Article

Network pharmacology of medicinal attributes and functions of Chinese herbal medicines: (III) Canonical correlation functions between attribute classes and linear eignmodels of Chinese herbal medicines

WenJun Zhang
School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China; International Academy of Ecology and Environmental Sciences, Hong Kong
E-mail: zhwj@mail.sysu.edu.cn, wjzhang@iaees.org

Received 10 April 2017; *Accepted 28 April 2017*; *Published 1 September 2017*

Abstract
In present study I used the data from CHM-DATA, the interactive database of 1127 Chinese herbal medicines. Canonical correlation functions were determined for taste attribute class (7 taste attributes), medicinal property class (5 medicinal properties), chemical composition class (22 chemical composition categories), meridians and colleterals class (12 meridians and colleterals), and medicinal function class (77 medicinal functions). Linear eignmodels were also developed for Chinese herbal medicines. Theoretically the attribute values of any Chinese herbal medicines meet the corresponding linear eignmodel. Matlab codes for canonical correlation analysis and linear eignmodel were given. Finally, the canonical correlation network for attribute classes of Chinese herbal medicines was constructed.

Keywords Chinese herbal medicine; medicinal function; attribute; canonical correlation; eignmodel.

Network Pharmacology
ISSN 2415-1084
URL: http://www.iaees.org/publications/journals/np/online-version.asp
RSS: http://www.iaees.org/publications/journals/np/rss.xml
E-mail: networkpharmacology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

The single drug-single target-single disease view in traditional western medicine (Hopkins，2007, 2008; Budovsky and Fraifeld, 2012) met various problems over the past 20 years (Zhang, 2016, 2017a-b)。Traditional Chinese Medicne takes biological network regulation as the theoretical basis, and thus provides a new thinking and new approach for drug design and disease treatment. However, the theory of Traditional Chinese Medicine has developed so slowly in the past thousands of years. So far we still lack of fundamental research on Chinese herbal medicines, which greatly retards the development and practice of the theory of Chinese herbal medicines. For this reason, Zhang (2017a) collected a total of 1127 Chinese herbal medicines mainly with recorded chemical composition, and calculated the basic statistics of medicinal attributes and functions, e.g., totals, frequencies or probabilities, percentages, etc., on the basis of total population of medicines and families. Thereafter, four relational networks, i.e., the networks for medicinal attributes and functions, for chemical composition and meridians and collaterals, for meridians and collaterals and medicinal functions, and for meridians and collaterals were constructed based on the significant point correlations (Zhang, 2017b). Network analysis indicated that the former three ones are scale-free complex networks and node degrees of the four networks followed power-law distribution. Detailed between-attribute relationships and medicinal mechanisms were revealed (Zhang, 2017b).

Based on the previous studies (Zhang, 2017a, b), this study will further determine the canonical correlations between attribute classes and develop the standard models of Chinese herbal medicines, in order to lay a foundation for further studies.

2 Material and Methods

2.1 Methods

2.1.1 Canonical Correlation Analysis

In present study, canonical correlation analysis is used to determine the correlation between two attribute classes (Zhang and Fang, 1982; Qi and Xu, 2009). Suppose the attribute class x has m attributes $x_1, x_2, \ldots, x_m$, $x=(x_1, x_2, \ldots, x_m)$, and the attribute class y has p attributes $y_1, y_2, \ldots, y_p$, $y=(y_1, y_2, \ldots, y_p)$, $m \leq p$. We want to analyze the correlation by determining the correlation between ux^T and vy^T, where $u=(u_1, u_2, \ldots, u_m)$, $v=(v_1, v_2, \ldots, v_p)$. The degree of correlation between ux^T and vy^T changes with different u and v. We need to determine u and v, such that the linear correlation between ux^T and vy^T is the strongest. First, assume there are n medicines, and the raw data are as follows

$$x=(x_{ij}), \quad j=1, 2, ..., m$$
$$y=(y_{ij}), \quad j=1, 2, ..., p$$
$$i=1, 2, ..., n$$

In present study, x_{ij} and y_{ij} take 0 or 1. Let $x_{ij}= x_{ij} - x_{bar\ j}$, $y_{ij}= y_{ij} - y_{bar\ j}$, where

$$x_{bar\ j}=\sum^n_{i=1} x_{ij}/n, \quad j=1, 2, ..., m$$
$$y_{bar\ j}=\sum^n_{i=1} y_{ij}/n, \quad j=1, 2, ..., p$$

Calculate

$$e_{ij}=\sum^n_{k=1} x_{ik}\, x_{kj}/n, \quad i, j=1, 2, ..., m$$
$$f_{ij}=\sum^n_{k=1} y_{ik}\, y_{kj}/n, \quad i, j=1, 2, ..., p$$
$$g_{ij}=\sum^n_{k=1} x_{ki}\, y_{kj}/n, \quad i=1, 2, ..., m; j=1, 2, ..., p$$
$$h_{ij}=\sum^n_{k=1} y_{ki}\, x_{kj}/n, \quad i=1, 2, ..., p; j=1, 2, ..., m$$

Let $E=(e_{ij})$, $F=(f_{ij})$, $G=(g_{ij})$, $H=(h_{ij})$. Determine eignvalues $l_1^2, l_2^2, \ldots, l_m^2$, and the corresponding eignvector pairs $u_1, v_1; u_2, v_2; \ldots; u_m, v_m$,

$$(E^{-1}*G*F^{-1}*H-l^2*I)u=0$$
$$(F^{-1}*H*E^{-1}*G-l^2*I)v=0$$

where I is the unit matrix. Finally, the canonical correlation coefficients (absolute values) are $l_1, l_2, \ldots, l_m$, and the canonical attribute pairs, or correlation functions are obtained as

$$u_i=\sum^{m}_{k=1} u_{ik}\, x_k$$
$$v_i=\sum^{p}_{k=1} v_{ik}\, y_k$$
$$i=1, 2, ..., m$$

2.1.2 Linear eignmodels

For the above (u_i, v_i), i=1, 2, ..., m, develop linear regression with u_i (or v_i) as independent variable, and v_i (or u_i) as dependent variable, and choose these models with statistic significance. For example

$$u_i = a + b\, v_i$$

Consequently, we achieve a linear eignmodel of Chinese herbal medicines as the following

$$\sum^{m}_{k=1} u_{ik}\, x_k = a + b \sum^{p}_{k=1} v_{ik}\, y_k$$

In a statistic sense, the attribute values of any Chinese herbal medicine meet the corresponding linear eignmodels.

