
Is Migration Kin Structured?

ELISABETH J. MANDERSCHIED,¹ JAMES A. BRANNAN,¹ AND ALAN R. ROGERS¹

Abstract We estimate the strength of kin-structured migration in six human populations (five from New Guinea and one from Finland) and in one population of nonhuman primates. We also test the hypothesis that migration is not kin structured by generating a sampling distribution of the estimator under the null hypothesis of independent random migration. We are unable to detect a statistically significant level of kin-structured migration in any population. However, five of our six human populations were from Papua New Guinea, and we cannot dismiss the possibility that migration is kin structured in other parts of the world.

The original architects of population genetics theory were masters of omission. They built models of simplified populations in which many of the complexities of real populations did not appear. The enduring interest in these simple models reflects the early population geneticists' excellent judgment about what to omit. Later work has often shown that models with greater realism and complexity give substantially the same results as their simpler predecessors.

This article concerns one form of complexity on which the jury is still out: kin-structured migration. Until the late 1970s all models of gene flow assumed that the individuals who move from place A to place B are an independent random sample of those born in place A. This was a simplification: In the real world individuals often migrate together with relatives. Consequently, the individuals migrating from place A to place B are likely to be more similar genetically than a random sample drawn from place A. Fix (1978, 1985) called this kin-structured migration and showed that it tended to increase genetic differences between groups. Genetic differentiation increases with the level of kin structure. By simulating different degrees of kin structure in plants, Levin and Fix (1989) demonstrated that the extreme levels of kin structure expected in plant populations with self-fertilization could result in high levels of variability.

Kin-structured migration may also affect phenotypic differentiation. Williams-Blangero (1989a,b) compared phenotypic differentiation

¹Department of Anthropology, University of Utah, Salt Lake City, UT 84112.

between villages of the Jirels in Nepal with that expected under random migration and attributed the difference to clan-structured migration.

Theoretical work (Rogers 1987, 1988) has shown that the magnitude of kin structure can be measured by the parameter α_M , which affects the ratio G of between- to within-group genetic variance. If G^* is the ratio that would be expected in the absence of kin structure, then (Rogers 1987, 1988)

$$G \approx G^*(1 + 2m_e\alpha_M), \quad (1)$$

where m_e is the effective migration rate (Rogers and Harpending 1986). If $m_e\alpha_M$ is large, then kin structure is an important effect. Otherwise, its effect may justifiably be ignored. Note that α_M is not a correlation coefficient and need not lie within the interval $[-1, +1]$.

There are two ways to estimate the magnitude of α_M . First, it can be estimated from behavioral data. Rogers (1987, 1988) showed that, if all migration involves groups of relatives of size γ and if the genetic correlation between the individuals within these groups is κ , then very approximately $\alpha_M \approx (\gamma - 1)\kappa$. Thus, if migration always involves groups of three siblings, then $\alpha_M \approx (3 - 1) \times 0.5 = 1$. Behavioral data from several species (Rogers 1987, pp. 420–421) suggest that α_M can be large.

It is also possible to infer kin structure from genetic data (Rogers and Jorde 1987). Rogers and Eriksson (1988) introduced a method for estimating α_M from genetic data. They applied their method to only a single population, the human population of the Åland Islands, Finland, and obtained a result that did not differ significantly from 0. Here, we use genetic data to estimate α_M in six additional human populations and in a single population of nonhuman primates.

Methods

The set of individuals who are born in group i and later reside in group j as adults is called the ij th migrant set, and its size is denoted N_{ij} . Kin structure tends to inflate the variances of migrant set allele frequencies. For a given allele, let q_{ijk} denote the allele frequency (0, 0.5, or 1) of the k th individual in the ij th migrant set. We calculate the following sums of squares:

$$I_i = \sum_{jk} q_{ijk}^2 \quad (2)$$

is the sum of squares between individuals within the i th birthplace,

$$M_i = \sum_j N_{ij} q_{ij}^2 \quad (3)$$

is the sum of squares between migrant sets within the i th birthplace, and

$$B_i = N_{i+} q_{i\cdot}^2 \quad (4)$$

is the contribution of the i th birthplace to the sum of squares between birthplaces. Here,

$$N_{i+} = \sum_j N_{ij} \quad (5)$$

is the size of the i th birthplace,

$$q_{ij\cdot} = N_{ij}^{-1} \sum_k q_{ijk} \quad (6)$$

is the allele frequency within the ij th migrant set, and

$$q_{i\cdot\cdot} = N_{i+}^{-1} \sum_j N_{ij} q_{ij\cdot} \quad (7)$$

is the allele frequency within the i th birthplace.

