

基于SoFDA病-证-方关联的本体表征、富集分析和相似度计算评价

刘毓东, 许佳, 于泽从, 陈同, 王宁, 杜霞
王萍, 周雪忠, 许海玉, 张彦琼

中国中医科学院中药研究所 1
北京交通大学计算机与信息技术学院医学智能研究所、交通数据分析与挖掘中心
北京市重点实验室 2
中国中医科学院广安门医院 3
中国中医科学院中药资源中心 4



Yudong Liu, Jia Xu, Zecong Yu, Tong Chen, Ning Wang, Xia Du, *et al.* 2022. Ontology characterization, enrichment analysis, and similarity calculation-based evaluation of disease–syndrome–formula associations by applying SoFDA. *iMeta* 1: e1. <https://doi.org/10.1002/imt2.80>

摘要

- SoFDA是首个开源的中医证候本体数据库
- SoFDA通过多维异构生物网络，揭示了疾病、中医证候、症状、基因、中药方剂、药物靶标和通路之间的关联性及其生物内涵
- 采用基于相似度的关联评价体系，通过计算Jaccard或余弦相似度，定量评价“疾病-证候”“证候-证候”“方剂-证候”与“疾病-证候-方剂”的关联度



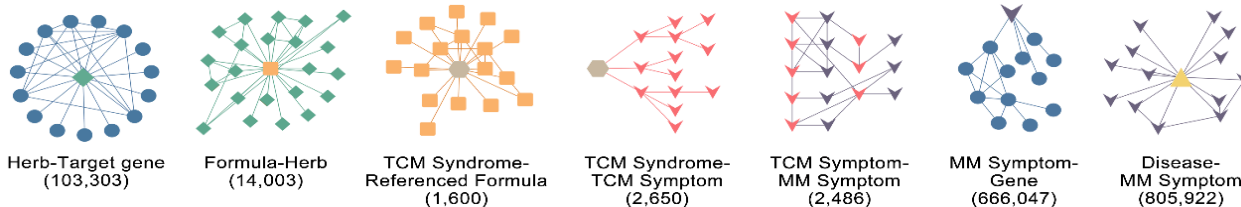
SOFDA平台的数据支持和来源

(A) Basic Data Sources

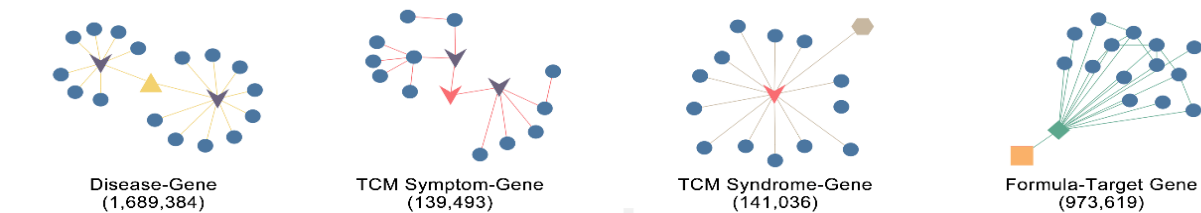
▲ Diseases(8,045)
 ▼ TCM Symptoms
 ◆ TCM Syndromes(319)
 ▼ MM Symptoms
 ■ Formulas(1,359)
 ◆ Herbs

Genes Pathways Network density Network module GO terms(BP,MF,CC)