The following are the Matlab codes for canonical correlation analysis and linear eignmodel calculation, CanonicalCorreAnaly.m

```
% Zhang WJ. 2017. Network pharmacology of medicinal attributes and functions of Chinese herbal medicines:
% (III) Canonical correlation functions between attribute classes and linear eignmodels of Chinese herbal medicines.
% Network Pharmacology, 2(3): 67-81
m=input('Input the number of variables x: ');
p=input('Input the number of variables y: ');
if (m>p) disp('Variables x should be less than variables y'); pause; end
file=input('Input the excel file name of data, e.g., cano.xls. The first m columns are for variables x and the followed p columns
are for variables y: ','s');
xy=xlsread(file);
n=size(xy,1);
x=xy(:,1:m);
y=xy(:,m+1:m+p);
xb=zeros(1,m); yb=zeros(1,p);sigx=zeros(m); sigy=zeros(p);
sigxy=zeros(m,p); sigyx=zeros(p,m); mat1=zeros(m); mat2=zeros(p);
u=zeros(m); v=zeros(p); val1=zeros(m); val2=zeros(p);
xb=mean(x);
yb=mean(y);
for i=1:m
x(:,i)=x(:,i)-xb(i);
end;
for i=1:p
y(:,i)=y(:,i)-yb(i);
end;
for i=1:m
for j=1:m
sigx(i,j)=0;
```

```
for k=1:n
sigx(i,j)=sigx(i,j)+x(k,i)*x(k,j);
end
sigx(i,j)=sigx(i,j)/n;
end; end
for i=1:p
for j=1:p
sigy(i,j)=0;
for k=1:n
sigy(i,j)=sigy(i,j)+y(k,i)*y(k,j);
end
sigy(i,j)=sigy(i,j)/n;
end; end
for i=1:m
for j=1:p
sigxy(i,j)=0;
for k=1:n
sigxy(i,j)=sigxy(i,j)+x(k,i)*y(k,j);
end
sigxy(i,j)=sigxy(i,j)/n;
end; end
for i=1:p
for j=1:m
sigyx(i,j)=0;
for k=1:n
sigyx(i,j)=sigyx(i,j)+y(k,i)*x(k,j);
end
sigyx(i,j)=sigyx(i,j)/n;
end; end
sigx=sigx^(-1);
sigy=sigy^(-1);
mat1=sigx*sigxy*sigy*sigyx;
mat2=sigy*sigyx*sigx*sigxy;
[u,val1]=eig(mat1);
[v,val2]=eig(mat2);
for i=1:m
p2(i)=i;
end
for i=1:m-1
k=i;
for j=i:m-1
if (val1(j+1,j+1)>val1(k,k)) k=j+1; end
end
i2=p2(i); p2(i)=p2(k); p2(k)=i2;
l=val1(i,i); val1(i,i)=val1(k,k); val1(k,k)=l;
```

```
end
iss='\n';
for k=1:m
iss=strcat(iss,'No. ',num2str(k),' canonical correlation coefficient: ',num2str(round(sqrt(val1(k,k))*10000)/10000.00),'\n');
iss=strcat(iss,'u',num2str(k),'=');
for i=1:m
e1=num2str(i);
if (u(i,p2(k))>0) e2=num2str(round(u(i,p2(k))*100000)/100000.00);
elseif (u(i,p2(k))<0) e2=num2str(round(abs(u(i,p2(k)))*100000)/100000.00);
end
if (u(i,p2(k))>0) iss=strcat(iss,'+',e2,'x',e1);
elseif (u(i,p2(k))<0) iss=strcat(iss,'-',e2,'x',e1);
end
end
iss=strcat(iss,'\n');
iss=strcat(iss,'v',num2str(k),'=');
for i=1:p
e1=num2str(i);
if (v(i,p2(k))>0) e2=num2str(round(v(i,p2(k))*100000)/100000.00);
elseif (v(i,p2(k))<0) e2=num2str(round(abs(v(i,p2(k)))*100000)/100000.00);
end
if (v(i,p2(k))>0) iss=strcat(iss,'+',e2,'y',e1);
elseif (v(i,p2(k))<0) iss=strcat(iss,'-',e2,'y',e1);
end
end
iss=strcat(iss,'\nLinear regression between u',num2str(k),' and', ' v',num2str(k),'\n');
for j=1:n
uxx(j)=x(j,:)*u(:,k);
vyy(j)=y(j,:)*v(:,k);
end
for j=1:2
if (j==1) xx=uxx'; yy=vyy';
else xx=vyy';yy=uxx';
end
[bb,bint,rr,rrint,stats]=regress(yy,[ones(n,1) xx]);
if (j==1) iss=strcat(iss,'u',num2str(k),'=',num2str(bb(1)),'+',num2str(bb(2)),'*','v',num2str(k),'\n');
else iss=strcat(iss,'v',num2str(k),'=',num2str(bb(1)),'+',num2str(bb(2)),'*','u',num2str(k),'\n');
iss=strcat(iss,'Pearson r=',num2str(sign(bb(2))*sqrt(stats(1))),', p=',num2str(stats(3)),'\n');
end
end
iss=strcat(iss,'\n');
end
fprintf(iss)
```

2.2 Data source

I used the interactive database of eight tables, CHM-DATA Version 1.0 (Zhang, 2017a, b), with 1127 Chinese herbal medicines mainly having recorded chemical composition, of which 210 families and approximately 2000 species of medicinal plants and fungi were involved, which account for approximately 1/5 of medicinal plants and fungi in China. Among them, medicinal plants accounted for 98.94%, and medicinal fungi accounted for 1.06%. The list included the most commonly used or important Chinese herbal medicines. The data with missing attribute values are ignored in the study. Finally, the taste attribute class (7 taste attributes, Table 1), medicinal property class (5 medicinal properties, Table 2), chemical composition class (22 chemical composition categories, Table 3), meridians and collaterals class (12 meridians and collaterals (Gui Jing), Table 4), and medicinal function class (77 medicinal functions (Gong Xiao), Table 5), were used for further analysis.

Table 1 Taste attribute class.

Taste	Bitter	Symplectic	Sweet	Light	Sour	Astringent	Salty
味	苦	辛	甘	淡	酸	涩	咸
x	x_1	x_2	x_3	x_4	x_5	x_6	x_7
y	y_1	y_2	y_3	y_4	y_5	y_6	y_7

Table 2 Medicinal property class.

