Jpn. J. Human Genet. **33**, 385–387, 1988

Letter to the Editor

HLA Polymorphism Information Content (PIC)

To the Editor:

Genetically polymorphic loci can be tested for linkage relationships in human pedigrees. While restriction fragment length polymorphisms (RFLPs) are commonly used as genetic markers, HLA complex are also currently employed for establishing linkage with disease in question (*e.g.* Carroll *et al.*, 1987). Pedigrees in which inherited traits are known to be segregating can then be analyzed, making possible the mapping of the gene(s) responsible for the trait with respect to HLA haplotype on chromosome 6.

For constructing gene map, the polymorphism information content (PIC) developed by Botstein *et al.* (1980) is the one most often used for linkage analysis. The PIC represents the probability that a given offspring of a random mating between a carrier of a rare dominant gene and a noncarrier is informative for linkage between the locus of the dominant gene and a codominant marker. The PIC may be interpreted in the following manner. If grandparents are not examined for the first child, the PIC is the probability of establishing linkage phase. Once phase is established, it is the probability that an offspring is informative for accepting or rejecting linkage.

PIC could be shown in terms of power-sum (Vasuda, 1986) as

$$\mathrm{PIC}=1-\mathrm{S}_2-\mathrm{S}_2{}^2+\mathrm{S}_4 \qquad (1)$$

where $\mathrm{S}_k=\sum_{i=1}^{m}\mathrm{p}_i{}^k$ or the k-th power-sum, m is the number of allele at a polymorphic locus and p_i is the allelic frequency.

While the formula (1) has been developed for the codominant system, HLA loci contain unidentified gene due to the lack of good typing reagents, especially in early days of HLA typing. A new formula may therefore be desirable in order that could assess the effect of recessiveness on PIC in genetic marker as HLA or ABO-like systems. In this case, the PIC is given by

$$\mathrm{PIC}=(1-2r+2r^2+r^3)\mathrm{S}_1'-(1-r)^2\mathrm{S}_2'+r\mathrm{S}_3'-(\mathrm{S}_2')^2+\mathrm{S}_4' \qquad (2)$$

where r is the frequency of recessive gene and $\mathrm{S}_k'=\sum_{i=1}^{m-1}\mathrm{p}_i{}^k$ and $m-1$ is the number of codominantly detectable antigens. When $r\to 0$, then $\mathrm{S}_k'\to\mathrm{S}_k$, the new formula (2) reduces to (1).

In order to see the influence of the presence of silent gene on PIC, we tabulated it for some values of r and $m-1$ setting equal frequency of codominant alleles. This situation gives the maximum value for PIC (max PIC). In Table 1, the figures in

Table 1. Maximum polymorphism information content (PIC) in ABO-like system.

m-1\r	0	0.02	0.05	0.10	0.20	0.30	0.40	0.50
1	0	.038	.088	.156	.243	.288	.307	.313
2	.375	.370	.361	.347	.320	.301	.291	.289
3	.593	.568	.533	.480	.395	.339	.307	.294
4	.703	.669	.621	.549	.435	.360	.317	.298
5	.768	.728	.673	.590	.459	.373	.324	.301
6	.810	.767	.706	.616	.475	.382	.328	.302
7	.840	.794	.730	.635	.487	.389	.331	.304
8	.861	.814	.748	.649	.495	.393	.334	.305
9	.878	.829	.761	.660	.502	.397	.336	.306
10	.891	.841	.771	.668	.506	.399	.337	.306
100	.989	.932	.851	.732	.546	.422	.349	.311
200	.994	.936	.855	.735	.548	.423	.349	.312

r=frequency of recessive gene. $m-1$=the number of codominant allele.

Table 2. HLA-PIC among Japanese.

Locus	No. of detected allele	Freq. of null (r)	Obs. PIC	Max PIC	(Obs./Max) ×PIC (%)
A	8	.00	.68	.86	79
B	19	.02	.82	.94	87
BW4-W6	2	.00	.34	.38	89
C	7	.39	.32	.34	94
DR	9	.05	.72	.77	93
DQ	4	.21	.35	.42	82
DP	3	.34	.32	.32	99
D	9	.22	.46	.47	97

column with no recessive gene (r=0) indicate the max PIC for $m-1$ codominant alleles or max PIC=$(1-1/m')^2(1+1/m')$ for $m'=m-1$. Increase of silent gene frequency reduces max PIC. Even if the number of codominant alleles ($m-1$) was 200, which gives more than 99% value of max PIC, and such situation might encounter with DNA minisatellite polymorphism (DNA finger prints, Jeffreys *et al.*, 1985), max PIC (.349) with .4 frequency of silent allele is less than the max PIC with two codominant system (.375).

Actual PICs in HLA system have been calculated for Japanese, based on gene frequency data reported by the Third Asia-Oceania Workshop (Aizawa, 1986). Only main antigens were taken into consideration. The results were summarized in Table 2. Max PIC for each locus was obtained for the observed frequency of silent

allele, derived from one minus the proportion of sum of detectable antigens, assuming equal frequency of all codominant antigens. The ratio observed to max PIC shows relative effectiveness of maker locus.

The influence of the proportion of silent allele is obvious. The lower the frequency of silent gene is, the higher the value of PIC is observed. As predicted, an increase of number of codominant allele also contributes to a high value of PIC.

It is concluded that HLA-loci, especially, A, B and DR, are indeed good genetic markers for linkage study with disease in question. Splitting antigens will lead HLA system to be more informative.

REFERENCES

Aizawa, M. 1986. HLA in Asia-Oceania. *Proceedings of the Third Asia-Oceania Histocompatibility Workshop and Conference.* pp. 1080–1081.

Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**: 314–331.

Carroll, C., Katzman, P., Alicot, E.M., Koller, B.E., Geraghty, D.E., Orr, H.T., Strominger, J.L., and Spies, T. 1987. Linkage map of the human major histocompatibility complex including the tumor necrosis factor genes. *Proc. Natl. Acad. Sci. USA* **84**: 8535–8539.

Jeffreys, A.J., Wilson, V., and Thein, S.L. 1985. Hypervariable 'minisatellite' regions in human DNA. *Nature* **314**: 67–73.

Yasuda, N. 1986. Handy formula for calculating the probability of parentage exclusion. *Jpn. J. Human Genet.* **31**: 113–114.

Norikazu YASUDA, Ph. D.

Division of Genetics,
National Institute of
Radiological Sciences,
Chiba 260
(Received January 29, 1988:
revised version received
March 7, 1988)