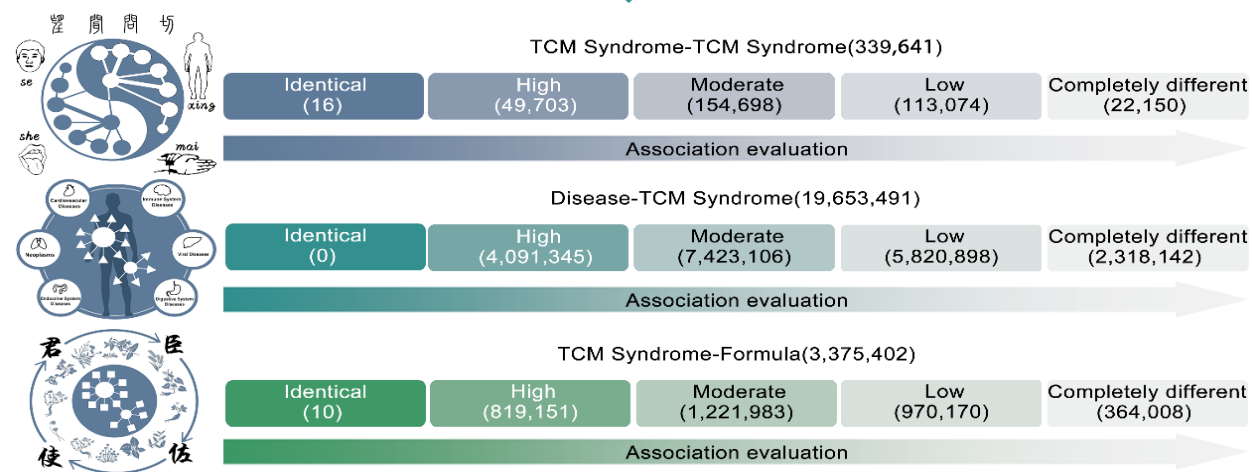
(B) Direct Associations



(C) Indirect Associations



(D) Evaluation



数据收集、关联性分析和SoFDA平台概况的方案

Data Collection

- Syndrome data
 - Syndrome name
 - Symptom
 - Alias of syndrome
 - Pathogenic syndrome element
 - Syndrome element of disease location
 - Formula against syndrome
- Disease data
 - MalaCards HUMAN DISEASE DATABASE
 - OMIM Human Genetics Knowledge for the World
 - human phenotype ontology
 - orphanet
 - DisGeNET
 - Disease name
 - Global category
 - Symptom
 - Anatomical category
 - Disease-related gene
- Formula data
 - ETCM "Chinese Medicine Diagnostics"
 - Formula name
 - Symptom
 - Herbs containing the formula
 - Syndromes treated by formula
 - Formula-related genes

Association evaluation

- Association
 - Syndrome
 - Disease
 - Formula
- Arithmetic
 - Cosine similarity
 - Functions
 - Pathways
 - Network features
 - Jaccard similarity
 - Genes
 - Symptoms
- Quantitative
 - Identical
 - High
 - Moderate
 - Low
 - Completely different

Association similarity

SoFDA

- Function
 - Syndrome-Syndrome association
 - Disease-Syndrome association
 - Formula-Syndrome association
 - Disease-Syndrome-Formula
 - Symptom/Gene Enrichment
 - Multi-level Network Visualization
 - Syndrome Classification Code Prediction
- Visualization
 - Network visualization
 - Heatmap visualization
- www.tcmip.cn/Syndrom/front/#

SoFDA平台主界面

SoFDA

Home

Browse ▾

Association Evaluation ▾

Tools ▾

Help

User center

(A) 中文 | English

太阳蓄水证(B)

Submit

e.g. Syndrome: 气阴两虚证 疫毒闭肺证 e.g. Disease: COVID-19 e.g. Formula: 化湿败毒方 清肺排毒汤

SoFDA is the first available collections of traditional Chinese medicine (TCM) syndrome ontology, the syndrome classification tool and the related-feature associations with disease and prescriptions for investigating the pathological links and therapeutic mechanisms. It contains the detailed information of 310 TCM syndromes recorded in the monograph of "Chinese Medicine Diagnostics" and 9 novel syndromes of COVID-19, as well as 8045 diseases and 1359 TCM prescriptions used for the corresponding syndromes. Considering the lack of classification information for most TCM syndromes affecting

Syndrome Browse Page

Classification and codes of diseases and ZHENG of traditional Chinese medicine (GB/T 15657-1995)



(C) Symptom Enrichment Analysis

Gene Enrich >

Please select the symptom type ▾

Click one or more Symptom to choose ▾

e.g. Primary Symptoms:

Fever Malignant Hyperthermia Pharyngeal Edema
Ulcer Abnormality Of The Pharynx

e.g. Secondary Symptoms:

Xerostomia Dysphoric Mood Skin Rash

e.g. All:

Fever Malignant Hyperthermia Pharyngeal Edema
Xerostomia Abnormality Of The Pharynx Ulcer
Dysphoric Mood Skin Rash

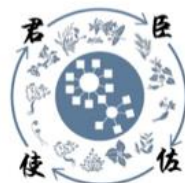
Submit



Syndrome-Syndrome
Association



Disease-Syndrome
Association



Formula-Syndrome
Association



Disease-Formula-
Syndrome Association



User center



浏览SoFDA平台的实用指南

SoFDA
Home
Browse ▾
Association Evaluation ▾
Tools ▾
Help
User center
中文 | English

Syndrome Browse Page (A)

Classification and codes of diseases and ZHENG of traditional Chinese medicine (GB/T 15657-1995, Released in 1995)

- Cause Of Disease Syndrome Type
- Syndrome Type Of Six Meridian
- Syndrome Type Of Defense, Qi, Nutrient And Blood
- Syndrome Type Of Yin, Yang, Qi, Blood, Fluid And Humor, Phlegm
- Syndrome Type Of Viscera And Meridians
- * Syndrome Type Of Six Meridian
- * Syndrome Type Of Yin, Yang, Qi, Blood, Fluid And Humor, Phlegm
- * Syndrome Type Of Viscera And Meridians
- * Cause Of Disease Syndrome Type
- * Syndrome Type Of Defense, Qi, Nutrient And Blood

*Note: The pre-coded classification and code for the corresponding syndrome.

Classification and Codes of Diseases and Patterns of Traditional Chinese Medicine (Revised edition of GB/T15657-1995, Released in 2020)

- Syndrome Type Of Cause Of Disease Terms
- Syndrome Type Of Viscera And Sense Organ Terms
- Syndrome Type Of Six Meridian Terms
- Syndrome Type Of Defense, Qi, Nutrient And Blood Aspect Terms
- Syndrome Type Of Qi, Blood, Yin, Yang, Essence And Fluid Terms
- Syndrome Type Of Eight Principles Terms
- Syndrome Type Of Meridian And Collatera Terms
- Syndrome Type Of Triple Energizer Terms
- Unknown

Primary coding of TCM Syndromes

Differentiation System for Syndrome Element of Disease Location

Differentiation System for Pathogenic Syndrome Element

Syndrome Name	Alias of Syndrome	Formula Against Syndrome
Syndrome Of Toxin Congesting The Upper Energizer	Syndrome Of Toxin Congesting Qifen	QING XIN LIANG GE SAN QING XAN ZHI CHI TANG XI LEI SAN YU YAO SHI
Syndrome Of Wind-Heat In The Qj Wheel	Syndrome Of Wind-Heat In The White Of The Eye	SANG BAI PI SAN XIE FEI TANG
Syndrome Of Nutritional Damage Of Greater Yang And Nutr	Syndrome Of Nutritional Damage And Generalized Pain Ache	GUI ZHI JIA SHAO YAO SHENG JING GE YI LIANG REN SHI
Syndrome Of Pathogen Invasion In Greater Yang And Spike	Syndrome Of Bamsu Cinnamon And Rituzabw Decoction	GUI ZHI JIA DA HUANG TANG
Syndrome Of Pathogen Stagnation In The Exterior Skin	Syndrome Of Similar To Malaria Of Greater Yang, Syndrome	GUI ZHI MA HUANG GE BAN TANG
Syndrome Of Heat Pathogenic Distressing The Lung In Great	Syndrome Of Pathogenic Heat Distressing The Lung As Auth	MA HUANG XING REN GAN CAO SHI GAO TANG
Syndrome Of Spleen Deficiency And Qi Stagnation In Greater	Syndrome Of Magnolia Officinalis, Ginger Plus Pinellia Tem	HOU PO SHENG JIANG BAN XIA GAN CAO REN SHEN TAN
Syndrome Of Stomach Qi Ascending Counterflow In Greater	Syndrome Of Inakufower And Hematite Decoction	XUAN FU DAI ZHE TANG
Syndrome Of Superficies Cold And Interior Fluid In Greater Y	Syndrome Of Exterior Excess And Retained Fluid, Syndrome C	XIAO QING LONG TANG
Syndrome Of Superficies Cold And Interior Heat In Greater Y	Syndrome Of Decoction Da Qing Long	DA QING LONG TANG

