreceptor type 2	Immunomodulation	0.0140
			Protein kinase C beta type	Immunomodulation	0.0230
			Protein kinase C eta type	Immunomodulation	0.0240
Cluster XII	Parasite destroying, dampness eliminating, and itchiness relieving		Dihydrofolate reductase	Cancer, bacterial infection	0.0532
			DNA-dependent protein kinase catalytic subunit	Cancer	0.0644
			Tumour necrosis factor	Cancer, bacterial infection	0.0687

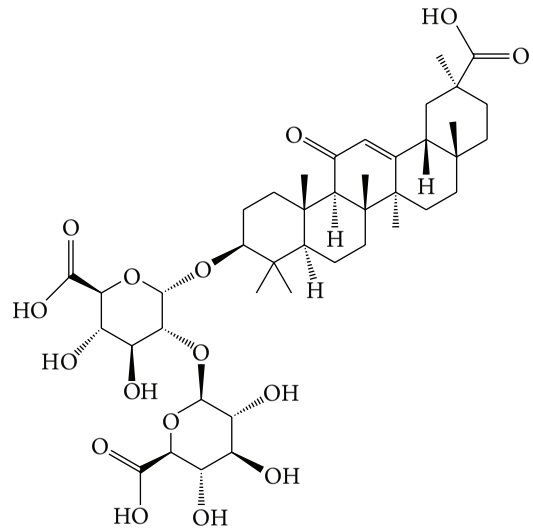
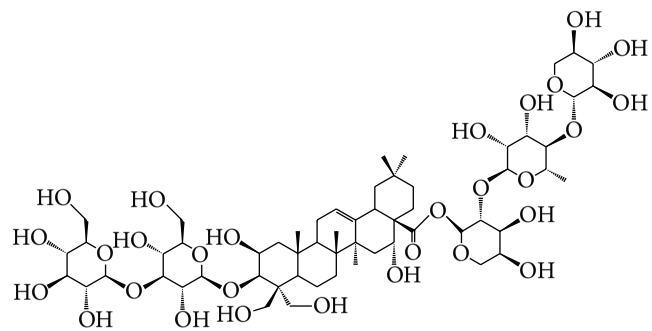
In the "wind-dampness dispelling, bone strengthening" (WD-bone) subclass, the top three enriched pathways are mineral absorption (ES = 0, AR = 0.0342), carbohydrate digestion and absorption (ES = 0, AR = 0.1427), and bile secretion (ES = 0, AR = 0.2050). To link the mineral absorption pathway to the indication of the subclass, it has been reported that minerals such as calcium are important for bone remodelling to strengthen the bone [97]. In the second pathway, carbohydrate digestion and absorption are associated with the indication of the subclass. Here, a study shows that the low carbohydrate-high fat (LH-FC) diets reduced the bone growth, bone structures, and mechanical properties

Table 3: Compounds from clusters VII, X, and XII, which were found to modulate the top three enriched targets of the therapeutic action classes from the respective clusters. The phytochemical triterpenoids were frequently found in different herbs. The predicted targets were supported by published literature based on at least one of the herbs, suggesting that the other herbs that also contain the same compound could contribute to related therapeutic effects.

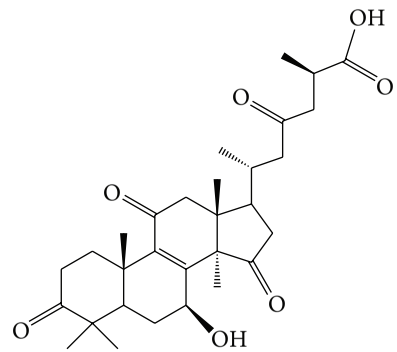
(a) Cluster X

Therapeutic action (sub)class	Herb	Compound	Predicted targets (based on the top three enriched targets)	Literature support
Wind-dampness dispelling, bone strengthening (WD-bone)	*Acanthopanax gracilistylus*	 Acankoreoside A	TOPOI SGLT1	The herb's extract inhibited cell proliferation of several types of cancer cells [72].

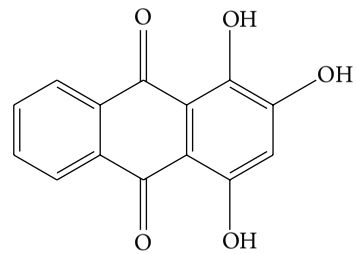
(a) Continued.

Therapeutic action (sub)class	Herb	Compound	Predicted targets (based on the top three enriched targets)	Literature support
Tonifying and replenishing, qi tonifying (TR-qi)	*Glycyrrhiza glabra* *Glycyrrhiza uralensis*	 Glycyrrhizin	PTPN2	An *in vivo* study suggested the compound ameliorates all established chronic histopathologic changes of lung in the mouse model of asthma [78].
Cough suppressing and panting-calming, clearing and heat phlegm resolving (CSPC-heat)	*Platycodon grandiflorum*	 Platycoside A	PTPN2 TOPO1 17-beta-HSD 3	There is no supporting literature.

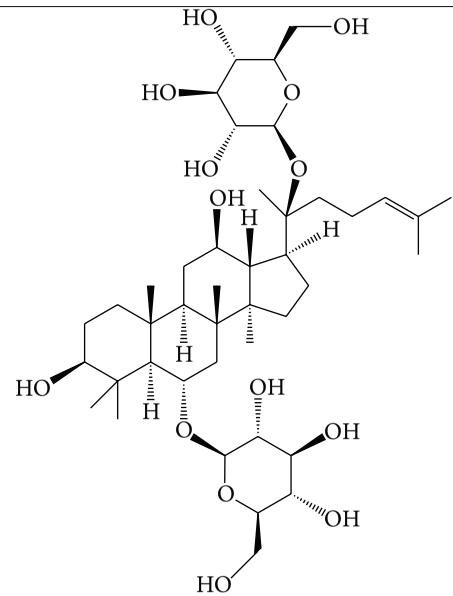
(a) Continued.

Therapeutic action (sub)class	Herb	Compound	Predicted targets (based on the top three enriched targets)	Literature support
Tranquilizing, heart nourishing tranquilizing (Tranquil-heart)	*Ganoderma lucidum*	 Ganoderic acid C1	CGPII	Triterpenoids of the herb exhibited nerve growth factor or brain-derived neurotrophic factor activities *in vitro*, which has the therapeutic potential in neurodegenerative diseases [81].

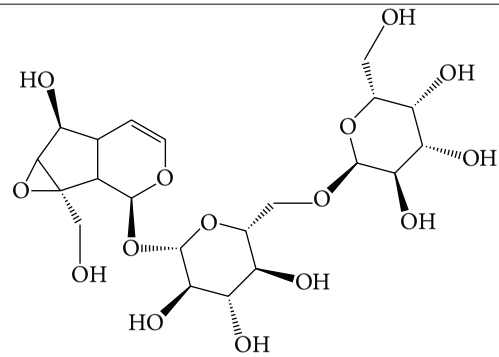
(b) Cluster VII

Therapeutic action (sub)class	Herb	Compound	Predicted targets (based on the top three enriched targets)	Literature support
Hemostatic, stasis resolving (Hemo-stasis)	*Rubia cordifolia*	 Purpurin	PKC-β PKC-ε	The ethanol extract showed wound healing activities in mice, which from histological evaluations indicated marked infiltration of the inflammatory cells, increased blood vessel formation, and enhanced proliferation of cells [85].

(b) Continued.

Therapeutic action (sub)class	Herb	Compound	Predicted targets (based on the top three enriched targets)	Literature support
Tonifying and replenishing, blood (TR-blood)	*Panax notoginseng*	 Ginsenoside Rg1	PTPN2 PKC-η	The compound ameliorated liver damage and suppressed proinflammatory cytokines secretion in concanavalin A-induced hepatitis in mice [86].

(b) Continued.

Therapeutic action (sub)class	Herb	Compound	Predicted targets (based on the top three enriched targets)	Literature support
Heat-clearing, blood cooling (HC-blood cool)	*Rehmannia glutinosa*	 Rehmannioside A	SGLT2	The stachyose extract from the herb showed a significant hypoglycaemic effect in diabetic mice [87].

