

# GENETIC STRUCTURE OF HUMAN POPULATIONS

## II. DIFFERENTIATION OF BLOOD GROUP GENE FREQUENCIES AMONG ISOLATED POPULATIONS\*

MASATOSHI NEI and YOKO IMAIZUMI

*Division of Genetics, National Institute of Radiological Sciences, Chiba, Japan*

Received 25.vii.65

### 1. INTRODUCTION

NEI and Imaizumi (1966) have shown that there is a significant local differentiation in the ABO and MN blood group gene frequencies in Japan, and presented data which suggest that this differentiation has occurred largely by genetic random drift. If drift is important, then the differentiation of gene frequencies among isolated populations of small size is expected to be higher than that of the general population, because the rate of migration, which prevents the differentiation, is lower in the former. Thus, a genetic analysis was undertaken of the differentiation of gene frequencies among isolated populations in Japan. In this paper we will be concerned only with the ABO blood groups, since data on the MN blood groups are limited.

### 2. ANALYSIS OF THE DATA

There are twelve isolated populations, for which the data on the ABO blood group frequencies have been published. These data are given in table 1, together with the gene frequency estimates. Of the twelve isolated populations, seven are inhabitants of small islands surrounding the main islands of Japan and others are located in mountainous regions (cf. fig. 1). As seen from table 1, some populations are not small but they had to be used because of the paucity of data. Actually, some populations include several subpopulations among which the blood group gene frequencies differ significantly (Nei and Imaizumi, unpublished). Also, the degree of isolation is not the same for all populations. Some island populations are highly isolated, whereas in some inland populations, it appears that considerable migration has taken place. It should be noted, however, that in almost all the populations the recently immigrated individuals have been excluded from the blood group typing.

The frequencies of the gene  $I^A$ ,  $I^B$ , and  $I^O$  controlling the ABO blood groups were estimated by the same method as used in the previous paper (Nei and Imaizumi, 1966), *i.e.* Bernstein's method. The results obtained are given in table 1, in which the frequencies of  $I^A$ ,  $I^B$ , and  $I^O$  are denoted by  $p$ ,  $q$ , and  $r$  respectively. The means of  $p$ ,  $q$ , and  $r$  weighted for the number of observations are 0.2688,

\* This investigation was supported in part by a grant from the Scientific Research Fund of the Ministry of Education, Japan.

TABLE I  
*ABO blood group frequencies and estimates of gene frequencies in isolated populations*

Population	N <sub>s</sub>	N	Per cent. frequency				P	q	r	χ <sup>2</sup>	Source
			O	A	B	AB					
Sado	1893	84,397	30.6	36.6	24.1	8.7	0.2611	0.1807	0.5582	2.0	Kobayashi and Katoh (1957)
Miomote	780	780*	27.3	28.6	33.7	10.4	0.2105	0.2531	0.5274	0.2	Ieta <i>et al.</i> (1937)
M-O	2460	4,000*	25.4	39.1	24.5	11.0	0.2948	0.1976	0.5076	1.3	Kanebako <i>et al.</i> (1958)
Okinoshima	646	773*	25.9	33.9	30.6	9.6	0.2593	0.2289	0.5208	3.5	Irako (1940)
Kusu	748	18,782	39.0	36.5	16.1	8.4	0.2561	0.1300	0.6139	5.3	Kobayashi and Okada (1956)
Okii	1993	29,641	28.2	37.8	23.0	11.0	0.2841	0.1876	0.5283	0.4	Kobayashi <i>et al.</i> (1957)
Kamitono	608	7,760	33.7	39.6	19.9	6.8	0.2689	0.1441	0.5870	1.3	Terada <i>et al.</i> (1956)
Iwajima	570	3,430*	43.2	30.4	19.6	6.8	0.2068	0.1421	0.6511	1.3	Kobayashi and Watari (1953)
Iki	674	50,593	24.8	37.2	25.8	12.2	0.2889	0.2126	0.4985	0.0	Kobayashi (1951)
Tsushima	3068	69,551*	32.4	37.1	21.2	9.3	0.2677	0.1660	0.5663	1.0	Kobayashi (1952)
Kuroshima	773	2,262*	26.6	39.2	21.7	12.5	0.3036	0.1882	0.5082	1.5	Schull <i>et al.</i> (1962)
Koshikijima	10,767	17,136	31.3	37.0	22.6	9.1	0.2660	0.1798	0.5602	0.2	Muranaga (1957)
Total	24,880		30.5	36.9	23.0	9.6	0.2683	0.1795	0.5522	0.1	

N<sub>s</sub>: number of observations. N: population size (those with asterisk refer to the numbers given in the respective papers and others to the census numbers in 1960).