Property	Cold	Cool	Temperate	Warm	Hot
性	寒	凉	平	温	热
x	x_1	x_2	x_3	x_4	x_5

Table 3 Chemical composition class.

Chemical composition Categories	Glycosides	Organic acids	Alkaloids	Amines	Sterols	Volatile oils or ordinary oils	Proteins or amino acids	Terpenoids	Phenols	Aldehydes	Esters or fats
成份	甙类	有机酸类	生物碱类	胺类	甾醇类	挥发油类/油类	蛋白质/氨基酸类	萜类	酚类	醛类	酯类/脂肪
y	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}
Chemical composition Categories	Carbohydrates or starch	Alcohols	Enzymes	Ketones or flavonoids	Alkanes or hydrocarbons	Ethers	Olefins	Anthracene or quinones	Tannins	Vitamins	Inorganic substances
成份	糖类/淀粉	醇类	酶类	(黄)酮类	烷类/烃类	醚类	烯类	蒽类/醌类	鞣质类	维生素类	无机物
y	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}	y_{21}	y_{22}

Table 4 Meridians and colleterals class.

Meridians & Collaterals	Liver meridians and collaterals	Gallbaldder meridians and collaterals	Urinary bladder meridians and collaterals	Kidney meridians and collaterals	Lung meridians and collaterals	Spleen meridians and collaterals
归经	肝	胆	膀胱	肾	肺	脾
x	x_1	x_2	x_3	x_4	x_5	x_6
Meridians & Collaterals	Stomach meridians and collaterals	Heart meridians and collaterals	Large intestine meridians and collaterals	Small intestine meridians and collaterals	Blood phase	Triple burner
归经	胃	心	大肠	小肠	血分	三焦
x	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}

Table 5 Medicinal function class.

Function	Clean liver, relax liver, consolidate liver, bright eyes or eliminate eye screens	Breed or blacked hair	Benefit gallbladder or cure jaundice	Reduce aminotransferase	Consolidate or warm kidney	Induce diuresis or treat strangurt
功效	清肝/补肝/舒肝/明目/退翳	生发/乌发	利胆/退黄	降转氨酶	补肾/温肾	利尿/通淋
y	y_1	y_2	y_3	y_4	y_5	y_6
Function	Activate water metabolism or excrete water	Invigorate male impotence (Yang) or strengthen male essence	Strengthen bones and muscles	Promote granulation	Remove lung-heat or nourish lung	Eliminate or relieve phlegm
功效	利水/行水	壮阳/温阳/益精	强筋骨	生肌	清肺/润肺	祛痰/化痰
y	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}
Function	Anti-asthma	Eliminate or relieve cough	Eliminate or relieve stuffy nose	Eliminate or relieve tuberculosis	Whet the appetite or reinforce stomach	Strengthen and reinforce spleen
功效	平喘/定喘	止咳	通鼻窍	祛肺结核	开胃/益胃	健脾/补脾
y	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}
Function	Improve digestion	Promote secretion of saliva or body	Relieve sore throat	Resolve food stagnation	Repel foulness	Prevent or arrest vomiting
功效	消食/化食	生津	利咽	消积/消滞	辟秽	止呕
y	y_{19}	y_{20}	y_{21}	y_{22}	y_{23}	y_{24}
Function	Strengthen heart or clean heart-fire	Relieve restlessness, calm the nerves, alleviate mental depression, or arrest convulsion	Arrest epilepsy	Relieve constipation	Loosen the bowels	Moisten dryness
功效	强心/清心	除烦/安神/解郁/定惊	定痫	通便	润肠	润燥
y	y_{25}	y_{26}	y_{27}	y_{28}	y_{29}	y_{30}
Function	Astringe intestine	Soften hardness or dissolve masses	Antidiarrheal	Stop diarrheal	Cool blood	Stop bleeding
功效	涩肠	散结/软坚	止痢	止泻	凉血	止血
y	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}
Function	Tonify blood	Invigorate blood circulation	Absorb clots, eliminate stasis, resolve carbuncle or promote wound healing	Reduce swelling	Antidiabetics	Antiatheroscloresis
功效	养血/补血	活血	化瘀/消痈/敛疮	消肿	降糖	降血脂
y	y_{37}	y_{38}	y_{39}	y_{40}	y_{41}	y_{42}
Function	Antihypertension	Nourish essential fluid (Yin)	Regulate menstruation or promote blood flow	Prevent miscarriage or abortion	Promote lactation or stimulate milk secretion	Regulate or enhance energy flow (Qi)
功效	降压	滋阴	调经/通淋	安胎	通乳/下乳	理气/养气
y	y_{43}	y_{44}	y_{45}	y_{46}	y_{47}	y_{48}
Function	Inhibit or break energy flow (Qi)	Anti-aging	Remove obstruction in meridians and collaterals, or relax the muscles and joints	Nourish, warm spleen, stomach or Qi	Relieve pain	Anticancer
功效	下气/破气	抗衰老	通络/活络/舒筋	温中/和中/补中	止痛	抗癌
y	y_{49}	y_{50}	y_{51}	y_{52}	y_{53}	y_{54}
Function	Clear away heat	Eliminate dampness	Detoxification	Decrease internal heat	Quench ones thirst	Relieve summer-heat
功效	清热	利湿	解毒	降火	止渴	解暑/消暑
y	y_{55}	y_{56}	y_{57}	y_{58}	y_{59}	y_{60}
Function	Dispel endogenous cold	Dispel endogenous damp	Dispel endogenous wind	Relieve rheumatism or lubricate the joints	Dry dampness	Suppress perspiration
功效	祛寒	祛湿	祛风	祛风湿/利关节	燥湿	止汗

y	y_{61}	y_{62}	y_{63}	y_{64}	y_{65}	y_{66}
Function	Induce perspiration	Relieve external syndrome	Promote astringent function	Discharge pus, diminish inflammation or anti-infection	Relieve itching	Kill or expel parasites
功效	发汗	解表/发表	收敛	排脓/消炎/抗感染	止痒	杀虫/驱虫
y	y_{67}	y_{68}	y_{69}	y_{70}	y_{71}	y_{72}
Function	Anti-malaria	Relieve muscular spasm	Expose exthanthema or promote eruption	Dispel evil spirit	Eliminate impediment	
功效	抗疟/截疟	解痉	透疹	逐邪	除痹	
y	y_{73}	y_{74}	y_{75}	y_{76}	y_{77}	