(c) Cluster XII

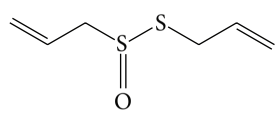
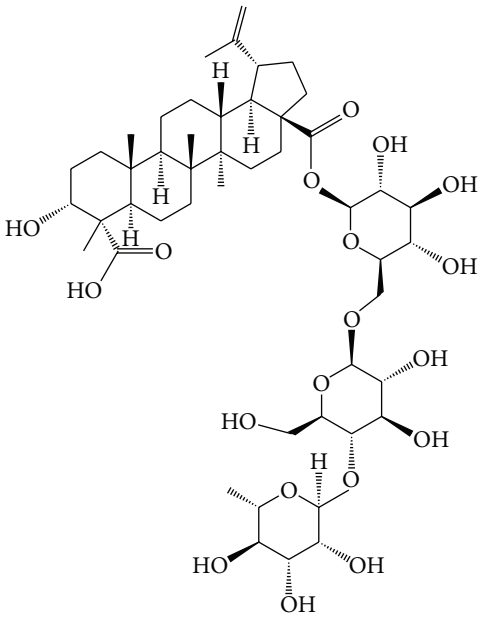
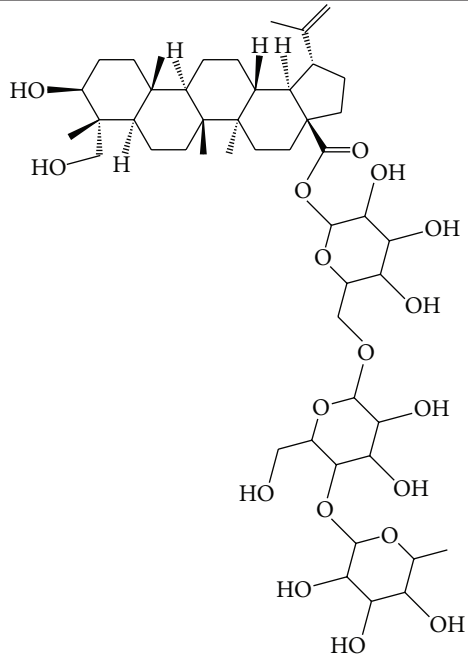
Therapeutic action (sub)class	Herb	Compound	Predicted targets (based on the top three enriched targets)	Literature support
Parasite destroying, dampness eliminating, and itchiness relieving (Parasite)	*Allium sativum*	 Allicin	DHFR	The extract of *Allium sativum* was found to inhibit the synthesis of proteins, nucleic acids, and lipids in *Candida albicans* where the major component of the herb was allicin [94].

Table 4: Compound from CHEMBL database that is most similar to acankoreoside A (Table 3) and its activity profile.

Reference compound	Closest similarity	Reported activity profile
 Acankoreoside A	 CHEMBL 1986122	The GI_{50} value of CHEMBL 1986122 was 43.05 nM in CCRF-CEF leukaemia cell line growth inhibition assay. *Tanimoto coefficient* = 0.91

in mice [98]. However, no study can be found to link bile secretion to the indication of the subclass.

"Tonifying and replenishing, qi tonifying" (TR-qi) subclass has mineral absorption (ES = 0, AR = 0.0321), carbohydrate digestion and absorption (ES = 0, AR = 0.1251), and bile secretion (ES = 0, AR = 0.2678) in the top three enriched pathways. Minerals play vital roles in maintaining the cell functions to optimise health and prevent diseases. The first enriched pathway, mineral absorption, the process of recycling iron in erythrocytes, is materialized in splenic macrophages in the red pulp [99] which agrees with the indication that the subclass which is qi tonifying is related to the maintenance of blood flow within the vessels which is implicated directly from spleen's activity [100]. In the second pathway, no study can be found to link carbohydrate digestion and absorption to the indication of the subclass. Relating to the third enriched pathway, bile secretion, it has been reported that the major components of bile acid, chenodeoxycholic acid and glycochenodeoxycholic acid, can induce cyclooxygenase 2 expression and cell proliferation in esophageal squamous cells, suggesting that bile acids may contribute to the inflammation and mucosal thickening [101]. Another paper demonstrates that bile acids may induce airway fibrosis through the production of TGF-β1 and fibroblast proliferation [102]. These findings can be associated with the indication of the subclass, which is related to lung deficiency [12].

The top enriched pathways of "cough suppressing and panting-calming, clearing and heat phlegm resolving" (CSPC-heat) subclass are minerals absorption (ES = 0, AR = 0.0380), carbohydrate digestion and absorption (ES = 0, AR = 0.1572), and bile secretion (ES = 0, AR = 0.1958). Relating to the first enriched pathway, mineral absorption, the abnormal distribution of trace elements such as zinc, selenium, and copper has been reported to aggravate oxidative damage and inflammation in the airways and subsequently decreased the lung's function in asthmatic patient [103], where an asthmatic condition is described as retention of heat phlegm in the lung [104]. In the second enriched pathway, carbohydrate digestion and absorption, the ingestion of carbohydrate is reported to attenuate the migration of T-lymphocytes to the bronchial epithelial cell line when it is infected with the common respiratory pathogen human rhinovirus during strenuous exercise [105]. This effect agrees with the phenotype of the subclass which is to dissolve phlegm upon infection in the lung [12, 74]. However, no strong evidence can be found to link bile secretion to the indication of the subclass.

The "tranquilizing, heart nourishing tranquilizing" (Tranquil-heart) subclass lists mineral absorption (ES = 0, AR = 0.0387), carbohydrate digestion and absorption (ES = 0, AR = 0.1163), and terpenoid backbone biosynthesis (ES = 0, AR = 0.2166). In the first enriched pathway, mineral absorption, it is reported that selenium plays an important role in the brain where its deficiency is implicated in senility and Alzheimer's disease [106]. In the third enriched pathway, terpenoid backbone biosynthesis includes mevalonate and nonmevalonate pathway. The mevalonate pathway is a pathway implicated in cholesterol biosynthesis in the brain and deficiencies in cholesterol metabolism can lead to diseases of the central nervous system (CNS) [107]. Both findings agree with the indication of the subclass, which is the pharmacological effects that are on the central nervous system [12]. Although there is no strong evidence to support

the second enriched pathway, carbohydrate digestion and absorption, the two enriched pathways can be linked to the indication of the subclass. In short, in many instances, the top three enriched pathways can be associated with the indications of the (sub)classes to explain their MOAs.

3.4.2. Cluster with Three Therapeutic Action (Sub)classes (Cluster VII). The top three enriched pathways in the cluster can be classified into digestive system (carbohydrate digestion and absorption, bile secretion, and mineral absorption), cellular communication (tight junction), and membrane transport (ABC transporters). The top three enriched targets in "hemostatic, stasis resolving" (Hemo-stasis) subclass are mineral absorption (ES = 0, AR = 0.0267), carbohydrate digestion and absorption (ES = 0, AR = 0.0987), and bile secretion (ES = 0, AR = 0.1872) and "tonifying and replenishing, blood" (TR-blood) subclass lists mineral absorption (ES = 0, AR = 0.0329), carbohydrate digestion and absorption (ES = 0, AR = 0.1055), and tight junction (ES = 0, AR = 0.2053). In "hemostatic, stasis resolving" (Hemo-stasis) subclass, the mineral absorption pathway can be related to zinc deficiency, which has been reported to delay wound healing [108]. In the bile secretion pathway, bile acids are implicated in platelet inhibition by solubilizing the platelets, in which patients with obstructive jaundice were exposed to abnormal hemostasis due to high level of bile acids [109]. Both enriched pathways are related to the indication of the subclass, which is to achieve hemostasis [12], and no study can be found to associate carbohydrate digestion and absorption pathway with the indication of the subclass. In "tonifying and replenishing, blood" (TR-blood), the first enriched pathway, mineral absorption, can be linked to the indication of the subclass which is to strengthen the heart that controls blood vessel [12], when selenium is reported to be protective against cardiovascular disease by contributing to the production of vasodilatory prostacyclin by the endothelium [106]. No study can be found to support the link between the second pathway, carbohydrate digestion and absorption, and the indication of the subclass. In the third enriched pathway, tight junction, the presence of tight junctions in the bile epitheliums acts as barrier to toxic diffusion of bile into hepatic interstitial tissue, which could impair the organ's function [110]. TCM enriched can be associated with the indication of the subclass, which is to strengthen the function of liver [12].

The top three enriched targets for "heat-clearing, blood cooling" (HC-blood cool) subclass are mineral absorption (ES = 0, AR = 0.0276), carbohydrate digestion and absorption (ES = 0, AR = 0.0753), and ABC transporter (ES = 0, AR = 0.0769). The mineral absorption pathway can be linked to the subclass when a decrease in intracellular magnesium concentration is implicated in type 2 diabetes [111] and the disease, according to TCM, is described as the deficiency of body liquid due to heat syndrome [112]. Similar to previous subclass, no study can be found to link the second enriched pathway, the carbohydrate digestion and absorption, to the indication of the subclass. In the third pathway, ABC transporter, it is reported that the high expression of ABCG5 and ABG8 in hypercholesterolemic condition of the heart is involved in cardiovascular protection by lowering plasma cholesterol level [113]. The pathway can be linked to the indication of the subclass when the herbs are described to act on liver and heart [12]. All in all, in many cases, the top three enriched pathways can be linked to the indications of the subclasses to explain their MOAs.