0.1781, and 0.5531 respectively, which are almost identical with those for the whole of Japan, *i.e.* 0.2773, 0.1706, and 0.5521 (Nei and Imaizumi, 1966). As seen from table 1, the  $\chi^2$ -value for the departure from the Hardy-Weinberg's equilibrium is significant at the 5 per cent. level in one of the twelve populations, thus agreeing fairly well with the theoretical expectation. The  $\chi^2$  for the heterogeneity of gene frequencies among the isolated populations is 228.80 with 22 degrees of freedom and is significant at the 0.1 per cent. level.

- |              |                |
|--------------|----------------|
| 1 Sado       | 7 Kamitono     |
| 2 Miomote    | 8 Iwaijima     |
| 3 M-O        | 9 Iki          |
| 4 Okinoshima | 10 Tsushima    |
| 5 Kusu       | 11 Kuroshima   |
| 6 Oki        | 12 Koshikijima |

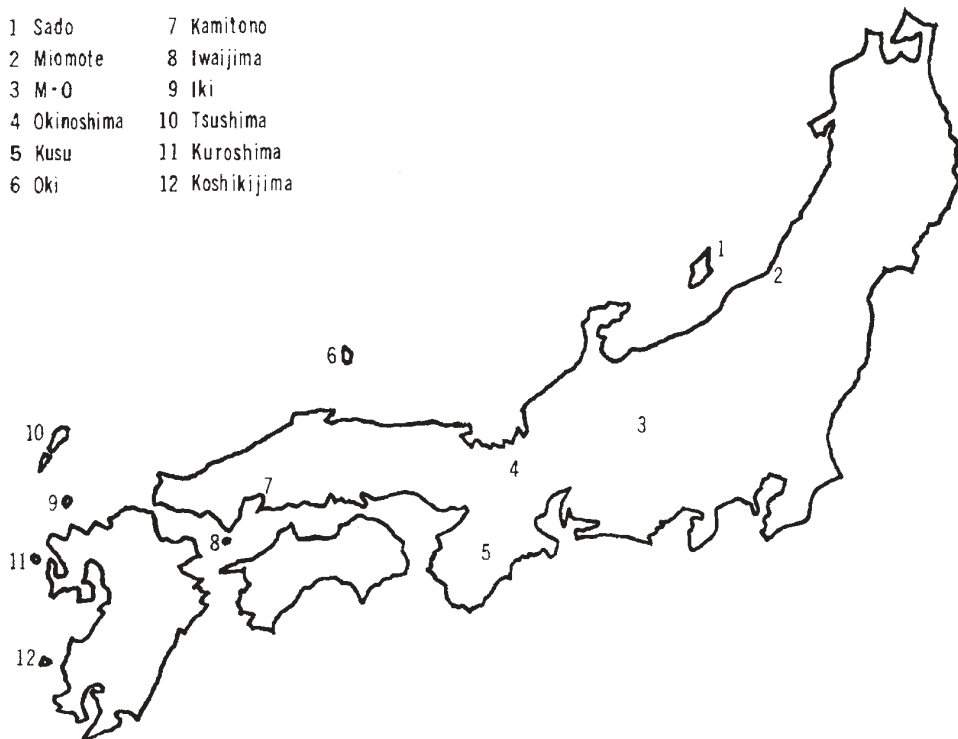


FIG. 1.—Geographical distribution of the isolated populations studied.

The true or genetic variances and covariances of  $p$ ,  $q$ , and  $r$  are obtained by subtracting the sampling variances and covariances from the observed variances and covariances respectively, as described in the previous paper. The genetic variances and covariances thus obtained are presented in table 2, together with those for the whole of Japan obtained in the previous paper. It is seen from this table that the variances of gene frequencies are all larger in the isolated populations than in the general population of Japan, as expected. If the differentiation of gene frequencies among subpopulations has occurred at random, the covariances between gene frequencies should all be negative (Nei, 1965). In the present case, the covariances between  $p$  and  $r$ , and  $q$  and  $r$ , are negative, but the covariance between  $p$  and  $q$  is positive. This may be due to sampling accident because of the

small number of isolated populations employed, or it may be due to natural selection disturbing the random differentiation. Inspection of fig. 2, in which the correlations between gene frequencies are graphically shown, reveals that the gene frequencies of Iwajjima deviate greatly from the others, contributing to the positive correlation between  $p$  and  $q$ . Thus, excluding the data on Iwajjima, the genetic variances and covariances were recomputed, the results of which are given also in table 2. The covariance between  $p$  and  $q$  is now negative.

