

Article

Pattern classification of HLA-DRB1 alleles, human races and populations: Application of self-organizing competitive neural network

WenJun Zhang^{1,2}, YanHong Qi³

¹School of Life Sciences, Sun Yat-sen University, Guangzhou, China; ²International Academy of Ecology and Environmental Sciences, Hong Kong

³Libraries of Sun Yat-sen University, Sun Yat-sen University, Guangzhou, China

E-mail: zhwj@mail.sysu.edu.cn, wjzhang@iaees.org

Received 29 May 2014; Accepted 6 July 2014; Published online 1 December 2014



Abstract

HLA-DRB1 gene is concerned with human immune systems. There are about 300 alleles of HLA-DRB1 gene. The self-organizing competitive neural network was used in present study to make non-supervisory classification on 14 HLA-DRB1 alleles, and 54 human races and populations (Zhang and Qi, 2005). It was found that HLA-DRB1-0901 and 1402 are similar to each other in the distribution in human races and populations. There were higher similarity between HLA-DRB1-0101 and 0302, and between HLA-DRB1-0701 and 0301. The results showed that there were significant differences among the various races and there were similarities among populations in the same race. South America Indians and Siberians are highly similar to each other. There was relatively significant difference between Northern Chinese and Southern Chinese. Han Chinese in Guangdong was similar to ethnic minorities such as the Lahu Chinese and Yao Chinese.

Keywords self-organizing competitive neural network; pattern classification; HLA-DRB1 alleles; human races and populations.

Selforganizology

URL: <http://www.iaees.org/publications/journals/selforganizology/online-version.asp>

RSS: <http://www.iaees.org/publications/journals/selforganizology/rss.xml>

E-mail: selforganizology@iaees.org

Editor-in-Chief: WenJun Zhang

Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

HLA is the most complex human genome, which locates on 6th human chromosome. The genome features a total of 128 genes and 96 pseudogenes (Jia, 2001). Among which there are 293 HLA-DRB1 alleles, which mainly determine the antigenic immune functions of human. A comparative study of HLA-DRB1 alleles will help trace the origin of mankind, migration, integration of history, and develop group-specific bio-engineering drugs. Self-organizing competitive neural network has been attracting attentions in various areas (Zhang and Qi, 2005; Zhang, 2007, 2010). The network can identify the mechanism and relationship from the input information and adjust the network for better adaptation. It is suitable for unsupervised pattern classification.

The self-organizing competitive neural network was used in present study to make non-supervisory classification on 14 HLA-DRB1 alleles, and 54 human races and populations (Zhang and Qi, 2005), in order to understand the relationship between HLA-DRB1 alleles and obtain some results on human races and populations.

2 Method

The principle of the self-organizing competitive network is to input a model vector, let nodes in output layer compete according to some given rules. If a node wins the competition, then adjust the weight structure, so make the winning node more sensitive to this vector model, while the other nodes are suppressed and they are hard to win under this vector model.

Matlab source codes of the algorithm are:

```
P=HLA_DRB1(:,:);
net=newc(minmax(P),10,0.01,0.001); %Set number of neurons as 10
net.trainParam.epochs=2000; %Set training epochs as 2000
net=init(net);
net=train(net,P);
w=net.iw{1}
a=0;
for i=1:size(P,2);
a=vec2ind(sim(net,P(:,i)));
outputclass(1,i)=i;
outputclass(2,i)=a;
end
outputclass
```

3 Results

Data of the world's 54 human races and populations and 14 common HLA-DRB1 alleles can be found in Jia (2001). First, analyze the similarities between HLA-DRB1 alleles according to the results of pattern classification of self-organizing competitive neural network. Then, the 54 human races and populations were classified using self-organizing competitive neural network, based on HLA-DRB1 allele polymorphism.

3.1 Pattern classification of HLA-DRB1 alleles

It was found that the more neurons in the neural network, the more classes we can obtain (Table 1). Thus the classification can be analyzed at different scales.

Main conclusions of HLA-DRB1 alleles' pattern classification are as follows:

- (1) At different scales, HLA-DRB1-0901 and 1402 show a stable similarity. On large-scale classification, HLA-DRB1-0901, 1402, and 1401 and the remaining 11 alleles belonging to two different classes, there are obvious differences between them.
- (2) At different scales, there is a strong similarity between HLA-DRB1-0101 and 0302. Both of them show a strong similarity to 1104 at mediate or larger scale. At the small-scale classification, 0101 and 0405 have a strong correlation.
- (3) At the mediate scale, HLA-DRB1-0701 and 0301 have a high similarity.
- (4) At certain scales, 1104 and 1502 are also significant associated.

Table 1 Pattern classification of HLA-DRB1 alleles using self-organizing competitive neural network.