3.4.3. Cluster with One Therapeutic Action (Sub)class (Cluster XII). In the "parasite destroying, dampness eliminating, and itchiness relieving" (Parasite) class, the top three enriched pathways are steroid biosynthesis (ES = 0, AR = 0.2530), glycerophospholipid metabolism (ES = 0, AR = 0.2941), and collecting duct acid secretion (ES = 0, AR = 0.3420). The first two pathways, steroid biosynthesis and glycerophospholipid metabolism, are part of lipid metabolism, and many reports have suggested their link to immune response. For instance, the steroid biosynthesis is downregulated by interferon type I upon viral infection [114] and the chlamydia exploits the nutrient-rich host cell cytosol by trafficking the glycerophospholipid from the host cell for survival [115]. In relation to the collecting duct acid secretion pathway, the bacterial infection in the kidney is reported to affect the collecting duct acid secretion because the presence of lipopolysaccharide (LPS) of the bacteria inhibits HCO_3^- absorption [116]. These pathways agree with the functions of the subclass, which is to kill and expel parasites [12, 74]. Altogether, the top enriched pathways can be linked to the indication of the class to explain the MOAs. Table 5 summarises the top three enriched pathways of the therapeutic action (sub)classes.

The above analysis can be summarized into three observations as follows. Firstly, a type of therapeutic action (sub)class is implicated in different enriched pathways and each pathway is involved in a different disease; secondly, more than one therapeutic action (sub)class is implicated in a pathway that is involved in multiple diseases; and thirdly, more than one therapeutic action (sub)class is implicated in different pathways but involved in only one type of a disease. The first observation is indicated, in many cases, by the (sub)classes in any of the three clusters such as bile secretion and mineral absorption in "tonifying and replenishing" (TR-qi). The top three enriched pathways in a therapeutic action (sub)class are implicated in different pathways and diseases. The modulation of one pathway to a disease could provide a better insight into the (sub)class MOA in the biological system. In the second observation, the mineral absorption in cluster VII and cluster X and carbohydrate digestion and absorption in cluster VII are implicated in various types of diseases. The redundancy of a pathway in the pathogenesis of various diseases implies that a pathway could serve multiple purposes; for instance, the mitogen-activated protein kinase (MAPK) signalling pathway was implicated in inflammation, cancer, cardiovascular dysfunction, and Alzheimer's disease [117]. The third observation shows that different pathways from different therapeutic action subclasses in a cluster are implicated in a type of physiological function, such as the ABC transporter and mineral absorption from "heat-clearing, blood cooling" (HC-blood cool) and "tonifying and replenishing, blood" (TR-blood), respectively, which is implicated in physiological function of the heart. The different pathways for a particular physiological function explain that

Table 5: The top three enriched pathways in clusters VII, X, and XII. It can be seen that, in many cases, similar pathways appear in the top three enriched targets regardless of clusters and subclasses. Estimation Score = 0, for all top three enriched pathways.

	TCM therapeutic action class	TCM therapeutic action subclass	Top three enriched pathways	Average Score
Cluster X	Wind-dampness dispelling	Bone (sinew) strengthening	hsa04978, mineral absorption	0.0342
			hsa04973, carbohydrate digestion and absorption	0.1427
			hsa04976, bile secretion	0.2050
	Tonifying and replenishing	Qi tonifying	hsa04978, mineral absorption	0.0321
			hsa04973, carbohydrate digestion and absorption	0.1251
			hsa04976, bile secretion	0.2678
	Cough suppressing and panting-calming	Clearing and heat phlegm resolving	hsa04978, mineral absorption	0.0380
			hsa04973, carbohydrate digestion and absorption	0.1572
			hsa04976, bile secretion	0.1958
	Tranquilizing	Heat nourishing tranquilizing	hsa04978, mineral absorption	0.0387
			hsa04973, carbohydrate digestion and absorption	0.1163
			hsa00900, terpenoid backbone biosynthesis	0.2166
Cluster VII	Heat-clearing medicinal	Heat-clearing and blood cooling	hsa04978, mineral absorption	0.0276
			hsa04973, carbohydrate digestion and absorption	0.0753
			hsa02010, ABC transporters	0.0769
	Hemostatic medicinal	Stasis resolving hemostatic	hsa04978, mineral absorption	0.0267
			hsa04973, carbohydrate digestion and absorption	0.0987
			hsa04976, bile secretion	0.1872
	Tonifying and replenishing	Blood tonifying	hsa04978, mineral absorption	0.0329
			hsa04973, carbohydrate digestion and absorption	0.1055
			hsa04530, tight junction	0.2053
Cluster XII	Parasite destroying, dampness eliminating, and itchiness relieving		hsa00100, steroid biosynthesis	0.253
			hsa00564, glycerophospholipid metabolism	0.294
			hsa04966, collecting duct acid secretion	0.342

a pathway does not function alone in the manifestation of a disease but through the interactions of multiple pathways [118], which account for the different clinical symptoms. To put it briefly, the pathways annotation not only is beneficial to suggest the MOAs of the (sub)classes based on their indications, but also shows that one pathway could have manifold functions and multiple pathways contribute to the pathogenesis of a disease. From the TCM perspective, it is suggested that the involvement of multiple pathways in the pathogenesis of a disease explains the complex TCM formulation, which consists of a set of herbs from various therapeutic action (sub)classes.

3.5. Comparison of Bioactivity Spaces of Clusters. In this part of study, the aim was to investigate the differences of bioactivity spaces among all clusters by classifying all the enriched targets in the cluster to their respective protein families. The 181 enriched targets were classified into 59 protein families. Out of 59 protein families, five protein families were frequently annotated in all clusters, which are G-protein coupled receptor (GPCR), protein kinase, nuclear hormone receptor, carbonic anhydrase, and cytochrome P450. The heatmap in Figure 3 compares the five major protein families that were annotated based on the enriched targets in each cluster. The more saturated colour represents the more significant protein family across all clusters. The numbers of protein families in the clusters were normalised because the distribution of enriched targets was not consistent because numbers of therapeutic action (sub)classes per cluster were different.

What can be seen from the graph is that GPCRs and protein kinases are the most highly classified protein families in almost all clusters. The results were expected as both of the families are the two largest protein families involved in many physiological processes [119, 120], thus explaining why these two protein families were observed to be significant in many of the clusters. In addition, the diversity of the compounds in the (sub)classes also contributed to the prediction of enriched targets from these protein families. At least four protein families were significant across all clusters, except for clusters VII and X in cytochrome P450. The numbers of enriched targets in the cluster were among the lowest across all clusters, and cytochrome P450 was less classified. Nuclear hormone receptors were found to be annotated in all clusters and the frequent predictions of the nuclear hormone receptor family can be explained by the presence of naturally occurring steroids in natural compounds [121]. For the remaining two major protein families, the cytochrome P450 family was also expected to be modulated by most of the subclasses because