Disease Browse Page (B)

Global category

Anatomical category

Disease MCID	Disease Name	Symptom	Category
11P002	11p154 Microduplication Syndrome	Appendicular Hypotonia Macrolalia Anteverted Nares Synophrys Aggressive Behavior Delayed Speech And Language Development	Global categoryAnatomical category
11Q201	11q22.2q22.3 Microdeletion Syndrome	Thick Eyebrow Obsessive-Compulsive Behavior Anxiety Delayed Speech And Language Development Autism With High Cognitive Abilities Poor Eye Contact	Global categoryAnatomical category

Formula Browse Page (C)

Formula Distribution according to the Corresponding Syndrome Coding

English Name of Formula	Herbs containing Formula	Syndrome Treated by Formula
BU ZHONG YI QI TANG	HUANG QI GAN CAO REN SHEN DANG GUI JU PI SHENG MA ZHAI XI	Syndrome Of Qi Deficiency In The Flesh Wheel Syndrome Of Qi Deficiency In Water Wheel Sunken Middle Qi Syndrome Syndrome Of Qing Yang Non-Promotion Sunken Spleen Qi Syndrome
DA CHENG QI TANG	DA HUANG HOU PO ZHI SHE MANG XIAO	Lesser Yin Combined Yang Brightness Syndrome Syndrome Of Liver Wind And Phlegm-Fire In Water Wheel Syndrome Of Deficiency Fire Invading Throat In Lesser Yin Large Intestinal Heat Bind Syndrome Syndrome Of Accumulated Heat In Stomach And Intestines Syndrome Of Yang Brightness And Bowel Excess

SoFDA平台证-证关联

Syndrome-Syndrome Association

Syndrome-Syndrome Association

* Syndrome Type 1: Syndrome 1: * Syndrome Type 2: Syndrome 2:

* Symptom: * Association Item:

Primary Symptom
Secondary Symptom
Primary & Secondary Symptom

Symptom
Symptom-related Genes
GO-BP
GO-CC
GO-MF
Network Density
Network Module
Reactome Pathway

4

Syndrome1	Syndrome2	Association Value	Association Level	Association Item
Syndrome Of Dampness Obstructing And Ar	Dual Deficiency Of The Lung-Spleen Syndro	0.533	High	Symptom
Cold-Dampness Obstructing The Lung Synd	Dual Deficiency Of The Lung-Spleen Syndro	0.5	High	Symptom
Dampness Toxin Depression Lung Syndrome	Cold-Dampness Obstructing The Lung Synd	0.467	High	Symptom
Syndrome Of Dampness Obstructing And Ar	Cold-Dampness Obstructing The Lung Synd	0.389	High	Symptom
Dampness Toxin Depression Lung Syndrome	Dual Deficiency Of The Lung-Spleen Syndro	0.25	High	Symptom

For visualization

Illustration Filter (Primary & Secondary Symptom)

Syndrome 1: Syndrome 2: Association Level: Association Item:



SoFDA平台病-证关联

Disease-Syndrome Association

* Syndrome Type: (1) Syndrome: Disease: * Symptom: (2)

* Association Item: (3)

Symptom
Symptom-related Genes
GO-BP
GO-CC
GO-MF
Network Density
Network Module
Reactome Pathway