3 Results and Analysis

3.1 Meridians and colleterals class *vs.* chemical composition class

The calculated canonical correlations and linear eignmodels for meridians and colleterals class (Table 4) *vs.* chemical composition class (Table 3) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: 0.3621

u_1=-0.1109x_1+0.10734x_2+0.06532x_3+0.0079x_4+0.01398x_5+0.16885x_6+0.12342x_7-0.00023x_8-0.16883x_9-0.14579x_{10}+0.14219x_{11}+0.92625x_{12}

v_1=-0.13876y_1+0.02522y_2-0.00081y_3+0.61501y_4-0.05993y_5+0.29265y_6+0.28537y_7-0.04131y_8-0.13212y_9+0.10198y_{10}+0.153y_{11}+0.15954y_{12}+0.03951y_{13}-0.13407y_{14}+0.15844y_{15}-0.00407y_{16}+0.07677y_{17}+0.38053y_{18}-0.22554y_{19}-0.28935y_{20}+0.06036y_{21}+0.13855y_{22}

Linear regression between u_1 and v_1

u_1=-2.2731*10^{-17}+0.81879*v_1

Linear regression between v_1 and u_1

v_1=6.7937*10^{-19}+0.16017*u_1

Pearson r=0.36214, p≈0

No.2 canonical correlation coefficient: 0.272

u_2=0.03819x_1-0.03038x_2+0.05457x_3+0.22421x_4+0.03326x_5+0.08942x_6+0.08308x_7+0.00758x_8-0.02874x_9+0.4307x_{10}+0.02628x_{11}-0.86098x_{12}

v_2=-0.065y_1-0.03642y_2-0.11035y_3-0.44763y_4+0.20343y_5-0.15458y_6-0.09163y_7-0.03051y_8+0.12594y_9+0.12925y_{10}+0.05717y_{11}+0.00476y_{12}+0.01791y_{13}-0.28158y_{14}-0.08924y_{15}+0.66085y_{16}-0.06079y_{17}+0.02648y_{18}-0.20732y_{19}-0.0557y_{20}+0.26351y_{21}+0.165y_{22}

Linear regression between u_2 and v_2

u_2=1.7621*10^{-17}+0.40353*v_2

Linear regression between v_2 and u_2

v_2=-4.848*10^{-19}+0.18337*u_2

Pearson r=0.27202, p=1.4544*10^{-14}

No.3 canonical correlation coefficient: -0.2564

u_3=0.20417x_1+0.46968x_2-0.03453x_3+0.01421x_4-0.18606x_5-0.1553x_6+0.24551x_7-0.05308x_8-0.23747x_9-0.23558x_{10}+0.59032x_{11}-0.39259x_{12}

v_3=0.03547y_1+0.31571y_2+0.08208y_3-0.33568y_4-0.04326y_5+0.14556y_6+0.25362y_7+0.6057y_8+0.09551y_9+0.27246y_{10}+0.04342y_{11}-0.25632y_{12}-0.1215y_{13}+0.18414y_{14}-0.06771y_{15}-0.0101y_{16}+0.09705y_{17}-0.11578y_{18}-0.20692y_{19}+0.15564y_{20}+0.15817y_{21}-0.07883y_{22}

Linear regression between u_3 and v_3

u_3=3.5186*10^{-17}-0.31822*v_3

Linear regression between v_3 and u_3

v_3=2.3784*10^{-17}-0.2066*u_3

Pearson r=-0.25641, p=4.6907*10^{-13}

All canonical correlations and linear eignmodels are statistically significant with $p \leq 4.6907*10^{-13}$.

3.2 Medicinal property class *vs.* taste attribute class

The calculated canonical correlations and linear eignmodels for medicinal property class (Table 2) *vs.* taste attribute class (Table 1) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: 0.518

u_1=-0.17722x_1-0.07275x_2+0.11248x_3+0.60466x_4+0.76488x_5

v_1=-0.4475y_1+0.80495y_2-0.03154y_3-0.33218y_4+0.0044y_5+0.071y_6-0.18812y_7

Linear regression between u_1 and v_1

u_1=-1.8564*10^{-17}+0.76742*v_1

Linear regression between v_1 and u_1

v_1=1.3387*10^{-17}+0.34969*u_1

Pearson r=0.51804, $p \approx 0$

No.2 canonical correlation coefficient: 0.2801

u_2=-0.5943x_1-0.18773x_2-0.0663x_3-0.45966x_4-0.62919x_5

v_2=-0.01513y_1+0.16663y_2+0.38585y_3+0.61173y_4+0.34056y_5+0.55242y_6-0.16663y_7

Linear regression between u_2 and v_2

u_2=-1.5874*10^{-17}+0.36872*v_2

Linear regression between v_2 and u_2

v_2=2.5131*10^{-17}+0.2128*u_2

Pearson r=0.28011, p=2.2204*10^{-15}

No.3 canonical correlation coefficient: -0.1656

u_3=-0.36991x_1-0.55525x_2-0.31421x_3-0.4035x_4-0.5416x_5

v_3=0.17287y_1+0.15472y_2-0.10007y_3+0.95268y_4-0.07593y_5-0.13697y_6-0.06357y_7

Linear regression between u_3 and v_3

u_3=1.3933*10^{-17}-0.42811*v_3

Linear regression between v_3 and u_3

v_3=3.3344*10^{-17}-0.064044*u_3

Pearson r=-0.16558, p=3.7421*10^{-6}

All canonical correlations and linear eignmodels are statistically significant with $p \leq 3.7421*10^{-6}$.

3.3 Taste attribute class *vs.* chemical composition class

The calculated canonical correlations and linear eignmodels for taste attrubute class (Table 1) *vs.* chemical composition class (Table 3) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: -0.4127

u_1=0.04686x_1-0.20019x_2+0.31934x_3+0.0961x_4+0.89134x_5+0.20428x_6+0.10154x_7

v_1=0.0215y_1-0.20678y_2+0.14947y_3+0.21587y_4+0.00813y_5+0.18111y_6+0.16217y_7+0.08229y_8+0.14266y_9-0.05774y_{10}+0.01542y_{11}-0.06934y_{12}+0.07574y_{13}-0.37881y_{14}+0.0051y_{15}-0.17602y_{16}+0.10964y_{17}+0.22318y_{18}+0.04404y_{19}-0.28013y_{20}-0.68449y_{21}+0.04637y_{22}