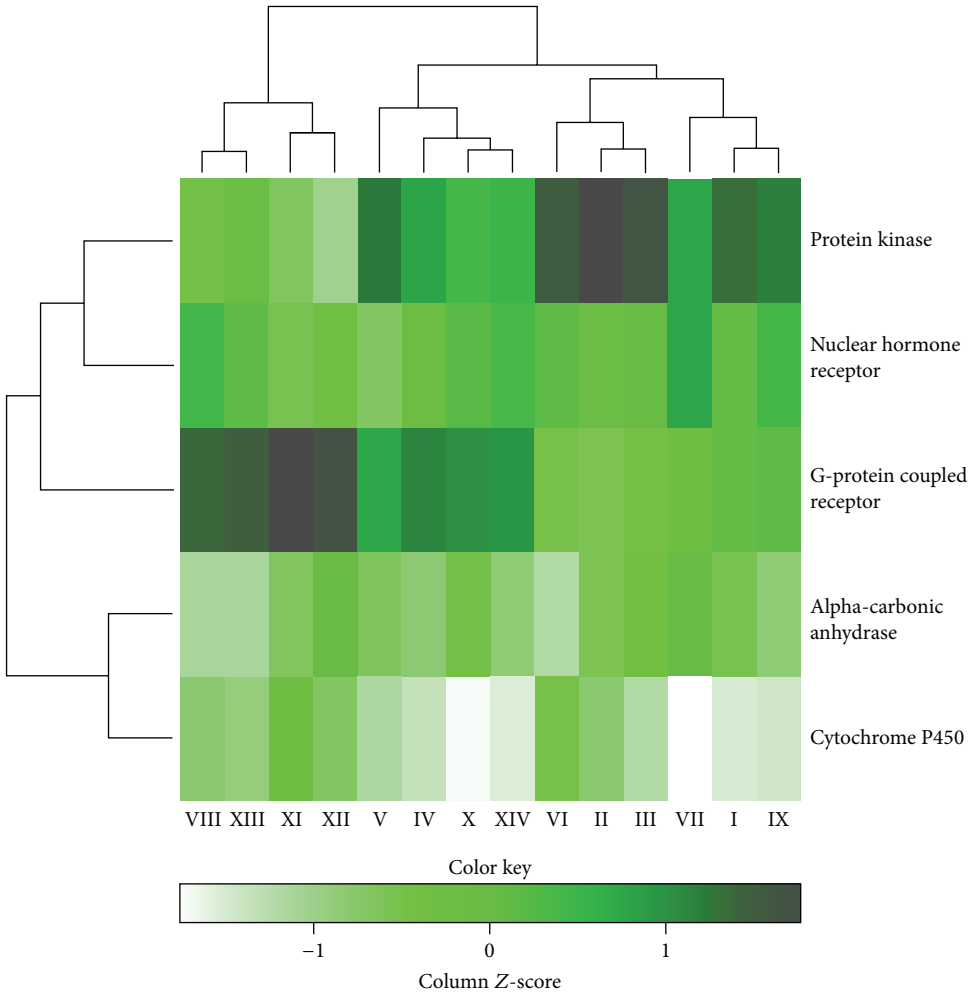


Figure 3: The heatmap compares the five major protein families that were annotated based on the enriched targets in each cluster, which were normalized. The more saturated colour across clusters represents the more significant protein family. GPCR and protein kinase are observed to be significant protein families in almost all clusters. It appears that all the protein families are heavily implicated in the biological processes such as cell regulation, sensory system, and steroid metabolism. The significant cluster for a particular protein family can be suggested to be further explored for a disease with the known protein family such as protein kinase in cancer.

this protein family plays an important role in the degradation of structurally rather diverse exogenous compounds [122]. The carbonic anhydrase family, which is a ubiquitous enzyme, is involved in the interconversion between carbon dioxide and the bicarbonate ion that is important for many physiological processes [123]. All in all, five major protein families were observed to frequently occur in most of the clusters and were heavily implicated in the biological processes such as cell regulation, sensory system, and steroid metabolism. This analysis has allowed the discovery of the bioactivity spaces connection between subclasses in TCM based on the sets of enriched targets from our *in silico* target prediction in which the compounds from the significant cluster can be further explored for diseases associated with the protein family such as cancer and protein kinase family.

3.6. Comparison of Pathways Annotation of Clusters. In this part of the study, we aimed to investigate the differences of pathway motifs among all clusters by classifying the enriched pathways according to KEGG ortholog. The 99 enriched pathways were classified to 33 pathway motifs, which were almost half the number of total pathway motifs available in KEGG. The major pathway motifs from the classification were infectious diseases, digestive system, immune system, signal transduction, lipid metabolism, cancer, and cellular communication. The heatmap in Figure 4 compares the seven major pathway motifs that were annotated based on the enriched pathways in each cluster. The more saturated colour across clusters represents the more significant pathway motif. The numbers of pathway motifs in the clusters were normalised due to the differences in numbers of the pathways that were enriched among all (sub)classes

In Figure 4, the digestive system is consistently classified in all clusters, where, in many cases, the plots' colours are more saturated compared to other pathway motifs. The digestive system includes the digestion and absorption of macro- and micronutrients and shows that the majority of the enriched pathways are bile secretion, pancreatic secretion,

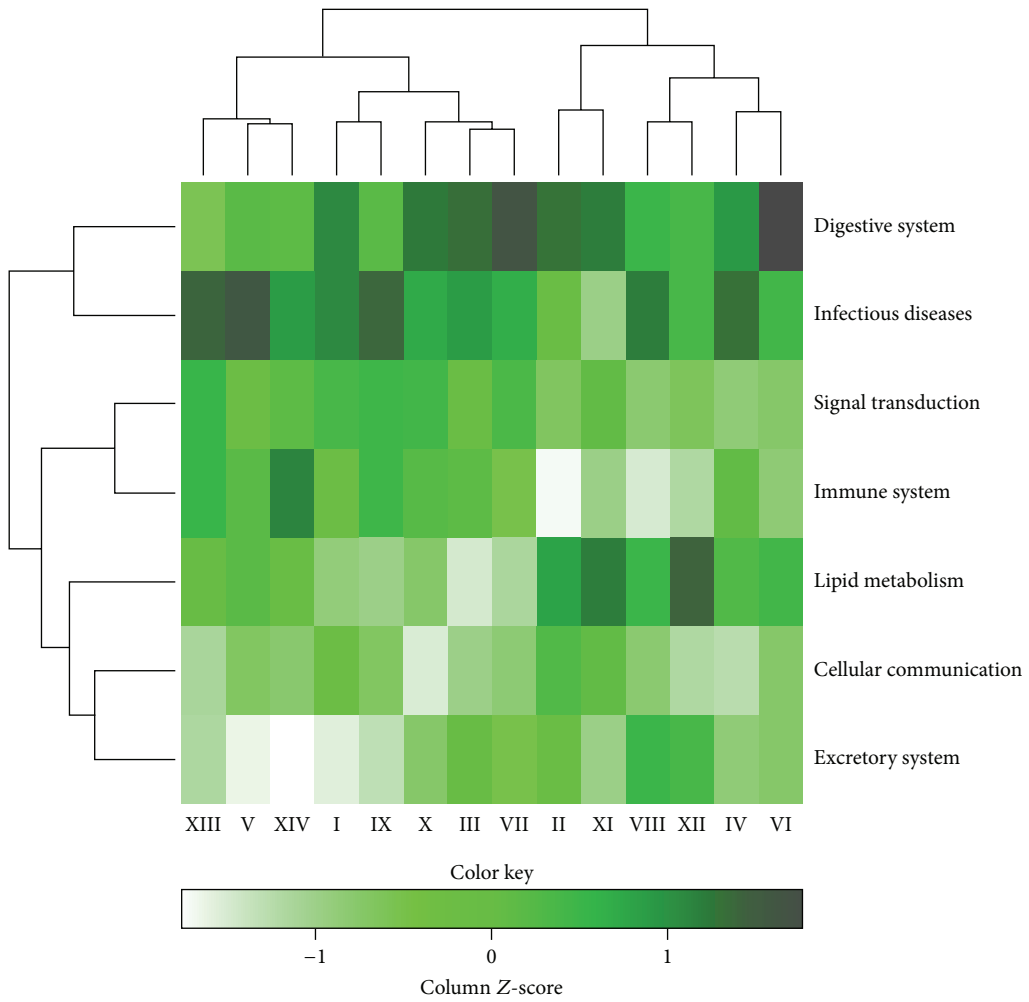


Figure 4: The heatmap compares the seven major pathway motifs that were annotated based on the number of enriched pathways in each cluster, which were normalized. The more saturated colour across clusters represents the more significant pathway motif. The significant cluster for a particular pathway can be suggested to be further explored for a disease with the known pathway motif such as digestive system in liver disease.

and gastric acid secretion. The bile secretion controls the cholesterol homeostasis by routing the elimination of cholesterol, in addition to harmful exogenous lipophilic substance [124]. In addition to digestive system, infectious diseases, signal transduction, and lipid metabolism are significant pathway motifs in many of the clusters. The significance of infectious diseases can be deduced from frequently predicted immunomodulatory targets as the pathogenic factors are described to attack weakened immune system [125]. The signal transduction pathway motif can be contributed from the classification of one of the major protein families, GPCR, which translates the extracellular signals for the downstream effectors that produce a physiological response in a target cell [126]. Many Chinese medicines have been reported to have lipid regulating effects by influencing the intestinal lipid absorption and lipid metabolism, to name a few [127]. Also, one of the highly observed enriched pathways in the lipid metabolism pathway motif was steroid hormone biosynthesis, which the cytochrome P450 protein family is involved in [128]. It also appears that a few of the pathway motifs are insignificant for some clusters such as immune system for cluster II and excretory system in cluster XIV, which resulted from low number of compounds to influence the targets prediction and pathways annotation. The list of compounds that were annotated for the Chinese medicine in the subclasses might be incomplete to influence the classification of the immune system as well as the remaining missing pathway motifs. In addition, only enriched pathways were used in the classification. Similar to significant clusters for the major protein families, the compounds from significant clusters for the major pathway motifs can be further explored for diseases associated with the protein family such as liver disease and digestive system.

Altogether, this analysis has allowed the discovery of the major pathway motifs in all clusters. Despite having the different therapeutic action (sub)classes in a cluster, in many cases, all the clusters can be classified of having the seven major pathway motifs. The classification of the major pathway motifs was associated with the major protein families analysed in the previous section.

