Movement, Relatedness and the Genetic Structure of the Population of Karkar Island

G. A. Harrison, R. W. Hiorns and A. J. Boyce

Phil. Trans. R. Soc. Lond. B 1974 **268**, 241-249
doi: 10.1098/rstb.1974.0027

Phil. Trans. R. Soc. Lond. B. **268**, 241–249 (1974) [241]
Printed in Great Britain

Movement, relatedness and the genetic structure of the population of Karkar Island

By G. A. Harrison,* R. W. Hiorns† and A. J. Boyce*
* *Anthropology Laboratory, Department of Human Anatomy, University of Oxford,*
† *Department of Biomathematics, University of Oxford*

Settlement on Karkar Island is confined to the perimeter, where some 59 villages are arranged more or less in a circle around a central volcanic cone. This cone is essentially impassable and communication links are also forced to follow a circular pattern. The effect of this topography on the genetic structure of the populations has been examined using a migration matrix model to predict patterns and levels of genetic heterogeneity from observations on the exchange of individuals between villages. The information incorporated in the matrix was obtained from the initial demographic survey of the island, in which place of birth and place of residence of the 16000 inhabitants was recorded. Social organization is patri-local and most of the permanent movement occurs through women moving from their village of birth to that of their husbands. The effects of the enforced circular nature of this movement are clearly apparent but other factors such as tribal affiliations and impassable areas of forest are also important.

Biological anthropologists interested in the nature, origins and causes of human genetic variation have in the past mainly focused their concerns on the broad patterns of global geographic inter-population differences. Recently, however, increasing attention has been directed at the microanalysis of relatively small-scale situations consisting of single or small clusters of populations in circumscribed areas and situations. It has become ever more widely appreciated that it is at this level that evolution is actually occurring and theoretical advances now allow a much more rigorous treatment of the factors affecting the genetic composition and structure of individual populations (for review, see Harrison & Boyce 1972; Cavalli-Sforza 1973). In these analyses, studies of population distribution and of movement are of paramount importance. Movement plays a critical role in delineating mendelian groups, and in determining the genetic variation both within and between them.

The geography of Karkar Island (figure 1) appeared to present a distribution of settlement, and to enforce a pattern of movement, which could be of some general interest in population genetics. The island is essentially oval in shape, but because the centre is formed by an inaccessible volcanic cone the habitation is more or less confined to the periphery. Here there are 59 villages, some of them actually on the coast; others somewhat inland and ranging in size from about 600 people to as few as 60. The inland villages tend to occur in clusters and to be connected by radially orientated tracks to the coastal ones, but the main communication link is a coastal road which runs more or less completely round the island. One might thus expect the major movement pattern and consequent gene flow to follow an essentially circular pattern.

However, there are other factors which are likely to influence movement, quantitatively if not qualitatively. The population is divided into two tribal groups: the Waskia in the northern half of the island, making up 30 of the villages, and the Takia in the south, occupying 29 villages. There are also variations in population density and the eastern side of the island is covered with impenetrable tropical rain-forest and is sparsely inhabited. Strips of rain-forests occur elsewhere and river drainage also imposes some radial barriers. Particularly important today are the cash-crop plantations and missions which have developed under European control. These occur at a number of different sites on the island, and not only attract people from the local villages, but also from elsewhere in New Guinea. The islanders would seem to return eventually to the villages but many of them were born on plantations or at missions.