Primary Symptom
Secondary Symptom
Primary & Secondary Symptom

↓

Syndrome	Disease	Association Value	Association Level	Association Item
Large Intestinal Heat Blind Syndrome	13q12.3 Microdeletion Syndrome	0.051	High	Symptom
Gallbladder Heat Syndrome	13q12.3 Microdeletion Syndrome	0.025	Moderate	Symptom
Gallbladder Heat Syndrome	12q14 Microdeletion Syndrome	0.193	High	Symptom-related Genes
Gallbladder Qi Deficiency Syndrome	16p12.1p12.3 Triplication Syndrome	0.184	High	Symptom-related Genes
Large Intestinal Heat Blind Syndrome	12q14 Microdeletion Syndrome	0.184	High	Symptom-related Genes

(4)

For visualization

Illustration Filter (Primary Symptom)

Syndrome: (5) Disease: Association Level: Association Item:



SoFDA平台方-证关联

Formula-Syndrome Association

* Syndrome Type: Syndrome: * Formula: * Symptom:

* Association Item:

Symptom
Symptom-related Genes
GO-BP
GO-CC
GO-MF
Network Density
Network Module
Reactome Pathway

Primary Symptom
Secondary Symptom
Primary & Secondary Symptom

syndrome1	Syndrome2	Association Value	Association Level	Association Item
syndrome Of Dampness Obstructing And Ar	Dual Deficiency Of The Lung-Spleen Syndro	0.533	High	Symptom
old-Dampness Obstructing The Lung Synd	Dual Deficiency Of The Lung-Spleen Syndro	0.5	High	Symptom
ampness Toxin Depression Lung Syndrom	Cold-Dampness Obstructing The Lung Synd	0.467	High	Symptom
syndrome Of Dampness Obstructing And Ar	Cold-Dampness Obstructing The Lung Synd	0.389	High	Symptom
ampness Toxin Depression Lung Syndrom	Dual Deficiency Of The Lung-Spleen Syndro	0.25	High	Symptom

Illustration Filter (Primary & Secondary Symptom)

Syndrome: Formula: Association Level: Association Item:



SoFDA平台病-证-方关联

(A) Disease-Syndrome-Formula Association with the Same Disease

Disease-Syndrome-Formula Association with the Same Disease

Disease-Syndrome-Formula Association with the Same Syndrome

Disease-Syndrome-Formula Association with the Same Disease

* Disease:
Please Select disease

* Syndrome:
Please Select syndrome

* Symptom:
Primary Symptom

* Association Item:
Symptom-related Genes

Association Level:
Low +2

* Formula:
Please Select formula

Submit

Completely Different
Low
Moderate
High
Identical

Primary Symptom
Secondary Symptom
Primary & Secondary Symptom

Symptom
Symptom-related Genes
GO-BP
GO-CC
GO-MF
Network Density
Network Module
Reactome Pathway

Disease-Syndrome-Formula Association with the Same Syndrome

Disease-Syndrome-Formula Association with the Same Disease

Disease-Syndrome-Formula Association with the Same Syndrome

* Syndrome:
Please Select syndrome

* Symptom:
Primary Symptom

* Disease:
Please Select disease

* Association Item:
Symptom-related Genes

Association Level:
Please Select syndrome

* Formula:
Please Select formula

Submit

Completely Different
Low
Moderate
High
Identical

Primary Symptom
Secondary Symptom
Primary & Secondary Symptom

Symptom
Symptom-related Genes
GO-BP
GO-CC
GO-MF
Network Density
Network Module
Reactome Pathway

Disease	Syndrome	Formula	Correlation
12q14 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BA ZHEN TANG	0.042
12q14 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BA ZHENG SAN	0.053
12q14 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BA ZHEN YI MU GAO	0
12q14 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BA ZHENG HE JI	0
12q14 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BAI BU WAN	0

(B) Disease-Syndrome-Formula Association with the Same Syndrome

Disease-Syndrome-Formula Association with the Same Disease

Disease-Syndrome-Formula Association with the Same Syndrome

Disease-Syndrome-Formula Association with the Same Disease

* Syndrome:
Please Select syndrome

* Symptom:
Primary Symptom

* Disease:
Please Select disease

* Association Item:
Symptom-related Genes

Association Level:
Please Select syndrome

* Formula:
Please Select formula

Submit

Completely Different
Low
Moderate
High
Identical

Primary Symptom
Secondary Symptom
Primary & Secondary Symptom

Symptom
Symptom-related Genes
GO-BP
GO-CC
GO-MF
Network Density
Network Module
Reactome Pathway