4. Discussion

As shown in this study, the global mapping of relationships between TCM therapeutic action classes and subclasses, based on their putative bioactivity spaces and annotated pathways, provides a novel approach to understand the MOAs of TCM formulations. The classification of the enriched targets and pathways according to the protein families and pathway motifs allows the discovery of the relationship between therapeutic action (sub)classes across clusters, defined in the dendrogram. In the first part of the study, we were able to rationalize the link between the top three enriched targets/pathways and the description of the respective therapeutic action (sub)classes. In the bioactivity space, in many cases, the supporting *in vitro* or *in vivo* studies of the herbs' extracts or isolated compounds were also included to illustrate the MOA of the compounds. Among the three most enriched targets, we observed that immunomodulatory targets such as PTPN2 and PKC family were frequently represented across selected therapeutic action classes and subclasses. In TCM, symptoms are usually regarded as the invasion of pathogenic factors, thus sensitizing the immune system to response, and this might provide a mechanistic link between TCM and Western thinking. The other frequently enriched targets were implicated in glucose homeostasis, namely, SGLT1 and SGLT2, as well as cancer, such as TOPO1. The analysis of the enriched pathways showed that the multiple enriched pathways were implicated for one type of disease, which was cardiovascular diseases, and one enriched pathway could be associated with different pathogenesis of diseases such as carbohydrate digestion and absorption and mineral absorption. This finding is in agreement with the utilisation of different herbs in one TCM formulation, in order to modulate biology in the desired polypharmacological manner. In addition, in many cases, the highly annotated pathway motifs are involved in digestive system such as bile secretion, carbohydrate digestion and absorption, and mineral absorption, which can be linked to a theory of "the foundation of acquired constitution" that includes stomach and spleen as the source of production of qi and blood [129].

The dendrogram which was generated based on the *in silico* target prediction of TCM therapeutic action (sub)classes has enabled the visualisation of their bioactivity space, of which, in most of the clusters, the five major protein families were observed to contribute the bioactivity space. The GPCR and the protein kinase family are the two major protein families that contribute to the diversity of the TCM bioactivity space in all clusters. The major KEGG pathways motifs annotated, such as signal transduction and lipid metabolism, were annotated from major protein families, GPCR and cytochrome P450, found in the previous analysis.

Although this study has successfully explained the link of the enriched targets/pathways to their respective therapeutic action (sub)classes and provided a global overview of their bioactivity space, this study is still limited to targets that are only available in the *in silico* target prediction, while the entire human proteome is much larger than 477 targets. Thus, extending the biological space of the *in silico* target prediction could provide a more comprehensive overview of targets that are involved in the therapeutic effects. Secondly, the chemical space in the chemogenomic database is limited to the version used when the model was developed and in this ChEMBL [49] v.14.0 database, natural compounds only represent approximately 3.85% of the total compounds available [130]. Thus, the limited coverage of chemical space from natural compounds led to zero prediction for targets in the subclass "tranquilizing, settling." Thirdly, as TCM's therapeutic principle works through *Jun-Chen-Zuo-Shi*, our *in silico* target prediction could only predict the putative targets and the top three enriched targets/pathways might not be modulated by the Emperor compounds of the herb, which play the leading role in treating the disease. In this study, similar top three enriched targets/pathways frequently appeared, which did not represent the actual therapeutic actions described in TCM's philosophy per therapeutic action classes because the definition of the (sub)classes could be limited to the English translation [131]. Although compounds in a TCM formula are known to work synergistically, the algorithm is unable to report whether a compound either activates or inhibits the predicted targets, which can be experimentally influenced by ADMET (Adsorption, Distribution, Metabolism, Excretion, and Toxicity) properties [55]. Therefore, only a general justification could be established between the enriched targets/pathways and therapeutic action (sub)classes.

Despite the limitations, our *in silico* target prediction was able to describe the putative MOA of the compounds from the selected therapeutic action (sub)classes by providing target- and pathway-based MOA hypotheses. With the global overview of the bioactivity space of the therapeutic action (sub)classes, we could observe the similarity and the differences between them, which are not apparent from the name given to the (sub)class itself. Hence, this analysis hopefully helps to bridge the gap between TCM and Western medicine a bit further.

Conflict of Interests

All authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the Centre for Molecular Informatics (Andreas Bender, Sonia Liggi, Georgios Drakakis, Fazlin Mohd Fauzi, and Siti Zuraidah Mohamad Zobir) and IPharm (Siti Zuraidah Mohamad Zobir). The authors also would like to thank Daniel Mason and Stephanie Ashenden for proofreading the paper.

References

[1] K. Chan, "Progress in traditional Chinese medicine," *Trends in Pharmacological Sciences*, vol. 16, no. 6, pp. 182–187, 1995.

[2] J. B. Waldram, "The efficacy of traditional medicine: current theoretical and methodological issues," *Medical Anthropology Quarterly*, vol. 14, no. 4, pp. 603–625, 2000.

[3] K. Chen and H. Xu, "The integration of traditional Chinese medicine and Western medicine," *European Review*, vol. 11, no. 2, pp. 225–235, 2003.

[4] E. Chan, M. Tan, J. Xin, S. Sudarsanam, and D. E. Johnson, "Interactions between traditional Chinese medicines and Western therapeutics," *Current Opinion in Drug Discovery & Development*, vol. 13, no. 1, pp. 50–65, 2010.

[5] Y. Wang and A. Xu, "Zheng: a systems biology approach to diagnosis and treatments," *Science*, vol. 346, no. 6216, pp. S13–S15, 2014.

[6] X. Zhao, X. Zheng, T.-P. Fan, Z. Li, Y. Zhang, and J. Zheng, "A novel drug discovery strategy inspired by traditional medicine philosophies," *Science*, vol. 347, no. 6219, pp. S38–S40, 2015.

[7] R. H. Liu, R. S. Runyon, Y. C. Wang, S. G. Oliver, T.-P. Fan, and W. D. Zhang, "Deciphering ancient combinatorial formulas: the Shexiang Baoxin pill," *Science*, vol. 347, no. 6219, pp. S40–S42, 2015.

[8] W. Lam, S.-H. Liu, Z. Jiang, and Y.-C. Cheng, "Lessons from the development of the traditional Chinese medicine formula PHY906," *Science*, vol. 347, no. 6219, pp. S43–S44, 2015.

[9] H. Sheridan, B. Kopp, L. Krenn, D. Guo, and J. Sendker, "Traditional Chinese herbal medicine preparation: invoking the butterfly effect," *Science*, vol. 350, no. 6262, pp. S64–S66, 2015.

[10] S. Li, "Mapping ancient remedies: applying a network approach to traditional Chinese medicine," *Science*, vol. 350, no. 6262, pp. S72–S74, 2015.

[11] B. He, C. Lu, M. Wang et al., "Drug discovery in traditional Chinese medicine: from herbal fufang to combinatory drugs," *Science*, vol. 350, no. 6262, pp. S74–S76, 2015.

[12] J. K. Chen and T. T. Chen, *Chinese Medical Herbalogy and Pharmacology*, Art of Medicine Press, Industry, Calif, USA, 2004.

[13] C. Y. Chen, "TCM Database@Taiwan: the world's largest traditional Chinese medicine database for drug screening *in silico*," *PLoS ONE*, vol. 6, no. 1, Article ID e15939, 2011.

[14] J. Zhao, P. Jiang, and W. Zhang, "Molecular networks for the study of TCM pharmacology," *Briefings in Bioinformatics*, vol. 11, no. 4, pp. 417–430, 2010.

[15] M. Zhao, Q. Zhou, W. Ma, and D.-Q. Wei, "Exploring the ligand-protein networks in traditional chinese medicine: current databases, methods, and applications," *Evidence-based Complementary and Alternative Medicine*, vol. 2013, Article ID 806072, 15 pages, 2013.

[16] J. Qiu, "A culture in the balance," *Nature*, vol. 448, no. 7150, pp. 126–128, 2007.

[17] L. Wang, G.-B. Zhou, P. Liu et al., "Dissection of mechanisms of Chinese medicinal formula Realgar-*Indigo naturalis* as an effective treatment for promyelocytic leukemia," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 12, pp. 4826–4831, 2008.

[18] T.-P. Fan, J.-C. Yeh, K. W. Leung, P. Y. K. Yue, and R. N. S. Wong, "Angiogenesis: from plants to blood vessels," *Trends in Pharmacological Sciences*, vol. 27, no. 6, pp. 297–309, 2006.

[19] S. Ma, C. Feng, X. Zhang et al., "The multi-target capabilities of the compounds in a TCM used to treat sepsis and their *in silico* pharmacology," *Complementary Therapies in Medicine*, vol. 21, no. 1, pp. 35–41, 2013.

[20] Z. Xu, "Modernization: one step at a time," *Nature*, vol. 480, no. 7378, pp. S90–S92, 2011.

[21] T. W. Corson and C. M. Crews, "Molecular understanding and modern application of traditional medicines: triumphs and trials," *Cell*, vol. 130, no. 5, pp. 769–774, 2007.

[22] Y. Cai, Q. Luo, M. Sun, and H. Corke, "Antioxidant activity and phenolic compounds of 112 traditional Chinese medicinal plants associated with anticancer," *Life Sciences*, vol. 74, no. 17, pp. 2157–2184, 2004.

[23] Q. Wang, H. Kuang, Y. Su et al., "Naturally derived anti-inflammatory compounds from Chinese medicinal plants," *Journal of Ethnopharmacology*, vol. 146, no. 1, pp. 9–39, 2013.