Our analysis is based on the estimator described by Rogers and Eriksson (1988, p. 453). There is a typographical error in the published formula, which should have been

$$\hat{\alpha}_M = \frac{\sum_i (M_i - B_i)(N_{i+} - 1)}{\sum_i (I_i - B_i)(k_i - 1)} - 1, \quad (8)$$

where k_i is the number of nonempty migrant sets originating in group i and i runs over the birthplaces within the population. Rogers and Eriksson point out that multiple alleles and loci can be accommodated by allowing i to run over all alleles and loci as well as over all birthplaces.

We can also estimate α_M using data from several populations simultaneously, under the assumption that the true value of α_M is the same in each population. For that case the index of summation in Eq. (8) is allowed to run over populations as well as over the alleles, loci, and birthplaces within them.

Statistical inference is based on an empirical approximation to the sampling distribution of $\hat{\alpha}_M$ under the hypothesis that $\alpha_M = 0$. This sampling distribution is obtained through Monte Carlo simulation, as described by Rogers and Eriksson (1988).

The software used in this analysis is available electronically through ftp from ANTHRO.UTAH.EDU.

Data

Data from several human populations were generously provided by J. Littlewood, J. Wood, and S. Serjeantson. Littlewood's data derive

from the University of Washington New Guinea Micro-Evolution Project, which took place in the early 1960s. The language groups studied are located in the Eastern Highlands of New Guinea. Wood's data result from his 1978 study of the Gainj located in the Takwi valley, at the northern edge of the Central Highlands of New Guinea. Serjeantson provided data from three populations in and around Papua New Guinea: the Tep Tep, located in the Finisterre Range; the Lufa of the Eastern Highlands; and the inhabitants of Karkar Island, located off the northeast coast of Papua New Guinea. Karkar Islanders and the Lufa were studied in the context of the Human Adaptability Section of the International Biological Programme.

The New Guinea groups practice horticulture, their main crops being sweet potatoes in the Highlands and taro on Karkar Island. Cash-crop growing (e.g., coffee) and migrant labor (plantations) have developed since the 1960s and provide some small cash income (Littlewood 1972; Watson 1983; Wood et al. 1982; Norgan et al. 1974; Walsh 1974; Long et al. 1986). Karkar Island, however, is much more developed economically because of its long contact history, and only 56.6% of adult men describe themselves as subsistence farmers (Hornabrook et al. 1974).

Precontact history in the Eastern and Central Highlands was characterized by endemic warfare, and segments of groups or families often sought refuge in another group (Feil 1987; Watson 1983; Robbins 1982; Long et al. 1986).

Contemporary migration consists of women moving to their husband's group as a result of preferred patrilocality (Feil 1987) and of men emigrating to work on plantations. The men usually return to their village (Healey 1985). Thus the ethnographic evidence suggests kin-structured migration during precontact times but provides no clear prediction about the kin structure of contemporary migration.

The rhesus monkey colony of Cayo Santiago, established in 1938, has been described in detail elsewhere (Sade et al. 1985). As with most Old World primates, the female rhesus monkeys remain in their natal group throughout life, whereas males emigrate at adulthood (Sade et al. 1985; Greenwood 1980; Koford 1966; Colvin 1983). Because only males migrate, we ignored females in our analysis. Observers report that male migration occurs at about 3 years of age and does not appear to be kin structured (Colvin 1983). The lack of behavioral evidence for kin structure is interesting in view of the ample evidence that among many Old World monkeys males often migrate in the company of peers and/or brothers (Cheney and Seyfarth 1977, 1983; Meikle and Vessey 1981; Sugiyama 1976; Kawanaka 1973; Froehlich et al. 1981). Furthermore, in one rhesus population monkeys often migrate into groups where older maternally related brothers already reside (Meikle and Vessey 1981). This allows the migrants to integrate into the new group more easily