TABLE 2  
*Genetic variances, covariances, and the degree of differentiation of the ABO blood group gene frequencies*

Gene frequency concerned	Isolated populations including Iwajjima		Isolated populations excluding Iwajjima		Whole of Japan	
	$\sigma_j^2$ or $\sigma_{jk}$	$f_{ST}$	$\sigma_j^2$ or $\sigma_{jk}$	$f_{ST}$	$\sigma_j^2$ or $\sigma_{jk}$	$f_{ST}$
p	0.000262	0.00133	0.000182	0.00092	0.000146	0.00073
q	0.000436	0.00296	0.000416	0.00281	0.000092	0.00065
r	0.000771	0.00312	0.000561	0.00226	0.000183	0.00074
p.q	0.000037	-0.00076	-0.000018	0.00038	-0.000028	0.00058
p.r	-0.000299	0.00201	-0.000163	0.00110	-0.000118	0.00077
q.r	-0.000476	0.00480	-0.000397	0.00401	-0.000065	0.00069
Mean		0.00224		0.00191		0.00069

The degree of differentiation of gene frequencies may be measured by the inbreeding coefficient of subpopulations relative to the total population. This inbreeding coefficient ( $f_{ST}$ ) is obtained by

$$f_{ST} = \frac{\sigma_j^2}{\bar{p}_j(1-\bar{p}_j)} \quad \text{or} \quad f_{ST} = \frac{-\sigma_{jk}}{\bar{p}_j\bar{p}_k}$$

provided that the differentiation has occurred at random, where  $\bar{p}_j$  and  $\sigma_j^2$  denote the mean and the genetic variance of the  $j$ th allele frequency, respectively, and  $\sigma_{jk}$  the genetic covariance of the  $j$ th and  $k$ th allele frequencies (Wright, 1943; Nei, 1965). The estimates of this inbreeding coefficient, obtained by including and excluding the data on Iwajjima, are given in table 2. There is a considerable heterogeneity in the estimate of  $f_{ST}$ , the values obtained from the variance of  $q$  and  $r$  being two or three times larger than that obtained from the variance of  $p$ . Further, the estimate of  $f_{ST}$  obtained from the covariance between  $p$  and  $q$  is negative when the Iwajjima population is included, although the estimates from covariances are generally subject to a larger sampling error than those obtained from variances. The heterogeneity of  $f_{ST}$  suggests that the isolated populations employed are not random samples of the whole, or that the differentiation of gene frequencies has not occurred completely at random (cf. Nei, 1965). Exclusion of the Iwajjima population makes the estimate of  $f_{ST}$  more

homogeneous but not so much as in the case of the whole of Japan (table 2). The mean estimate of  $f_{ST}$  obtained by excluding the data on Iwaijima is slightly smaller than that obtained by including them.

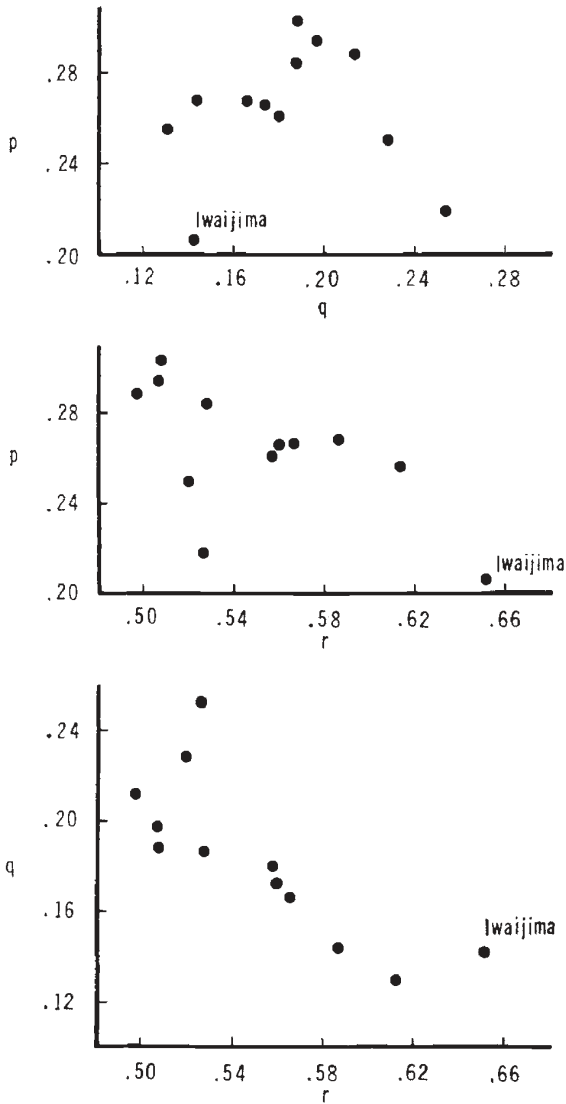


FIG. 2.—Correlations between different gene frequencies.