Linear regression between u_1 and v_1

u_1=-6.2815*10^{-18}-0.35265*v_1

Linear regression between v_1 and u_1

v_1=9.4197*10^{-18}-0.48309*u_1

Pearson r=-0.41275, $p\approx 0$

No.2 canonical correlation coefficient: 0.3378

u_2=0.71443x_1-0.48748x_2-0.10732x_3+0.44241x_4-0.17876x_5-0.09808x_6-0.05594x_7

v_2=0.2426y_1-0.05867y_2+0.20069y_3-0.0215y_4-0.06164y_5-0.15208y_6-0.28724y_7+0.13741y_8-0.07541y_9-0.34951y_{10}-0.18074y_{11}+0.00479y_{12}-0.15872y_{13}-0.15044y_{14}+0.17092y_{15}-0.18414y_{16}-0.15651y_{17}-0.30307y_{18}+0.31954y_{19}+0.23849y_{20}-0.46478y_{21}+0.07628y_{22}

Linear regression between u_2 and v_2

u_2=-2.9681*10^{-18}+0.2791*v_2

Linear regression between v_2 and u_2

v_2=-1.0703*10^{-18}+0.40893*u_2

Pearson r=0.33783, $p\approx 0$

No.3 canonical correlation coefficient: -0.2617

u_3=0.30342x_1+0.29661x_2-0.08946x_3-0.37942x_4+0.81177x_5-0.05866x_6-0.07469x_7

v_3=-0.0796y_1-0.22354y_2-0.08376y_3+0.14587y_4+0.2728y_5-0.13554y_6+0.31414y_7+0.10453y_8+0.03226y_9-0.07441y_{10}+0.01787y_{11}+0.26868y_{12}-0.02162y_{13}+0.38821y_{14}-0.19486y_{15}+0.23667y_{16}+0.16262y_{17}-0.00395y_{18}+0.43165y_{19}-0.25989y_{20}-0.25808y_{21}+0.20452y_{22}

Linear regression between u_3 and v_3

u_3=-2.8082*10^{-18}-0.24437*v_3

Linear regression between v_3 and u_3

v_3=3.0321*10^{-17}-0.28017*u_3

Pearson r=-0.26166, p=1.4966*10^{-13}

All canonical correlations and linear eignmodels are statistically significant with $p\leq 1.4966*10^{-13}$.

3.4 Medicinal property class *vs.* chemical composition class

The calculated canonical correlations and linear eignmodels for medicinal property class (Table 2) *vs.* chemical composition class (Table 3) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: -0.3723

u_1=-0.08775x_1-0.04354x_2+0.05009x_3+0.71101x_4+0.69452x_5

v_1=0.0989y_1+0.0559y_2+0.0374y_3-0.5289y_4-0.14064y_5-0.47068y_6+0.03507y_7+0.16535y_8+0.03438y_9-0.08874y_{10}-0.17874y_{11}-0.00847y_{12}-0.08475y_{13}+0.05523y_{14}+0.07855y_{15}-0.31922y_{16}-0.24531y_{17}-0.38085y_{18}+0.04074y_{19}+0.10758y_{20}+0.2378y_{21}-0.04105y_{22}

Linear regression between u_1 and v_1

u_1=-3.156*10^{-17}-0.40118*v_1

Linear regression between v_1 and u_1

v_1=-1.6295*10^{-17}-0.34547*u_1

Pearson r=-0.37228, $p\approx 0$

No.2 canonical correlation coefficient: -0.2817

$u_2=0.40483x_1+0.36083x_2+0.30733x_3+0.35588x_4+0.69628x_5$

$v_2=-0.12512y_1-0.01661y_2-0.32913y_3-0.32527y_4+0.10112y_5+0.07457y_6-0.00013y_7-0.00566y_8+0.09068y_9+0.01763y_{10}-0.02882y_{11}-0.00068y_{12}+0.11194y_{13}+0.25656y_{14}-0.04017y_{15}-0.71883y_{16}+0.08554y_{17}-0.23867y_{18}-0.12695y_{19}-0.01813y_{20}-0.01006y_{21}+0.25932y_{22}$

Linear regression between u_2 and v_2

$u_2=1.0498*10^{-17}-0.99578*v_2$

Linear regression between v_2 and u_2

$v_2=-3.8825*10^{-18}-0.079667*u_2$

Pearson $r=-0.28166$, $p=1.5543*10^{-15}$

No.3 canonical correlation coefficient: 0.1889

$u_3=-0.2568x_1-0.34642x_2-0.50593x_3-0.30834x_4-0.68045x_5$

$v_3=0.1908y_1-0.01518y_2+0.07877y_3-0.49439y_4+0.10412y_5-0.06303y_6+0.0986y_7+0.05814y_8+0.22266y_9+0.11985y_{10}+0.07737y_{11}-0.13853y_{12}-0.04033y_{13}-0.38498y_{14}-0.05785y_{15}-0.4866y_{16}-0.0101y_{17}+0.19452y_{18}+0.01162y_{19}-0.18955y_{20}+0.35499y_{21}-0.07564y_{22}$

Linear regression between u_3 and v_3

$u_3=-1.3625*10^{-17}+0.37177*v_3$

Linear regression between v_3 and u_3

$v_3=2.4076*10^{-18}+0.095984*u_3$

Pearson $r=0.1889$, $p=1.2381*10^{-7}$

All canonical correlations and linear eignmodels are statistically significant with $p\leq 1.2381*10^{-7}$.

3.5 Meridians and colleterals class *vs.* medicinal function class

The calculated canonical correlations and linear eignmodels for meridians and colleterals class (Table 4) *vs.* medicinal function class (Table 5) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: 0.6934

$u_1=0.49467x_1-0.15557x_2+0.15316x_3+0.01776x_4-0.16345x_5-0.44071x_6-0.36465x_7+0.17153x_8-0.15786x_9+0.19911x_{10}-0.02969x_{11}+0.50847x_{12}$