14 neurons	13 neurons	12 neurons	11 neurons	10 neurons	9 neurons	8 neurons	7 neurons	6 neurons	5 neurons	4 neurons	3 neurons	2 neurons	1 neuron
0302	0302	1401	0101	0101	0101	0101	1104	1104	1104	1104	1202	1202	1202
0901	1401	0101	0302	0302	0302	0302	0101	0101	0101	0101	1501	1501	0901
1501	1602	0302	0701	1402	1401	1401	0302	0302	1602	1602	0701	0701	1501
0701	0803	1402	1402	1401	1402	0901	1401	1401	0302	0302	1104	1104	0701
1402	1402	0901	1401	0803	0803	1402	0901	0901	1401	0301	0101	0101	1104
0803	0901	0803	0803	0701	0701	0803	1402	1402	0901	1401	1502	1502	0101
1202	0701	0701	1501	0301	0301	0701	0803	1502	1402	0901	1602	1602	1402
1401	0301	1602	1602	1501	1602	0301	0701	0803	1502	1402	0803	0803	1401
1502	1104	0301	1202	0405	1202	1602	0301	1501	0803	1202	0405	0405	1502
0101	1502	1104	0901	1602	0405	0405	1501	0701	0405	1501	0302	0302	1602
0405	0101	1202	1104	1202	0901	1202	1602	0301	1202	0701	0301	0301	0803
1104	0405	1501	0301	0901	1501	1501	1202	1202	1501	1502	1401	0901	0405
0301	1501	1502	1502	1104	1104	1104	1502	1602	0701	0803	0901	1402	0302
1602	1202	0405	0405	1502	1502	1502	0405	0405	0301	0405	1402	1401	0301

3.2 Pattern classification of human races and populations

Some conclusions for pattern classification of human races and populations are listed as below (Table 2).

(1) Overall, there are significant differences among the various races and there are similarities among populations in the same race. For example, there are relatively strong similarities between populations of the following races:

- Siberian populations;
- Australia's native populations;
- Black populations;
- South American Indian populations;
- Jewish populations;
- Japanese populations;
- European and American whites.

(2) South American Indians and Siberians are highly similar to each other, which is coincident with the conclusion that Amerindian were from Siberia.

(3) There was relatively significant difference between Northern Chinese and Southern Chinese. There are significant differences between ethnic minorities in Southern China. Han Chinese in Guangdong was similar to ethnic minorities such as the Lahu Chinese and Yao Chinese. Northern Han and Manchu are highly similar. Hunan Han and Singapore Han are highly similar to each other.

(4) There is a high similarity between Pumi and Australia's native populations, and a remarkable similarity between Thais and the Dai in China.

(5) At the large-scale classification, Japanese populations and Australia's native populations are highly similar; at the mediate/small scale, Japanese populations are similar to Northern Han and some ethnic minorities.

(6) Siberian Kets population is somewhat different from other Siberian populations.

(7) The following races/populations are highly similar to each other:

Geeks, Macedonians, Iranian Jews;
 Romanians, Turks;
 American whites, Spanish, German, Polish.

Table 2 Pattern classification of human races and populations using self-organizing competitive neural network.

20 neurons	16 neurons	12 neurons	8 neurons	4 neurons
Pumi-China	Pumi-China	Siberian Evenki population	Uighur-China	Uighur-China
Japanese	Native population-Australia's central desert	Siberian Kets population	Siberian Kets population	Kazak-China
Hokkaido-Japan	Kimberley native population-Australia	North American blacks	USA whites	Siberian Kets population
Romanians	Cape York native population-Australia	South African blacks	Spanish	USA whites
Turks	Romanians	Uighur-China	German	Spanish
Native population-Australia's central desert	Turks	USA whites	Romanians	German
Kimberley native population-Australia	Israeli Arabs	Spanish	Polish	Romanians
Cape York native population-Australia	Liaoning Han-China	German	Ethiopian Jews	Bulgarian
Liaoning Han-China	Northwest Han-China	Polish	North American blacks	Greek
Northern Han-China	Northern Han-China	Kazak-China	Dulong-China	Polish
Manchu-China	Siberian Kets population	Romanians	Kazak-China	Turks
Siberian Evenki population	North American blacks	Bulgarian	Bulgarian	Macedonians
North American blacks	South African blacks	Turks	Turks	Israeli Arabs
South African blacks	Siberian Eskimo	Ethiopian Jews	Macedonians	Iranian Jews
Siberian Kets population	Siberian Chukchi population	Pumi-China	South African blacks	Ashkenazi Jews-Germany
Polish	Siberian Evenki population	Native population-Australia's central desert	Siberian Nivkhs population	Libyan Jews
Siberian Koryaks population	USA whites	Yuendumu Native population-Australia	Siberian Udegeys population	Yemeni Jews
Siberian Eskimo	Spanish	Kimberley native population-Australia	Siberian Koryaks population	Moroccan Jews
Siberian Chukchi population	German	Cape York native population-Australia	Siberian Eskimo	Ethiopian Jews
Dai-China	Polish	Buyi-China	Siberian Chukchi population	North American blacks
Thais	Greek	Siberian Eskimo	Pumi-China	South African blacks
Shanghai Han-China	Macedonians	South American Indians	Japanese	Dulong-China
Shenyang Han-China	Iranian Jews	Ticuna	Native population-Australia's central desert	Buyi-China
Northwest Han-China	Dulong-China	South American Indians	Terena	
Uighur-China	Yuendumu Native population-Australia	Dulong-China	Yuendumu Native population-Australia	Pumi-China
USA whites	Buyi-China	Greek	Kimberley native population-Australia	Japanese
Spanish	South American Indians	Macedonians	Cape York native population-Australia	Hokkaido-Japan
	Ticuna	Iranian Jews	Lahu-China	Native population-Australia's central desert