This is because the gene frequencies of Iwaijima deviate considerably from the others. But even the former estimate is larger by a factor of 3 than that for the whole of Japan.

The randomness of differentiation of gene frequencies can be tested also by comparing the observed and expected values of correlation

between different gene frequencies (Nei, 1965). The expected correlation between the  $j$ th and  $k$ th allele frequencies is given by

$$r(j, k) = -\sqrt{\frac{\bar{p}_j \bar{p}_k}{(1 - \bar{p}_j)(1 - \bar{p}_k)}}$$

The observed correlation is obtained from the genetic variances and covariances. The results obtained are presented in table 3. The difference between the expected and observed correlations is not significant in any pair of gene frequencies, regardless of whether the Iwajima population is included or excluded.

TABLE 3  
*Correlations between the ABO blood group gene frequencies*

Gene frequency concerned	Isolated populations including Iwajima		Isolated populations excluding Iwajima	
	Expected	Observed	Expected	Observed
p.q	-0.2832	0.1087	-0.2851	-0.0670
p.r	-0.6725	-0.6644	-0.6718	-0.6421
q.r	-0.5193	-0.8213	-0.5184	-0.7941

### 3. DISCUSSION

Nei and Imaizumi (1966) showed that the differentiation of ABO blood group gene frequencies in Japan has occurred largely at random, although a slight cline was observed from north to south, and this random differentiation was attributed to a genetic drift owing to the limited effective size of neighbourhoods in Wright's (1946) sense. The fact that the estimate of  $f_{ST}$  is larger in isolated populations than in the general population supports this view.

The estimate of  $f_{ST}$ , however, is less homogeneous in the isolated populations than in the general population. The deviations of the observed from the expected correlations between gene frequencies are also larger in the former than in the latter (cf. Nei and Imaizumi, 1966). If the sample sizes were large, these results would indicate that the differentiation of gene frequencies in isolated populations has not occurred at random. However, in view of the relatively small sample sizes employed, these may well be due to sampling accident, perhaps with the exception of Iwajima. In the case of the whole of Japan the total number of observations was 599,342, divided into 45 subpopulations, while in the present case it is 24,880, divided into 12 isolated populations.

As seen from table 1, the phenotype frequencies summed for all the isolated populations are in good agreement with the Hardy-Weinberg expectations, the  $\chi^2$ -value for the discrepancy being 0.1. If the data

on Iwajima are excluded, the  $\chi^2$  becomes 0.2. This and the fact that there is a significant differentiation of gene frequencies among the isolated populations suggest that the heterozygotes for the ABO blood groups have a selective advantage over the homozygotes, as discussed in the previous report. Assuming that the genotype frequencies are in complete agreement with the Hardy-Weinberg expectations, it may be conjectured that the heterozygotes have a selective advantage at least of the order of  $f_{ST}$  or 0.2 per cent. on the average.

It may be noted that the mating structure of our isolated populations closely approximates to Wright's island model. In these populations certain proportions of genes are considered to have been exchanged every generation with the large general population in the main islands. The average proportion of gene exchange or the rate of migration per generation ( $m$ ) can be estimated by Wright's (1943) formula

$$m = 1 - \sqrt{2N_e f_{ST} / [(2N_e - 1)f_{ST} + 1]}$$

where  $N_e$  is the effective population size. It is not easy to know the exact effective size of a population in which generations are completely overlapped. Recently, Kimura and Crow (1963) developed a formula but its applicability appears to be limited. Therefore, we use the following approximate formula

$$N_e = \tau N_m$$

where  $\tau$  is the generation time or the mean reproductive age and  $N_m$  the number of individuals who reach the mean reproductive age per year (see the Appendix to this paper on p. 344). In a stationary population  $N_m$  may be written as  $bpN$ , *i.e.* the product of the crude birth rate per year ( $b$ ), the probability that a newborn individual reaches the mean reproductive age ( $p$ ), and the total population size ( $N$ ). The values of  $\tau$ ,  $b$  and  $p$  around 1900 in Japan are approximately 30, 0.32, and 0.65 respectively. Thus,  $N_e$  is given by 0.65 $N$  approximately. The sizes of isolated populations in Japan appear to have been almost constant (*cf.* Schull, *et al.*, 1962). Hence, using the population sizes given in table 1, the effective size of each population can be obtained. The average effective size is then given by the harmonic mean of these individual effective sizes. The value thus obtained is 1993, excluding the Iwajima population. Using this number and the previous estimate of  $f_{ST}$  ( $= 0.00191$ ), the rate of migration becomes 0.06. In this connection it should be noted that the heterozygote advantage, which was advocated here, causes the rate of migration to be overestimated. It seems, therefore, that our isolated populations have exchanged about six per cent. or less of individuals every generation with the general population, excepting those who emigrated because of the population overflow.