Table 1. Description of Data

<i>Population</i>	<i>Size</i>	<i>Number of Groups</i>	<i>Number of Systems</i>	<i>% Migrating</i>	<i>K</i>
Åland	2408	16	6	29	17
Gainj	361	22	10	31	23
Lufa	293	10	7	17	11
E. Highlands	816	67	4	5	68
Tep Tep	289	6	8	18	13
Rhesus a	158	10	3	79	11
Rhesus b	156	9	3	77	10
Karkar	454	9	6	16	14

K is the number of component tests used in the sequential Bonferroni test for multiple comparisons (see text).

because genetically related rhesus monkey migrants often support each other in aggressive encounters and because monkeys with relatives are higher in dominance than solitary migrants.

In the analysis of each population we attempted to exclude local groups (villages, parishes, or bands) that were not censused by the investigator. Typically, the data included some immigrants from outside the study area. To exclude these individuals, we included only those local groups that appeared both as birthplaces and as adult residences in the data. We also excluded individuals who were younger than age 18 when the data were collected and individuals with missing values for all relevant genetic systems. These exclusions reduced the size of some data sets significantly. To facilitate comparison, we treated the Åland data anew. The sample size and number of groups in this article differ from those published previously (Rogers and Eriksson 1988).

The rhesus monkey data comprise all male monkeys born after 1960. Male monkeys migrate repeatedly over the course of their life. The population was analyzed twice: first, analyzing migration from birthplace to last residence recorded (Rhesus a) and, second, analyzing migration from birthplace to first adult residence (Rhesus b).

The sample size, number of groups, number of genetic systems, and migration rate for each data set are given in Table 1. The sample size ranges from as little as 156 to as many as 2408 individuals. The number of groups ranges from 6 to 67, and the number of genetic systems ranges from 3 to 10. Only codominant systems were used. The percentage of individuals migrating varies from 5% to 79%. The very low rate of migration (5%) in Littlewood's data set is probably due to the facts that (1) the data are predominantly male, (2) the local residence pattern is patrilocal, and (3) villages that were closely related through

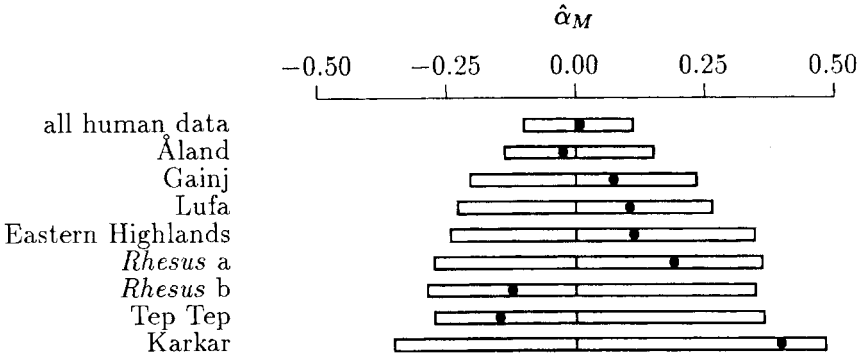


Figure 1. Critical regions and observed values. To reject the hypothesis that $\alpha_M = 0$ at the 0.05 significance level, the estimate, $\hat{\alpha}_M$, must fall outside the critical regions shown by the bars. The filled circles indicate the observed values of $\hat{\alpha}_M$. "Rhesus a" refers to our analysis of migration to last residence group recorded, whereas "Rhesus b" traces migration from natal group to first adult group.

marriage or had split recently were pooled for the purpose of his study (Littlewood 1972).

Results

Our estimates of α_M are shown in Figure 1. The rectangles in the figure indicate the rejection region of a test of the hypothesis that $\alpha_M = 0$. This hypothesis can be rejected at the 0.05 significance level for a particular population only if the estimate for that population falls outside its rectangle. However, Figure 1 shows that none of our estimates is significantly different from 0.

We searched for significant results in several ways. First, we considered the hypothesis that all human populations share a single value of α_M . We estimated α_M under this hypothesis, as described in the Methods section. Combining data sets in this fashion increased the power of our test, as indicated by the narrowness of the rectangle at the top of Figure 1. However, our estimate, $\hat{\alpha}_M = 0.010$, is very close to 0 and remains insignificantly different from 0.