[24] Y. J. Kang, "Herbogenomics: from traditional Chinese medicine to novel therapeutics," *Experimental Biology and Medicine*, vol. 233, no. 9, pp. 1059–1065, 2008.

[25] K. Birchall, V. J. Gillet, G. Harper, and S. D. Pickett, "Training similarity measures for specific activities: application to reduced graphs," *Journal of Chemical Information and Modeling*, vol. 46, no. 2, pp. 577–586, 2006.

[26] A. F. Fliri, W. T. Loging, P. F. Thadeio, and R. A. Volkmann, "Biological spectra analysis: linking biological activity profiles to molecular structure," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 2, pp. 261–266, 2005.

[27] J. Klekota, E. Brauner, F. P. Roth, and S. L. Schreiber, "Using high-throughput screening data to discriminate compounds with single-target effects from those with side effects," *Journal of Chemical Information and Modeling*, vol. 46, no. 4, pp. 1549–1562, 2006.

[28] Nidhi, M. Glick, J. W. Davies, and J. L. Jenkins, "Prediction of biological targets for compounds using multiple-category bayesian models trained on chemogenomics databases," *Journal of Chemical Information and Modeling*, vol. 46, no. 3, pp. 1124–1133, 2006.

[29] A. Lagunin, A. Stepanchikova, D. Filimonov, and V. Poroikov, "PASS: prediction of activity spectra for biologically active substances," *Bioinformatics*, vol. 16, no. 8, pp. 747–748, 2000.

[30] A. Bender, D. W. Young, J. L. Jenkins et al., "Chemogenomic data analysis: prediction of small-molecule targets and the advent of biological fingerprints," *Combinatorial Chemistry & High Throughput Screening*, vol. 10, no. 8, pp. 719–731, 2007.

[31] N. Paul, E. Kellenberger, G. Bret, P. Müller, and D. Rognan, "Recovering the true targets of specific ligands by virtual screening of the protein data bank," *Proteins: Structure, Function and Genetics*, vol. 54, no. 4, pp. 671–680, 2004.

[32] X. Chen, C. Y. Ung, and Y. Chen, "Can an *in silico* drug-target search method be used to probe potential mechanisms of medicinal plant ingredients?" *Natural Product Reports*, vol. 20, no. 4, pp. 432–444, 2003.

[33] T. M. Ehrman, D. J. Barlow, and P. J. Hylands, "In silico search for multi-target anti-inflammatories in Chinese herbs and formulas," *Bioorganic & Medicinal Chemistry*, vol. 18, no. 6, pp. 2204–2218, 2010.

[34] S. Zhang, W. Lu, X. Liu et al., "Fast and effective identification of the bioactive compounds and their targets from medicinal plants via computational chemical biology approach," *Medicinal Chemistry Communication*, vol. 2, no. 6, pp. 471–477, 2011.

[35] J. L. Jenkins, A. Bender, and J. W. Davies, "*In silico* target fishing: predicting biological targets from chemical structure," *Drug Discovery Today: Technologies*, vol. 3, no. 4, pp. 413–421, 2006.

[36] J. Xu and A. Hagler, "Chemoinformatics and drug discovery," *Molecules*, vol. 7, no. 8, pp. 566–600, 2002.

[37] P. Willett, J. M. Barnard, and G. M. Downs, "Chemical similarity searching," *Journal of Chemical Information and Computer Sciences*, vol. 38, no. 6, pp. 983–996, 1998.

[38] A. Koutsoukas, B. Simms, J. Kirchmair et al., "From *in silico* target prediction to multi-target drug design: current databases, methods and applications," *Journal of Proteomics*, vol. 74, no. 12, pp. 2554–2574, 2011.

[39] V. Poroikov, D. Filimonov, A. Lagunin, T. Gloriozova, and A. Zakharov, "PASS: identification of probable targets and mechanisms of toxicity," *SAR and QSAR in Environmental Research*, vol. 18, no. 1-2, pp. 101–110, 2007.

[40] M. J. Keiser, B. L. Roth, B. N. Armbruster, P. Ernsberger, J. J. Irwin, and B. K. Shoichet, "Relating protein pharmacology by ligand chemistry," *Nature Biotechnology*, vol. 25, no. 2, pp. 197–206, 2007.

[41] H. Yu, J. Chen, X. Xu et al., "A systematic prediction of multiple drug-target interactions from chemical, genomic, and pharmacological data," *PLoS ONE*, vol. 7, no. 5, Article ID e37608, 2012.

[42] B. Li, X. Xu, X. Wang et al., "A systems biology approach to understanding the mechanisms of action of Chinese herbs for treatment of cardiovascular disease," *International Journal of Molecular Sciences*, vol. 13, no. 10, pp. 13501–13520, 2012.

[43] B. Li, W. Tao, C. Zheng et al., "Systems pharmacology-based approach for dissecting the addition and subtraction theory of traditional Chinese medicine: an example using Xiao-Chaihu-Decoction and Da-Chaihu-Decoction," *Computers in Biology and Medicine*, vol. 53, pp. 19–29, 2014.

[44] F. Mohd Fauzi, A. Koutsoukas, R. Lowe et al., "Chemogenomics approaches to rationalizing the mode-of-action of traditional chinese and ayurvedic medicines," *Journal of Chemical Information and Modeling*, vol. 53, no. 3, pp. 661–673, 2013.

[45] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.

[46] A. A. Lagunin, R. K. Goel, D. Y. Gawande et al., "Chemo- and bioinformatics resources for *in silico* drug discovery from medicinal plants beyond their traditional use: a critical review," *Natural Product Reports*, vol. 31, no. 11, pp. 1585–1611, 2014.

[47] *Molecular Operating Environemnt (MOE)*, Chemical Computing Group, Montreal, Canada, 2010.

[48] A. Koutsoukas, R. Lowe, Y. Kalantar-Motamedi et al., "*In silico* target predictions: comparing multiclass Naïve Bayes and Parzen-Rosenblatt Window and the definition of a benchmarking dataset for target prediction," *Journal of Chemical Information and Modeling*, vol. 53, no. 8, pp. 1957–1966, 2013.

[49] A. Gaulton, L. J. Bellis, A. P. Bento et al., "ChEMBL: a large-scale bioactivity database for drug discovery," *Nucleic Acids Research*, vol. 40, no. 1, pp. D1100–D1107, 2012.

[50] A. Bender, H. Y. Mussa, R. C. Glen, and S. Reiling, "Similarity searching of chemical databases using atom environment descriptors (MOLPRINT 2D): evaluation of performance," *Journal of Chemical Information and Computer Sciences*, vol. 44, no. 5, pp. 1708–1718, 2004.

[51] P. A. Flach and N. Lachiche, "Naive Bayesian classification of structured data," *Machine Learning*, vol. 57, no. 3, pp. 233–269, 2004.

[52] M. Olah, R. Rad, L. Ostopovici et al., "WOMBAT and WOMBAT-PK: bioactivity databases for lead and drug discovery," in *Chemical Biology*, chapter 13, pp. 760–786, Wiley-VCH, Weinheim, Germany, 2007.

[53] G. Drakakis, A. Koutsoukas, S. Brewerton, M. Bodkin, D. Evans, and A. Bender, "Comparing global and local likelihood score thresholds in multiclass laplacian-modified naive bayes protein target prediction," *Combinatorial Chemistry & High Throughput Screening*, vol. 18, no. 3, pp. 323–330, 2015.

[54] M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe, "Data, information, knowledge and principle: back to metabolism in KEGG," *Nucleic Acids Research*, vol. 42, no. 1, pp. D199–D205, 2014.

[55] S. Liggi, G. Drakakis, A. E. Hendry et al., "Extensions to *in silico* bioactivity predictions using pathway annotations and differential pharmacology analysis: application to *Xenopus laevis* phenotypic readouts," *Molecular Informatics*, vol. 32, no. 11-12, pp. 1009–1024, 2013.

[56] S. Liggi, G. Drakakis, A. Koutsoukas et al., "Extending in silico mechanism-of-action analysis by annotating targets with pathways: application to cellular cytotoxicity readouts," *Future Medicinal Chemistry*, vol. 6, no. 18, pp. 2029–2056, 2014.

[57] E. E. Bolton, Y. Wang, P. A. Thiessen, and S. H. Bryant, "PubChem: integrated platform of small molecules and biological activities," in *Annual Reports in Computational Chemistry*, vol. 4, chapter 12, pp. 217–241, Elsevier, Washington, DC, USA, 2008.

[58] J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad, and R. G. Coleman, "ZINC: a free tool to discover chemistry for biology," *Journal of Chemical Information and Modeling*, vol. 52, no. 7, pp. 1757–1768, 2012.