German	South American Indians Terena	Siberian Nivkhs population	Yao-China	Yuendumu Native population-Australia
Naxi-China	Siberian Nivkhs population	Siberian Udegeys population	Guangdong Han-China	Kimberley native population-Australia
Yi-China	Siberian Udegeys population	Siberian Koryaks population	Hunan Han-China	Cape York native population-Australia
Kazak-China	Siberian Koryaks population	Siberian Chukchi population	Southern Han-China	Lahu-China
Bulgarian	Yi-China	Lahu-China	Singapore Han-Singapore	Dai-China
Ethiopian Jews	Manchu-China	Yao-China	Shanghai Han-China	Naxi-China
Siberian Nivkhs population	Japanese	Guangdong Han-China	Dai-China	Yao-China
Siberian Udegeys population	Hokkaido-Japan	Hunan Han-China	Naxi-China	Guangdong Han-China
Lahu-China	Lahu-China	Southern Han-China	Buyi-China	Thais
Yao-China	Naxi-China	Singapore Han-Singapore	Thais	Yi-China
Guangdong Han-China	Yao-China	Shanghai Han-China	South American Indians Ticuna	Hunan Han-China
Hunan Han-China	Guangdong Han-China	Liaoning Han-China	South American Indians Terena	Southern Han-China
Southern Han-China	Hunan Han-China	Shenyang Han-China	Yi-China	Singapore Han-Singapore
Singapore Han-Singapore	Singapore Han-Singapore	Northwest Han-China	Liaoning Han-China	Shanghai Han-China
Dulong-China	Shanghai Han-China	Dai-China	Shenyang Han-China	Liaoning Han-China
Yuendumu Native population-Australia	Shenyang Han-China	Naxi-China	Northwest Han-China	Shenyang Han-China
Israeli Arabs	Dai-China	Thais	Northern Han-China	Northwest Han-China
Ashkenazi Jews-Germany	Thais	Yi-China	Manchu-China	Northern Han-China
Libyan Jews	Southern Han-China	Northern Han-China	Hokkaido-Japan	Manchu-China
Yemeni Jews	Uighur-China	Manchu-China	Siberian Evenki population	Siberian Nivkhs population
Moroccan Jews	Kazak-China	Japanese	Greek	Siberian Udegeys population
Buyi-China	Bulgarian	Hokkaido-Japan	Israeli Arabs	Siberian Koryaks population
South American Indians Ticuna	Ethiopian Jews	Israeli Arabs	Iranian Jews	Siberian Eskimo
South American Indians Terena	Ashkenazi Jews-Germany	Ashkenazi Jews-Germany	Ashkenazi Jews-Germany	Siberian Chukchi population
Greek	Libyan Jews	Libyan Jews	Libyan Jews	South American Indians Ticuna
Macedonians	Yemeni Jews	Yemeni Jews	Yemeni Jews	South American Indians Terena
Iranian Jews	Moroccan Jews	Moroccan Jews	Moroccan Jews	Siberian Evenki population

References

- Jia ZJ. 2001. Polymorphism of HLA-DRB1 gene in southern Chinese populations. PhD Thesis. 46-47, Sun Yat-sen University, Guangzhou, China
- Zhang WJ. 2010. Computational Ecology: Artificial Neural Networks and Their Applications. World Scientific, Singapore
- Zhang WJ. 2007. Pattern classification and recognition of invertebrate functional groups using self-organizing neural networks. Environmental Monitoring and Assessment, 130: 415-422
- Zhang WJ, Qi YH. 2005. Pattern classification of HLA-DRB1 alleles and human races using Self-Organizing Competitive Neural Network. Modern Computer, 4: 10-13, 30