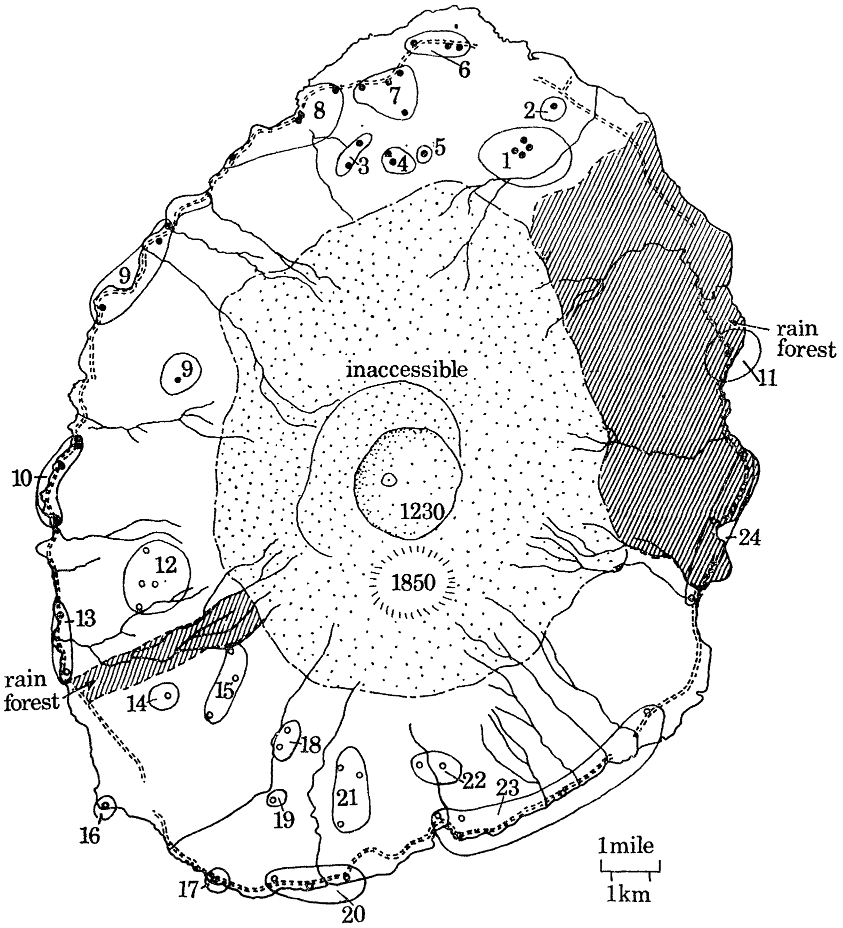

FIGURE 1. Diagrammatic map of Karkar Island, showing the main topographical features, the disposition of the villages and the way they were grouped for analysis.

In the census study which was undertaken at the beginning of the I.B.P. study, the place of birth and place of residence of every locatable villager on the island was recorded. It is these data which have most kindly been made available to us by Dr Hornabrook and his colleagues to enable the effects of the population distribution and movement on the genetic structure of the whole island to be examined. In all, data are available on 16095 individuals, and in addition to place of birth and residence there is information on marital status – whether a person is (*a*) married, (*b*) never married, (*c*) widowed, (*d*) divorced or (*e*) separated. The last three categories form a small fraction of the total and from the point of view of movement analysis have been ignored. The situation with regard to those married and never married was preliminarily examined by calculating for each village the proportion of people in each group who were at the time of the Census resident in their village of birth. These proportions were established for males and females separately and the sex difference is plotted in figure 2. It is clearly evident that in almost all the villages there is very little movement among the 'never marrieds' – who of course in the main are the young – and that such movement as does occur is equal in males and females. Among the group of married people the situation is very different – the male movement is still small, but there is in most villages a very considerable proportion of women who are living in a different place from the one in which they were born. It is concluded from this that most of the permanent movement which occurs is marital movement and that the populations are patrilocal.

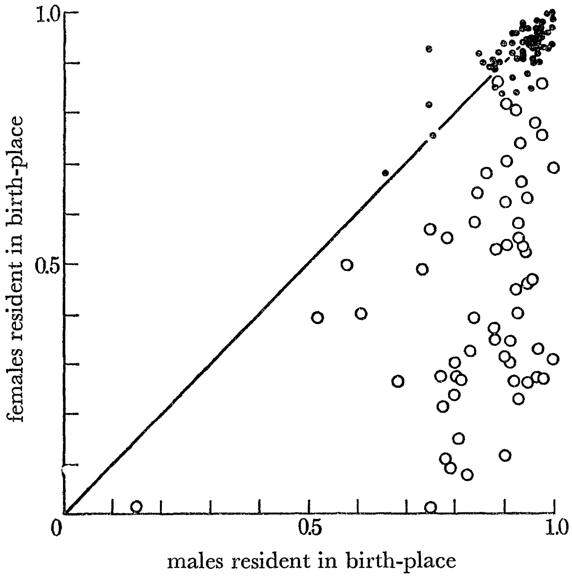

Figure 2. Proportions of married (○) and unmarried (●) males and females resident in the village in which they were born. Each point represents the situation for a village.