Disease-Syndrome-Formula Association with the Same Syndrome

Disease-Syndrome-Formula Association with the Same Disease

Disease-Syndrome-Formula Association with the Same Syndrome

* Syndrome:
Please Select syndrome

* Symptom:
Primary Symptom

* Disease:
Please Select disease

* Association Item:
Symptom-related Genes

Association Level:
Please Select syndrome

* Formula:
Please Select formula

Submit

Completely Different
Low
Moderate
High
Identical

Primary Symptom
Secondary Symptom
Primary & Secondary Symptom

Symptom
Symptom-related Genes
GO-BP
GO-CC
GO-MF
Network Density
Network Module
Reactome Pathway

Disease	Syndrome	Formula	Correlation
11q22.2q22.3 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BAI HU TANG	0.121
11q22.2q22.3 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BAO HE WAN	0.078
11q22.2q22.3 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BAI TONG TANG	0.032
11q22.2q22.3 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BAO AN WAN LING DAN	0.082
11q22.2q22.3 Microdeletion Syndrome	Syndrome Of Cold Congealing In The Uterus	BAI ZI ZI XIN WAN	0.037



SoFDA平台多维网络可视化工具

Custom Network

Syndrome: Syndrome Of Cold... +2 Symptom: Primary & Secondary Symptom Disease: Bone Pain Period... +2 Formula: 柴胡桂枝汤... +2

Show: Disease +3

Submit

Disease
Syndrome
Formula
Symptom
Gene
Reactome Pathway

Primary Symptom
Secondary Symptom
Primary & Secondary Symptom

The list of core network targets and their topological eigenvalues

Gene	network target (True/False)	Degree	Closeness	Betweenness
RP529	true	21	0.015	0.292
RP517	true	21	0.015	0.292
RP526	true	21	0.015	0.292
RP520	true	21	0.015	0.292
RP514	true	22	0.015	24.292

For visualization

Edit

The visualization shows a network graph with a central hub node (a red diamond) and several clusters of peripheral nodes (blue circles). The nodes are interconnected by edges, forming a complex network structure. The clusters are arranged around the central hub, with some clusters being more densely connected than others. The overall structure is symmetrical and highly interconnected.



总结

SoFDA是一个颇具前景的平台，可以将中医证候、疾病、方剂和分子机制的信息联系起来，以深化我们对古代医学、中医学以及相应医疗干预的剖析。

Yudong Liu, Jia Xu, Zecong Yu, Tong Chen, Ning Wang, Xia Du, et al. 2022. Ontology characterization, enrichment analysis, and similarity calculation-based evaluation of disease–syndrome–formula associations by applying SoFDA. *iMeta* 1: e1. <https://doi.org/10.1002/imt2.80>





“**iMeta**”由威立、肠菌分会和华人科学家出版的开放获取期刊，主编由中科院微生物所刘双江和荷兰格罗宁根大学傅静远教授共同担任。目的是发表原创研究、方法和综述以促进宏基因组学、微生物组和生物信息学发展。目标是发表前10%(IF > 15)的高影响力论文。期刊特色包括视频投稿、可重复分析、图片打磨、青年编委、前3年免出版费、50万用户的社交媒体宣传等。2022年的**三月**、**六月**、**九月**和**十二月**期已正式在线出版发行，相继被[Google Scholar](https://scholar.google.com/)、[PubMed\(部分\)](https://pubmed.ncbi.nlm.nih.gov/)、[DOAJ](https://doaj.org/)、[Scopus](https://scopus.com/)等数据库收录！



主页: <http://www.imeta.science>

出版社: <https://wileyonlinelibrary.com/journal/imeta>



投稿: <https://mc.manuscriptcentral.com/imeta>



office@imeta.science



宣传片



[iMeta](https://www.imeta.science)