We also examined subdivisions and subsets of the study groups. This yielded several apparently significant results. However, when many hypotheses are tested, a certain number of spuriously significant results are to be expected because of chance alone. We corrected for this problem by using the sequential Bonferroni test for multiple comparisons (Rice 1989), which eliminated all our seemingly significant results. The sequential Bonferroni test compares the p values, starting with the small-

est, to α/K , where α is the desired significance level (in this case 0.05) and K is the number of component tests. We defined K for each population by adding the number of tests performed on the populations as a whole and, when applicable, the tests performed on the subsets. The tests per population and per subset are obtained by adding the number of birthplaces tested and the aggregate test over those birthplaces. For example, there would be four component tests for a population or subset containing three different birthplaces.

Discussion

Our results provide no support for the notion that kin structure is pervasive in human migration. None of the populations, taken as a whole, shows any evidence of kin structure. Why should this be?

First, it is possible that kin structure is indeed pervasive but weak in magnitude. The rectangles in Figure 1 indicate that for most populations α_M is unlikely to be judged significantly different from 0 unless its true value is above 0.4 or so. This is an appreciable amount of kin structure. For example, if every instance of migration involved a pair of brothers, α_M would equal roughly 0.5. Thus our statistical method is not powerful enough to detect low levels of kin structure with data sets of the size that are presently available.

Moreover, we used a two-tailed statistical test, and it is noteworthy that Williams-Blangero's (1989b) significant evidence for nonrandom migration in the Jirels of Nepal was obtained with a one-tailed test.

Second, our estimates may be contaminated by at least two kinds of bias. The first kind of bias arises when populations are sampled so as to minimize the relatedness of individuals within the sample. For example, an investigator might decide to include only one member of each sibship in the sample. Such a sample would be useful for some purposes, because it would reduce the sampling variance of estimates of allele frequencies. However, if two members of a sibship migrate together, this event would be invisible in such a sample. Thus we suspect that this sampling strategy would bias estimates of α_M downward.

A second kind of bias arises when the relatives in a sample are too many rather than too few. For example, suppose that in some village we include only a single sibship in the sample. If two brothers from that sibship migrate together, they would appear to our program as merely a random pair of individuals. Thus $\hat{\alpha}_M$ is probably also biased downward when the sampling strategy exaggerates the relatedness of individuals within local groups. Although it is doubtful that any investigator would sample just a single sibship, a less extreme version of this bias might easily arise. A truly random sample of even a small population would

be prohibitively difficult to draw, so most investigators adopt a fairly informal approach to sampling. These informal samples probably often include larger clusters of relatives than would appear in a random sample.

Finally, five of our six human populations are from Papua New Guinea. The focus on one region precludes any generalizing conclusion. Migration may well be kin structured in other parts of the world.

Acknowledgments This research was supported by the US Department of Health and Human Services under grant MGN 1 R29 GM39593. Data were kindly supplied by Lynn Jorde and Aldur Eriksson (Åland Islands), Susan Serjeantson (Karkar, Tep Tep, Lufa), Robert Littlewood (Eastern Highlands), and James Wood (Gainj).

Received 10 December 1992; revision received 17 March 1993.

Literature Cited

- Cheney, D.L., and R.M. Seyfarth. 1977. Behaviour of adult and immature male baboons during intergroup encounters. *Nature* 269:404–406.
- Cheney, L.D., and R.M. Seyfarth. 1983. Nonrandom dispersal in free-ranging vervet monkeys: Social and genetic consequences. *Am. Natur.* 122:392–412.
- Colvin, J. 1983. Influence of the social situation on male migration. In *Primate Social Relationships*, R.A. Hinde, ed. Sunderland, MA: Sinauer Associates, 160–171.
- Feil, D. 1987. *The Evolution of Highland Papua New Guinea Societies*. Cambridge, England: Cambridge University Press.
- Fix, A. 1978. The role of kin-structured migration in genetic microdifferentiation. *Ann. Hum. Genet.* 41:329–339.
- Fix, A. 1985. Evolution of altruism in kin-structured and random subdivided populations. *Evolution* 39:928–939.
- Frochlich, J.W., R.W. Thorington, and J.S. Otis. 1981. The demography of howler monkeys (*Alouatta palliata*) on Barro Colorado Island. *Int. J. Primatol.* 2:197–205.
- Greenwood, P.J. 1980. Mating systems, philopatry, and dispersal in birds and mammals. *Anim. Behav.* 30:1140–1162.
- Healey, C.J. 1985. New Guinea inland trade: Transformation and resilience in the context of capitalist penetration. *Mankind* 15:127–144.
- Hornbrook, R.W., G.G. Crane, and J.M. Stanhope. 1974. Karkar and Lufa: An epidemiological and health background to the human adaptability studies of the International Biological Programme. *Phil. Trans. R. Soc. Lond.*, ser. B, 268:293–308.
- Kawanaka, K. 1973. Intertroop relations among Japanese monkeys. *Primates* 14:113–159.
- Koford, C.B. 1966. Population changes in rhesus monkeys: Cayo Santiago, 1960–1964. *Tulane Stud. Zool.* 13:1–7.
- Levin, D.A., and A.G. Fix. 1989. A model of kin migration in plants. *Theor. Appl. Genet.* 77:332–336.