#### 4. SUMMARY

1. The ABO blood group gene frequencies among isolated populations with scattered distributions in Japan are differentiated nearly

at random and the degree of the differentiation, as measured by the inbreeding coefficient, is about three times higher than that of the general population.

2. This may be attributed to a lower rate of migration in the isolated populations. This rate was estimated to be about six per cent. or less per generation. These findings support our view that the local differentiation of ABO blood group gene frequencies in Japan has occurred largely by genetic random drift.

3. An analysis of genotype frequency distributions has suggested that the heterozygotes have a higher viability than the homozygotes by at least 0.2 per cent. on the average.

## 5. REFERENCES

- IETA, S., YAMANOUCHI, S., OGATA, S., AND OHNISHI, S. 1937. The blood group frequencies in Miomote-mura. *Hokuetsu Igakkai Zasshi*, 52, 1285-1288.
- IRAKO, M. 1940. The frequencies and inheritance of blood groups in an isolated island of Biwa lake. *Hanzaigaku Zasshi*, 14, 153-156.
- KANEBAKO, F., AKABANE, J., AIZAWA, Y., SUZUKI, H., AND OHNO, K. 1958. Report on the hematypological investigation upon M-O villagers (Nagano Prefecture). *Hanzaigaku Zasshi*, 24, 24-27.
- KIMURA, M., AND CROW, J. F. 1963. The measurement of effective population number. *Evolution*, 17, 279-288.
- KOBAYASHI, H. 1951. Racial biological studies of the inhabitants in Iwai Island. I. The ABO blood groups. *Igaku to Seibutsugaku*, 18, 321-323.
- KOBAYASHI, H. 1952. On the ABO blood groups of the inhabitants of Iki Island (Nagasaki Prefecture). *Igaku to Seibutsugaku*, 22, 170-171.
- KOBAYASHI, H. 1952. Racial biological studies of the inhabitants in Tsushima Island. *Hiroshima Ikadaigaku Ronbunshu*, 4, 95-102.
- KOBAYASHI, H., AND WATARI, K. 1953. On the ABO blood groups of the inhabitants in Kamitono-mura, Hiroshima Prefecture. *Hiroshima Igaku*, 6, 185-186.
- KOBAYASHI, H., AND OKADA, K. 1956. A report of the investigation of blood groups of the inhabitants in Kusu-mura, Nara Prefecture. *Igaku to Seibutsugaku*, 38, 120-122.
- KOBAYASHI, H., AND KATOH, S. 1957. On the distribution of blood groups of the inhabitants of Sado Island, Niigata Prefecture. *Zinruigaku Zihō*, 18, 67-72.
- KOBAYASHI, H., TERADA, K., AND KAJITANI, K. 1957. On the distribution of the blood groups of the inhabitants of Oki-togo Island. *Igaku to Seibutsugaku*, 42, 83-85.
- MURANAGA, K. 1957. A study on the islanders of Koshikijima, from the viewpoint of their fingerprint and blood type. II. Frequencies of ABO blood groups. *Okayama Igakaki Zasshi*, 69, 3019-3030.
- NEI, M. 1965. Variation and covariation of gene frequencies in subdivided populations. *Evolution*, 19, 256-258.
- NEI, M., AND IMAIZUMI, Y. 1966. Genetic structure of human populations. I. Local differentiation of blood group gene frequencies in Japan. *Heredity*, 21, 9-36.
- SCHULL, W. J., YANASE, T., AND NEMOTO, H. 1962. Kuroshima: the impact of religion on an island's genetic heritage. *Human Biology*, 34, 271-298.
- TERADA, K., KAJIYA, K., FURUTOKU, T., TANAKA, Y., AND FUJITANI, Y. 1956. The blood groups of the inhabitants of Oki-Island (Tozen) *Nihon Hoigaku Zasshi*, 10, 325.
- WRIGHT, S. 1943. Isolation by distance. *Genetics*, 28, 114-138.
- WRIGHT, S. 1946. Isolation by distance under diverse systems of mating. *Genetics*, 31, 39-59.