[59] W. Liu and D. E. Johnson, "Clustering and its application in multi-target prediction," *Current Opinion in Drug Discovery & Development*, vol. 12, no. 1, pp. 98–107, 2009.

[60] R. D. Browne and Y. C. Martin, "Use of structure-activity data to compare structure-based clustering methods and descriptors for use in compound selection," *Journal of Chemical Information and Computer Sciences*, vol. 36, no. 3, pp. 572–584, 1996.

[61] RDC Team, *RStudio: Integrated Development Environment for R*, Rstudio, Boston, Mass, USA, 2010.

[62] J. H. Ward, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.

[63] The UniProt Consortium, "Activities at the universal protein resource (UniProt)," *Nucleic Acids Research*, vol. 42, no. 1, pp. D191–D198, 2014.

[64] A. Bender and R. C. Glen, "Molecular similarity: a key technique in molecular informatics," *Organic and Biomolecular Chemistry*, vol. 2, no. 22, pp. 3204–3218, 2004.

[65] P. Nussbaumer and A. Billich, "Steroid sulfatase inhibitors," *Medicinal Research Reviews*, vol. 24, no. 4, pp. 529–576, 2004.

[66] K. M. Doody, A. Bourdeau, and M. L. Tremblay, "T-cell protein tyrosine phosphatase is a key regulator in immune cell signaling: lessons from the knockout mouse model and implications in human disease," *Immunological Reviews*, vol. 228, no. 1, pp. 325–341, 2009.

[67] A. M. Cameron, J. Frederick, C. Nucifora, E. T. Fung et al., "FKBP12 binds the inositol 1,4,5-trisphosphate receptor at leucine-proline (1400-1401) and anchors calcineurin to this FK506-like domain," *Journal of Biological Chemistry*, vol. 272, no. 44, pp. 27582–27588, 1997.

[68] S. A. Jabbour and B. J. Goldstein, "Sodium glucose co-transporter 2 inhibitors: blocking renal tubular reabsorption of glucose to improve glycaemic control in patients with diabetes," *International Journal of Clinical Practice*, vol. 62, no. 8, pp. 1279–1284, 2008.

[69] J. L. Nitiss, "Investigating the biological functions of DNA topoisomerases in eukaryotic cells," *Biochimica et Biophysica Acta (BBA)—Gene Structure and Expression*, vol. 1400, no. 1–3, pp. 63–81, 1998.

[70] S. Andersson, W. M. Geissler, L. Wu et al., "Molecular genetics and pathophysiology of 17 beta-hydroxysteroid dehydrogenase 3 deficiency," *The Journal of Clinical Endocrinology & Metabolism*, vol. 81, no. 1, pp. 130–136, 1996.

[71] C. Bařinka, C. Rojas, B. Slusher, and M. Pomper, "Glutamate carboxypeptidase II in diagnosis and treatment of neurologic disorders and prostate cancer," *Current Medicinal Chemistry*, vol. 19, no. 6, pp. 856–870, 2012.

[72] B. E. Shan, K. Zeki, T. Sugiura, Y. Yoshida, and U. Yamashita, "Chinese medicinal herb, *Acanthopanax gracilistylus*, extract induces cell cycle arrest of human tumor cells in vitro," *Japanese Journal of Cancer Research*, vol. 91, no. 4, pp. 383–389, 2000.

[73] L. Zhang, A. S. Ravipati, S. R. Koyyalamudi et al., "Antioxidant and anti-inflammatory activities of selected medicinal plants containing phenolic and flavonoid compounds," *Journal of Agricultural and Food Chemistry*, vol. 59, no. 23, pp. 12361–12367, 2011.

[74] *Chinese Herbal Medicine Materia Medica*, Eastlan Press, Seattle, Wash, USA, 3rd edition, 2004.

[75] T. N. Hilton, L. J. Tuttle, K. L. Bohnert, M. J. Mueller, and D. R. Sinacore, "Excessive adipose tissue infiltration in skeletal muscle in individuals with obesity, diabetes mellitus, and peripheral neuropathy: association with performance and function," *Physical Therapy*, vol. 88, no. 11, pp. 1336–1344, 2008.

[76] A. Giordano, M. Calvani, O. Petillo, M. Carteni, M. R. A. B. Melone, and G. Peluso, "Skeletal muscle metabolism in physiology and in cancer disease," *Journal of Cellular Biochemistry*, vol. 90, no. 1, pp. 170–186, 2003.

[77] P. Pouliot, S. Bergeron, A. Marette, and M. Olivier, "The role of protein tyrosine phosphatases in the regulation of allergic asthma: implication of TC-PTP and PTP-1B in the modulation of disease development," *Immunology*, vol. 128, no. 4, pp. 534–542, 2009.

[78] A. B. Hocaoglu, O. Karaman, D. O. Erge et al., "Glycyrrhizin and long-term histopathologic changes in a murine model of asthma," *Current Therapeutic Research*, vol. 72, no. 6, pp. 250–261, 2011.

[79] J. H. Choi, Y. P. Hwang, H. G. Kim et al., "Saponins from the roots of platycodon grandiflorum suppresses TGFβ1-induced epithelial-mesenchymal transition via repression of PI3K/Akt, ERK1/2 and Smad2/3 pathway in human lung carcinoma A549 cells," *Nutrition and Cancer*, vol. 66, no. 1, pp. 140–151, 2014.

[80] S. R. Kim, H. S. Seo, H.-S. Choi et al., "*Trichosanthes kirilowii* ethanol extract and cucurbitacin D inhibit cell growth and induce apoptosis through inhibition of STAT3 activity in breast cancer cells," *Evidence-Based Complementary and Alternative Medicine*, vol. 2013, Article ID 975350, 9 pages, 2013.

[81] X.-Q. Zhang, F. C. F. Ip, D.-M. Zhang et al., "Triterpenoids with neurotrophic activity from *Ganoderma lucidum*," *Natural Product Research*, vol. 25, no. 17, pp. 1607–1613, 2011.

[82] T. Kawakami, Y. Kawakami, and J. Kitaura, "Protein kinase Cβ (PKCβ): nomal functions and dieases," *Journal of Biochemistry*, vol. 132, no. 5, pp. 677–682, 2002.

[83] M. Kashiwagi, M. Ohba, K. Chida, and T. Kuroki, "Protein kinase cη (PKCη): its involvement in keratinocyte differentiation," *Journal of Biochemistry*, vol. 132, no. 6, pp. 853–857, 2002.

[84] N. Saito and Y. Shirai, "Protein kinase Cγ (PKCγ): function of neuron specific isotype," *Journal of Biochemistry*, vol. 132, no. 5, pp. 683–687, 2002.

[85] R. Karodi, M. Jadhav, R. Rub, and A. Bafna, "Evaluation of the wound healing activity of a crude extract of *Rubia cordifolia* L. (Indian madder) in mice," *International Journal of Applied Research in Natural Products*, vol. 2, no. 2, pp. 12–18, 2009.

[86] L. Cao, Y. Zou, J. Zhu, X. Fan, and J. Li, "Ginsenoside Rg1 attenuates concanavalin A-induced hepatitis in mice through inhibition of cytokine secretion and lymphocyte infiltration," *Molecular and Cellular Biochemistry*, vol. 380, no. 1-2, pp. 203–210, 2013.

[87] R. X. Zhang, Z. P. Jia, L. Y. Kong et al., "Stachyose extract from *Rehmannia glutinosa* Libosch. to lower plasma glucose in normal and diabetic rats by oral administration," *Pharmazie*, vol. 59, no. 7, pp. 552–556, 2004.

[88] W. L. Li, H. C. Zheng, J. Bukuru, and N. De Kimpe, "Natural medicines used in the traditional Chinese medical system for therapy of diabetes mellitus," *Journal of Ethnopharmacology*, vol. 92, no. 1, pp. 1–21, 2004.

[89] W. L. W. Hsiao and L. Liu, "The role of traditional Chinese herbal medicines in cancer therapy—from TCM theory to mechanistic insights," *Planta Medica*, vol. 76, no. 11, pp. 1118–1131, 2010.

[90] V. I. Polshakov, "Dihydrofolate reductase: structural aspects of mechanisms of enzyme catalysis and inhibition," *Russian Chemical Bulletin*, vol. 50, no. 10, pp. 1733–1751, 2001.

[91] F.-M. Hsu, S. Zhang, and B. P. C. Chen, "Role of DNA-dependent protein kinase catalytic subunit in cancer development and treatment," *Translational Cancer Research*, vol. 1, no. 1, pp. 22–34, 2012.

[92] P. E. Wakefield, W. D. James, C. P. Samlaska, and M. S. Meltzer, "Tumor necrosis factor," *Journal of the American Academy of Dermatology*, vol. 24, no. 5, pp. 675–685, 1991.

[93] W.-M. Chu, "Tumor necrosis factor," *Cancer Letters*, vol. 328, no. 2, pp. 222–225, 2013.