$v_1=0.13649y_1-0.09307y_2+0.13187y_3+0.29866y_4-0.00691y_5+0.03045y_6-0.04238y_7-0.00393y_8+0.05936y_9+0.0644y_{10}-0.03536y_{11}-0.07812y_{12}-0.0772y_{13}+0.01804y_{14}-0.05598y_{15}+0.25528y_{16}-0.12064y_{17}-0.24571y_{18}-0.07139y_{19}-0.04577y_{20}-0.13601y_{21}-0.15049y_{22}-0.07495y_{23}-0.18769y_{24}-0.00838y_{25}+0.06461y_{26}+0.11783y_{27}+0.01996y_{28}-0.13667y_{29}-0.05599y_{30}-0.03833y_{31}-0.01582y_{32}-0.04764y_{33}-0.09478y_{34}+0.07603y_{35}-0.00203y_{36}+0.09609y_{37}+0.0487y_{38}+0.0661y_{39}-0.00396y_{40}+0.18242y_{41}-0.00418y_{42}+0.07974y_{43}-0.01147y_{44}+0.11266y_{45}-0.06142y_{46}-0.01383y_{47}-0.03661y_{48}-0.12066y_{49}-0.39311y_{50}+0.04015y_{51}-0.23655y_{52}+0.03633y_{53}+0.08624y_{54}-0.00645y_{55}-0.00171y_{56}-0.02411y_{57}+0.05085y_{58}-0.08999y_{59}-0.02453y_{60}-0.0539y_{61}-0.02482y_{62}+0.06042y_{63}+0.07667y_{64}-0.10442y_{65}-0.31382y_{66}-0.01055y_{67}-0.19222y_{68}+0.00545y_{69}+0.0016y_{70}-0.02364y_{71}-0.05152y_{72}+0.1043y_{73}+0.1241y_{74}-0.1475y_{75}+0.12906y_{76}+0.0445y_{77}$

Linear regression between u_1 and v_1

$u_1=1.0577*10^{-17}+0.28372*v_1$

Linear regression between v_1 and u_1

$v_1=-1.1712*10^{-17}+1.6948*u_1$

Pearson $r=0.69344$, $p\approx 0$

No.2 canonical correlation coefficient: 0.642

$u_2=-0.02594x_1+0.34996x_2+0.13919x_3+0.7065x_4-0.47434x_5+0.25881x_6+0.0498x_7-0.11854x_8+0.15169x_9-0.02658x_{10}-0.14876x_{11}+0.06305x_{12}$

v_2=-0.05718y_1-0.09725y_2+0.11827y_3+0.03042y_4+0.33338y_5+0.07635y_6+0.12309y_7+0.12355y_8+0.1136y_9-0.08907y_{10}-0.16118y_{11}-0.11538y_{12}-0.14925y_{13}-0.06186y_{14}-0.16157y_{15}-0.09974y_{16}-0.02988y_{17}+0.11625y_{18}+0.07177y_{19}-0.01486y_{20}+0.00217y_{21}+0.07963y_{22}+0.02743y_{23}+0.14218y_{24}-0.18586y_{25}-0.0417y_{26}-0.11175y_{27}+0.04257y_{28}+0.07794y_{29}-0.02564y_{30}+0.18975y_{31}-0.05725y_{32}+0.05317y_{33}+0.15256y_{34}+0.00541y_{35}-0.04797y_{36}+0.11159y_{37}+0.00752y_{38}-0.0132y_{39}-0.00668y_{40}+0.11641y_{41}-0.16541y_{42}-0.03908y_{43}+0.06415y_{44}+0.06334y_{45}+0.05264y_{46}-0.06521y_{47}-0.02699y_{48}-0.0349y_{49}+0.34454y_{50}+0.07858y_{51}-0.02831y_{52}-0.00941y_{53}+0.29596y_{54}-0.05626y_{55}+0.03358y_{56}-0.04869y_{57}-0.0028y_{58}+0.00304y_{59}-0.09212y_{60}+0.06052y_{61}+0.09518y_{62}-0.0672y_{63}+0.0333y_{64}+0.1653y_{65}-0.08444y_{66}-0.28706y_{67}-0.08007y_{68}-0.05711y_{69}-0.12115y_{70}+0.098y_{71}-0.05477y_{72}-0.04852y_{73}-0.00402y_{74}-0.06633y_{75}-0.09029y_{76}+0.20663y_{77}

Linear regression between u_2 and v_2

$u_2=4.2019*10^{-18}+0.27965*v_2$

Linear regression between v_2 and u_2

$v_2=-2.3996*10^{-17}+1.474*u_2$

Pearson r=0.64203, $p\approx 0$

No.3 canonical correlation coefficient: -0.5807

u_3=0.15424x_1+0.2169x_2-0.24008x_3-0.45988x_4-0.38404x_5+0.1403x_6+0.15062x_7+0.22277x_8+0.17801x_9-0.44047x_{10}-0.41109x_{11}+0.16079x_{12}

v_3=-0.01495y_1+0.09547y_2-0.11743y_3-0.12479y_4+0.26804y_5+0.18104y_6+0.14225y_7+0.02223y_8+0.03199y_9+0.01492y_{10}+0.14704y_{11}+0.12018y_{12}+0.14377y_{13}+0.06358y_{14}+0.1551y_{15}-0.11239y_{16}+0.00549y_{17}-0.07201y_{18}-0.12257y_{19}+0.06791y_{20}+0.07759y_{21}-0.10637y_{22}-0.02166y_{23}-0.09282y_{24}-0.04642y_{25}-0.08974y_{26}-0.29102y_{27}-0.00671y_{28}-0.11823y_{29}+0.10677y_{30}-0.08302y_{31}+0.04944y_{32}-0.08483y_{33}+0.00079y_{34}-0.05164y_{35}-0.02157y_{36}-0.01083y_{37}-0.0271y_{38}-0.04485y_{39}+0.01506y_{40}-0.08502y_{41}-0.3371y_{42}-0.04491y_{43}+0.07967y_{44}+0.03352y_{45}+0.05982y_{46}-0.03551y_{47}+0.03506y_{48}+0.00255y_{49}-0.4097y_{50}+0.00839y_{51}-0.01087y_{52}-0.03877y_{53}-0.00613y_{54}+0.03612y_{55}+0.11123y_{56}+0.02982y_{57}+0.08211y_{58}+0.01416y_{59}-0.09318y_{60}+0.09175y_{61}+0.00523y_{62}+0.04333y_{63}+0.03289y_{64}-0.11441y_{65}-0.20519y_{66}+0.16499y_{67}+0.07476y_{68}-0.04746y_{69}-0.01706y_{70}-0.00924y_{71}-0.00954y_{72}-0.06098y_{73}-0.00212y_{74}+0.00872y_{75}+0.05784y_{76}+0.31212y_{77}

Linear regression between u_3 and v_3

$u_3=5.5161*10^{-18}-0.24472*v_3$

Linear regression between v_3 and u_3

$v_3=-3.3233*10^{-17}-1.3781*u_3$

Pearson r=-0.58073, $p\approx 0$

All canonical correlations and linear eignmodels are statistically significant with $p\leq 10^{-15}$.