- Littlewood, R.A. 1972. *Anthropological Studies in the Eastern Highlands of New Guinea*, v. 2, *Physical Anthropology of the Eastern Highlands of New Guinea*. Seattle, WA: University of Washington Press.
- Long, J.C., J.M. Naidu, H.W. Mohrenweiser et al. 1986. Genetic characterization of Gainj- and Kalam-speaking people of Papua New Guinea. *Am. J. Phys. Anthropol.* 70:75-96.
- Meikle, D.B., and S.H. Vessey. 1981. Nepotism among rhesus monkey brothers. *Nature* 294:160-161.
- Norgan, N.G., A. Ferro-Luzzi, and J.V.G.A. Durnin. 1974. The energy and nutrient intake and the energy expenditure of 204 New Guinean adults. *Phil. Trans. R. Soc. Lond.*, ser. B, 268:309-348.
- Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* 43:223-225.
- Robbins, S. 1982. *Anthropological Studies in the Eastern Highlands of New Guinea*, v. 6, *Auyana: Those Who Held onto Home*. Seattle, WA: University of Washington Press.
- Rogers, A.R. 1987. A model of kin-structured migration. *Evolution* 41:417-426.
- Rogers, A.R. 1988. Three components of genetic drift in subdivided populations. *Am. J. Phys. Anthropol.* 77:435-449.
- Rogers, A.R., and A.W. Eriksson. 1988. Statistical analysis of the migration component of genetic drift. *Am. J. Phys. Anthropol.* 77:451-457.
- Rogers, A.R., and H.C. Harpending. 1986. Migration and genetic drift in human populations. *Evolution* 40:1312-1327.
- Rogers, A.R., and L.B. Jorde. 1987. The effect of nonrandom migration on genetic differences between populations. *Ann. Hum. Genet.* 51:169-176.
- Sade, D.S., B.D. Chepko-Sade, J.M. Schneider et al. 1985. Basic demographic observations on free-ranging rhesus monkeys. New Haven, CT: Human Relations Area Files.
- Sugiyama, Y. 1976. Life histories of male Japanese monkeys. In *Advances in the Study of Behavior*, J.S. Rosenblatt, R.A. Hinde, E. Shaw et al., eds. New York: Academic Press, v. 3, 255-285.
- Walsh, R.J. 1974. Geographical, historical, and social background of the peoples studied in the IBP. *Phil. Trans. R. Soc. Lond.*, ser. B, 268:223-228.
- Watson, J.B. 1983. *Anthropological Studies in the Eastern Highlands of New Guinea*, v. 5, *Tairora Culture: Contingency and Pragmatism*. Seattle, WA: University of Washington Press.
- Williams-Blangero, S. 1989a. Clan-structured migration and phenotypic differentiation in the Jirels of Nepal. *Hum. Biol.* 61(2):143-157.
- Williams-Blangero, S. 1989b. Phenotypic consequences of nonrandom migration in the Jirels of Nepal. *Am. J. Phys. Anthropol.* 80:115-125.
- Wood, J.W., P.L. Johnson, R.L. Kirk et al. 1982. The genetic demography of the Gainj of Papua New Guinea. I. Local differentiation of blood group, red cell enzyme, and serum protein allele frequencies. *Am. J. Phys. Anthropol.* 57:15-25.