[94] M. Adetumbi, G. T. Javor, and B. H. S. Lau, "*Allium sativum* (garlic) inhibits lipid synthesis by *Candida albicans*," *Antimicrobial Agents and Chemotherapy*, vol. 30, no. 3, pp. 499–501, 1986.

[95] H.-D. Ma, Y.-R. Deng, Z. Tian, and Z.-X. Lian, "Traditional Chinese medicine and immune regulation," *Clinical Reviews in Allergy and Immunology*, vol. 44, no. 3, pp. 229–241, 2013.

[96] S.-L. Tan and P. J. Parker, "Emerging and diverse roles of protein kinase C in immune cell signalling," *Biochemical Journal*, vol. 376, no. 3, pp. 545–552, 2003.

[97] R. P. Heaney and C. M. Weaver, "Newer perspectives on calcium nutrition and bone quality," *Journal of the American College of Nutrition*, vol. 24, no. 6, pp. 574S–581S, 2005.

[98] M. Bielohuby, M. Matsuura, N. Herbach et al., "Short-term exposure to low-carbohydrate, high-fat diets induces low bone mineral density and reduces bone formation in rats," *Journal of Bone and Mineral Research*, vol. 25, no. 2, pp. 275–284, 2010.

[99] R. E. Mebius and G. Kraal, "Structure and function of the spleen," *Nature Reviews Immunology*, vol. 5, no. 8, pp. 606–616, 2005.

[100] S. Dharmananda, "The Qi Keeps the Blood Within the Vessels," 1998, http://www.itmonline.org/arts/guipitang.htm.

[101] F. Zhang, N. K. Altorki, Y.-C. Wu, R. A. Soslow, K. Subbaramaiah, and A. J. Dannenberg, "Duodenal reflux induces cyclooxygenase-2 in the esophageal mucosa of rats: evidence for involvement of bile acids," *Gastroenterology*, vol. 121, no. 6, pp. 1391–1399, 2001.

[102] D.-W. Perng, K.-T. Chang, K.-C. Su et al., "Exposure of airway epithelium to bile acids associated with gastroesophageal reflux symptoms: a relation to transforming growth factor-α1 production and fibroblast proliferation," *Chest*, vol. 132, no. 5, pp. 1548–1556, 2007.

[103] C.-H. Guo, P.-J. Liu, S. Hsia, C.-J. Chuang, and P.-C. Chen, "Role of certain trace minerals in oxidative stress, inflammation, CD4/CD8 lymphocyte ratios and lung function in asthmatic patients," *Annals of Clinical Biochemistry*, vol. 48, no. 4, pp. 344–351, 2011.

[104] W. Morse, "TCM Treatment of Bronchiol Asthma in Clinical Practice," http://www.acupuncture.com/newsletters/m_feb08/TCM and Asthma.htm.

[105] N. C. Bishop, G. J. Walker, M. Gleeson, F. A. Wallace, and C. R. A. Hewitt, "Human T lymphocyte migration towards the supernatants of human rhinovirus infected airway epithelial cells: influence of exercise and carbohydrate intake," *Exercise Immunology Review*, vol. 15, pp. 127–144, 2009.

[106] M. P. Rayman, "The importance of selenium to human health," *The Lancet*, vol. 356, no. 9225, pp. 233–241, 2000.

[107] M. Orth and S. Bellosta, "Cholesterol: its regulation and role in central nervous system disorders," *Cholesterol*, vol. 2012, Article ID 292598, 19 pages, 2012.

[108] Y. Lim, M. Levy, and T. M. Bray, "Dietary zinc alters early inflammatory responses during cutaneous wound healing in weanling CD-1 mice," *The Journal of Nutrition*, vol. 134, no. 4, pp. 811–816, 2004.

[109] Y.-J. Shiao, J.-C. Chen, C.-N. Wang, and C.-T. Wang, "The mode of action of primary bile salts on human platelets," *Biochimica et Biophysica Acta (BBA)—Biomembranes*, vol. 1146, no. 2, pp. 282–293, 1993.

[110] R. K. Rao and G. Samak, "Bile duct epithelial tight junctions and barrier function," *Tissue Barriers*, vol. 1, no. 4, Article ID e25718, 2013.

[111] D. P. Chaudhary, R. Sharma, and D. D. Bansal, "Implications of magnesium deficiency in type 2 diabetes: a review," *Biological Trace Element Research*, vol. 134, no. 2, pp. 119–129, 2010.

[112] J. Guo, H. Chen, J. Song, J. Wang, L. Zhao, and X. Tong, "Syndrome differentiation of diabetes by the traditional Chinese medicine according to evidence-based medicine and expert consensus opinion," *Evidence-Based Complementary and Alternative Medicine*, vol. 2014, Article ID 492193, 7 pages, 2014.

[113] J. F. Oram and A. M. Vaughan, "ATP-binding cassette cholesterol transporters and cardiovascular disease," *Circulation Research*, vol. 99, no. 10, pp. 1031–1043, 2006.

[114] M. Blanc, W. Y. Hsieh, K. A. Robertson et al., "Host defense against viral infection involves interferon mediated down-regulation of sterol biosynthesis," *PLoS Biology*, vol. 9, no. 3, Article ID e1000598, 2011.

[115] J. L. Wylie, G. M. Hatch, and G. Mcclarty, "Host cell phospholipids are trafficked to and then modified by *Chlamydia trachomatis*," *Journal of Bacteriology*, vol. 179, no. 23, pp. 7233–7242, 1997.

[116] D. W. Good, T. George, and B. A. Watts III, "Lipopolysaccharide directly alters renal tubule transport through distinct TLR4-dependent pathways in basolateral and apical membranes," *American Journal of Physiology—Renal Physiology*, vol. 297, no. 4, pp. F866–F874, 2009.

[117] A. Cuenda and S. Rousseau, "p38 MAP-kinases pathway regulation, function and role in human diseases," *Biochimica et Biophysica Acta—Molecular Cell Research*, vol. 1773, no. 8, pp. 1358–1375, 2007.

[118] Y. Li and P. Agarwal, "A pathway-based view of human diseases and disease relationships," *PLoS ONE*, vol. 4, no. 2, Article ID e4346, 2009.

[119] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam, "The protein kinase complement of the human genome," *Science*, vol. 298, no. 5600, pp. 1912–1934, 2002.

[120] K. L. Pierce, R. T. Premont, and R. J. Lefkowitz, "Seven-transmembrane receptors," *Nature Reviews Molecular Cell Biology*, vol. 3, no. 9, pp. 639–650, 2002.

[121] T. M. Ehrman, D. J. Barlow, and P. J. Hylands, "Phytochemical informatics of traditional Chinese medicine and therapeutic relevance," *Journal of Chemical Information and Modeling*, vol. 47, no. 6, pp. 2316–2334, 2007.

[122] P. A. Williams, J. Cosme, V. Sridhar, E. F. Johnson, and D. E. McRee, "Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity," *Molecular Cell*, vol. 5, no. 1, pp. 121–131, 2000.

[123] C. T. Supuran and A. Scozzafava, "Carbonic anhydrases as targets for medicinal chemistry," *Bioorganic & Medicinal Chemistry*, vol. 15, no. 13, pp. 4336–4350, 2007.

[124] J. L. Boyer, "Bile formation and secretion," *Comprehensive Physiology*, vol. 3, no. 3, pp. 1035–1078, 2013.

[125] J. Chen, "TCM and infectious disease," *Acupuncture Today*, vol. 12, no. 5, 2011, http://www.acupuncturetoday.com/mpacms/at/article.php?id=32393.

[126] A. E. Brady and L. E. Limbird, "G protein-coupled receptor interacting proteins: emerging roles in localization and signal transduction," *Cellular Signalling*, vol. 14, no. 4, pp. 297–309, 2002.

[127] W.-J. Bei, J. Guo, H.-Y. Wu, and Y. Cao, "Lipid-regulating effect of traditional Chinese medicine: mechanisms of actions," *Evidence-Based Complementary and Alternative Medicine*, vol. 2012, Article ID 970635, 10 pages, 2012.

[128] H. K. Ghayee and R. J. Auchus, "Basic concepts and recent developments in human steroid hormone biosynthesis," *Reviews in Endocrine and Metabolic Disorders*, vol. 8, no. 4, pp. 289–300, 2007.

[129] B. Zhu and H. Wang, *Diagnostics of Traditional Chinese Medicine*, Jessica Kingsley, London, UK, 2011.

[130] A. Bender, "Compound bioactivities go public," *Nature Chemical Biology*, vol. 6, article 309, 2010.

[131] K. M. Wu, J. G. Farrelly, R. Upton, and J. Chen, "Complexities of the herbal nomenclature system in traditional Chinese medicine (TCM): lessons learned from the misuse of *Aristolochia*-related species and the importance of the pharmaceutical name during botanical drug product development," *Phytomedicine*, vol. 14, no. 4, pp. 273–279, 2007.