3.6 Taste attribute class *vs.* medicinal function class

The calculated canonical correlations and linear eignmodels for taste attribute class (Table 1) *vs.* medicinal function class (Table 5) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: -0.6338

u_1=0.28926x_1-0.82418x_2+0.31268x_3+0.0237x_4+0.04424x_5+0.3203x_6-0.18486x_7

v_1=-0.0332y_1-0.26233y_2+0.02572y_3-0.08497y_4+0.02954y_5-0.0785y_6-0.04282y_7-0.01083y_8-0.02063y_9-0.15673y_{10}-0.07291y_{11}+0.08505y_{12}+0.04256y_{13}-0.09672y_{14}+0.31969y_{15}-0.01848y_{16}+0.05785y_{17}-0.05413y_{18}-0.0534y_{19}-0.12881y_{20}+0.02541y_{21}-0.0064y_{22}+0.25388y_{23}+0.07963y_{24}+0.00418y_{25}-0.01056y_{26}+0.23591y_{27}-0.09371y_{28}-0.08141y_{29}-0.0506y_{30}-0.05715y_{31}+0.03276y_{32}-0.15818y_{33}-0.00685y_{34}-0.02566y_{35}-0.01038y_{36}-0.15766y_{37}+0.02167y_{38}+0.0509y_{39}+0.07519y_{40}-0.11402y_{41}+0.15375y_{42}-0.20422y_{43}-0.03689y_{44}-0.05402y_{45}+0.00558y_{46}+0.01445y_{47}+0.09042y_{48}+0.15415y_{49}-0.1823y_{50}-0.03832y_{51}+0.12012y_{52}+0.04054y_{53}-0.20563y_{54}-0.08429y_{55}-0.04256y_{56}-0.00105y_{57}-0.04563y_{58}-0.05054y_{59}-0.05401y_{60}+0.18851y_{61}+0.03768y_{62}+0.06017y_{63}+0.07242y_{64}-0.01704y_{65}-0.16795y_6

$_{6}+0.19228y_{67}+0.11497y_{68}-0.20792y_{69}-0.04998y_{70}+0.07441y_{71}+0.01772y_{72}+0.07639y_{73}+0.0836y_{74}+0.19523y_{75}+0.27443y_{76}+0.0243y_{77}$

Linear regression between u_1 and v_1

$u_1=1.761*10^{-18}-0.21171*v_1$

Linear regression between v_1 and u_1

$v_1=-2.1662*10^{-17}-1.8971*u_1$

Pearson $r=-0.63375$, $p\approx 0$

No.2 canonical correlation coefficient: 0.5504

$u_2=0.18308x_1-0.1924x_2-0.18029x_3+0.21799x_4-0.44698x_5-0.7372x_6+0.32587x_7$

$v_2=-0.01159y_1-0.00954y_2+0.07119y_3-0.13492y_4-0.03786y_5+0.00955y_6+0.0577y_7-0.04274y_8+0.01538y_9-0.04336y_{10}-0.01726y_{11}-0.02446y_{12}+0.08333y_{13}-0.02469y_{14}-0.0213y_{15}+0.03696y_{16}-0.08658y_{17}-0.05291y_{18}-0.04127y_{19}-0.09463y_{20}-0.04701y_{21}-0.05572y_{22}-0.15381y_{23}+0.05231y_{24}+0.05945y_{25}+0.02736y_{26}+0.04318y_{27}+0.05327y_{28}+0.02528y_{29}-0.08415y_{30}-0.57102y_{31}+0.0857y_{32}-0.16753y_{33}-0.25775y_{34}+0.00926y_{35}-0.09308y_{36}-0.1204y_{37}+0.03288y_{38}-0.04416y_{39}-0.00316y_{40}-0.12331y_{41}+0.08424y_{42}+0.04922y_{43}+0.08336y_{44}-0.00195y_{45}+0.12733y_{46}-0.12671y_{47}+0.02716y_{48}-0.00848y_{49}+0.2424y_{50}+0.05544y_{51}-0.01649y_{52}+0.0441y_{53}-0.05354y_{54}+0.09121y_{55}+0.0561y_{56}-0.02928y_{57}+0.18367y_{58}-0.14679y_{59}-0.01526y_{60}-0.05496y_{61}+0.00059y_{62}-0.01864y_{63}-0.05203y_{64}-0.03578y_{65}-0.06965y_{66}-0.04654y_{67}-0.02516y_{68}-0.43329y_{69}+0.03302y_{70}-0.04909y_{71}+0.06264y_{72}+0.07615y_{73}-0.08262y_{74}-0.0565y_{75}-0.03334y_{76}+0.08524y_{77}$

Linear regression between u_2 and v_2

$u_2=-1.0851*10^{-18}+0.28397*v_2$

Linear regression between v_2 and u_2

$v_2=-1.1615*10^{-18}+1.0668*u_2$

Pearson $r=0.5504$, $p\approx 0$

No.3 canonical correlation coefficient: -0.4956

$u_3=0.28106x_1-0.03992x_2-0.52959x_3+0.35233x_4+0.50751x_5+0.44897x_6+0.23594x_7$

$v_3=-0.01795y_1-0.17773y_2-0.03254y_3+0.43084y_4+0.058y_5+0.0825y_6-0.02927y_7+0.08337y_8-0.02811y_9+0.15236y_{10}+0.08615y_{11}+0.03257y_{12}-0.03044y_{13}-0.03063y_{14}-0.01008y_{15}+0.11275y_{16}+0.03494y_{17}+0.16997y_{18}+0.01069y_{19}+0.03256y_{20}-0.12734y_{21}-0.02827y_{22}+0.0982y_{23}+0.05541y_{24}-0.01359y_{25}+0.05573y_{26}-0.07746y_{27}+0.03008y_{28}+0.11028y_{29}+0.20428y_{30}-0.34813y_{31}+0.02019y_{32}-0.09972y_{33}-0.12497y_{34}-0.0139y_{35}+0.00632y_{36}+0.01415y_{37}-0.00601y_{38}-0.03162y_{39}-0.0257y_{40}+0.26903y_{41}-0.18881y_{42}+0.02997y_{43}+0.1309y_{44}+0.00696y_{45}+0.15326y_{46}+0.10719y_{47}-0.00191y_{48}+0.01541y_{49}+0.17711y_{50}-0.05167y_{51}+0.05782y_{52}-0.05412y_{53}+0.10498y_{54}+0.00151y_{55}-0.00779y_{56}-0.02377y_{57}-0.09778y_{58}+0.00166y_{59}+0.09903y_{60}+0.05303y_{61}-0.0019y_{62}+0.03244y_{63}+0.04623y_{64}-0.13736y_{65}-0.1187y_{66}-0.00533y_{67}-0.00853y_{68}-0.2286y_{69}-0.05422y_{70}-0.16049y_{71}-0.03008y_{72}+0.01754y_{73}+0.21378y_{74}+0.14244y_{75}-0.13231y_{76}+0.06702y_{77}$

Linear regression between u_3 and v_3

$u_3=-2.1028*10^{-17}-0.17214*v_3$

Linear regression between v_3 and u_3

$v_3=-5.4882*10^{-18}-1.4269*u_3$

Pearson $r=-0.4956$, $p\approx 0$

All canonical correlations and linear eignmodels are statistically significant with $p\leq 10^{-15}$.

3.7 Canonical correlation network of attribute classes

Similar to my previous research (Zhang, 2017b), I constructed the canonical correlation network of attribute classes, as indicated in Fig. 1. In the network, only the greatest Pearson correlations are labeled, and the corresponding canonical correlation functions are listed in the sections above. All canonical correlations in the network are statistically significant ($p \leq 10^{-5}$).). Different from Zhang (2017b), the canonocal correlation network takes attribute classes as nodes and represents the correlations between attribute classes。

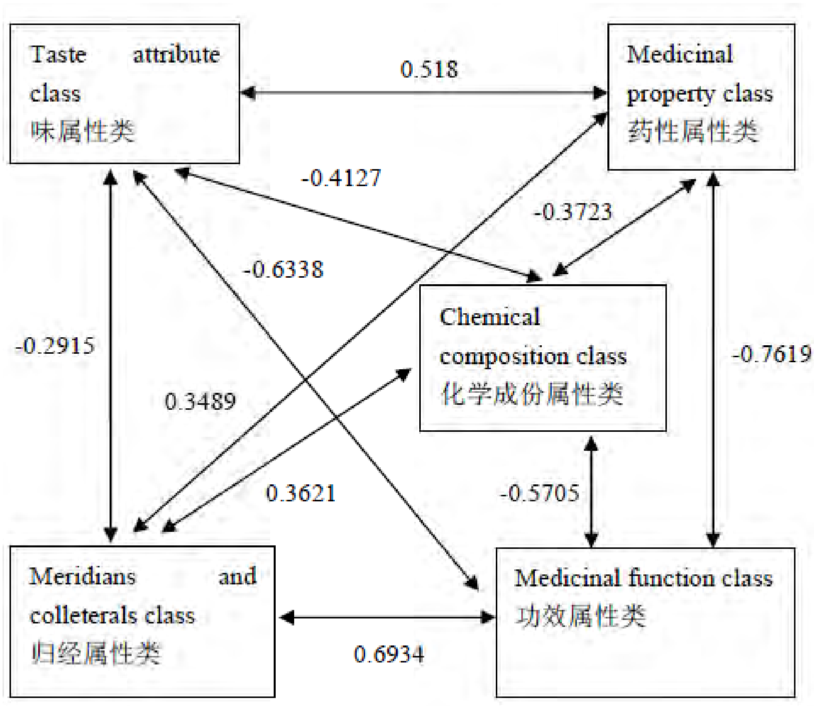

Fig. 1 Canonocal correlation network of attribute classes.

4 Discussion

Similar to between-attribute correlation, the correlation between attribute classes may mostly tend to be the quasi-linear correlation, or nonlinear correlation (Zhang, 2012, 2015), as seen by the linear correlation coefficients above. On the other hand, taking into account extreme complexity of quasi-linear correlation or nonlinear correlation between attribute classes, linear correlation is a reasonable approximation to quasi-linear correlation and nonlinear correlation.

It should be noted that the canonical correlation function is an optimal linear combination of attributes (variables) to maximize the fitting goodness. Thus, they, as well as the derived linear eignmodels from them, may be used as empirical models, but not as mechanism models. Its coefficient and sign are not suggested being used as importance or positive / negative interactions of attributes.

In a certain sense, the linear eignmodel serves more as a pattern of attributes of Chinese herbal medicines, rather than a predictive model. To predict the herbal attributes, other methods such as regression, discriminant analysis, etc (Qi, 2006; Liu et al., 2014) can be used.

Acknowledgment

We are thankful to the support of Discovery and Crucial Node Analysis of Important Biological and Social Networks (2015.6-2020.6), from Yangling Institute of Modern Agricultural Standardization, China.

References

Liu N, Li J, Li BG. 2014. Application of multivariable statistical analysis and thinking in quality control of Chinese medicine. China Journal of Chinese Materia Medica, 39(21): 4268-4271

Qi YH. 2006. A web computational software for stepwise discrimination analysis in information recognition. Journal of Information, 11: 64-65

Qi YH, Xu LH. 2009. Web Implementation of Canonical Correlation Analysis and Its Applications in Information Researches. Journal of Modern Information, 29(1): 134-139, 143

Budovsky A, Fraifeld VE. 2012. Medicinal plants growing in the Judea region: network approach for searching potential therapeutic targets. Network Biology, 2(3): 84-94

Hopkins AL. 2007. Network pharmacology. Nature Biotechnology, 25(10): 1110-1111

Hopkins AL. 2008. Network pharmacology: the next paradigm in durg discovery. Nature Chemical Biology, 4(11): 682-690

Zhang WJ. 2012. Computational Ecology: Graphs, Networks and Agent-based Modeling. World Scientific, Singapore

Zhang WJ. 2015. Calculation and statistic test of partial correlation of general correlation measures. Selforganizology, 2(4): 65-77

Zhang WJ. 2016. Network pharmacology: A further description. Network Pharmacology, 1(1): 1-14

Zhang WJ. 2017a. Network pharmacology of medicinal attributes and functions of Chinese herbal medicines: (I) Basic statistics of medicinal attributes and functions for more than 1100 Chinese herbal medicines. Network Pharmacology, 2(2): 17-37

Zhang WJ. 2017b. Network pharmacology of medicinal attributes and functions of Chinese herbal medicines: (II) Relational networks and pharmacological mechanisms of medicinal attributes and functions of Chinese herbal medicines. Network Pharmacology, 2(2): 38-66

Zhang YT, Fang KT. 1982. Introduction to Multivariable Analysis. Science Press, Beijing, China