Subsequent analysis has been confined to the married group, who are seen as a single generation of people distributed in a set of demographically distinct clusters between which movement is occurring mainly, but not entirely, through marital exchange. It is presumed that this married group contributes to the next generation in their village of residence; and the evidence from the young would support the view that there is little post-marital movement which needs to be taken into genetic account.

The movement data have been analysed in two different ways but both employing a migration matrix approach. The first uses the model developed by Hiorns, Harrison, Boyce & Küchemann (1969), which envisages a series of population units developing increasing relationship under the influence of the migration. The model moves from a pattern where each unit has its own unique ancestry to one where 95 % of the ancestors in pairs of units are shared and demonstrates the pattern whereby this relatedness develops and the generation time required to achieve an equilibrium situation of evenly distributed ancestry throughout the system. It corresponds to a genetic system moving from maximum heterogeneity to homogeneity under the deterministic effect of gene flow.

The second model is that devized by Bodmer & Cavalli-Sforza (1968) and is formulated directly in genetic terms. It envisages a cluster of populations tending to diverge genetically from one another under the stochastic forces of drift but whose divergence is limited and patterned by the form of the migration. It predicts the within- and between-population variances in gene frequency, and the co-variances between the units. It is particularly applicable to situations, as in New Guinea, where populations are small and may be expected to be subject to marked stochastic events.

In applying these models to the Karkar Island situation there are a number of difficulties. Multiplying up a 60 × 60 matrix (59 villages and a plantation population) was not easy to handle on the computer available and it was decided to reduce it to a more manageable size by combining together the data from some of the villages. This combination was undertaken on the basis of close geographic proximity and village size and led to the formation of 24 village groups in addition to the plantation population. The grouping is given in table 1 and shown in figure 1.

TABLE 1. GROUPING OF KARKAR VILLAGES FOR MIGRATION MATRIX ANALYSIS

Waskia division				Takia division			
no.	village	total no. of individuals	effective nos	no.	village	total no. of individuals	effective nos
1	Kaul 1	418	400	12	Kuduk	395	452
	Kaul 2	262			Moban	273	
	Kaul 3	402			Kilden	253	
	Kaul 4	202			Bafor	347	
2	Mapor	299	99	13	Mangar I	206	115
3	Sikentiga	302	251		Mangar II	127	
	Narer	416		14	Marup	755	264
4	Urugen	636	236	15	Kevasop	334	342
	Urara	106			Wakon	260	
5	Sangana	380	128		Liloi	404	
6	Koropak	257	210	16	Kurum	448	143
	Deltigu	139		17	Dumad	302	111
	Tugatuga	207		18	Gamog	305	279
7	Gial	382	400		Boroman	442	
	Bangame	371		19	Did	451	172
	Langlang	128		20	Biu	324	302
	Dimer	305			Dangsai	150	
8	Apara	116	268		Patilo	321	
	Kinim	198		21	Komaria	513	231
	Dorogadam	126			Daup	135	
	Kaviak	346		22	Ulun	311	113
9A	Keng	274	248	23	Pain	157	234
	Mater	213			Warat	61	
	Marangis	252			Kavailo	189	
9B	Mom	261	96	24	Katom	145	242
10	Kurumtaur	123	218		Kubam	88	
	Kurumlang	212			Yagadum	318	
	Kulkul	129			Wadau	189	
	Buson	159			Muluk	147	
11	Gnor	124	45				

The plantation population also raises a problem since, although there is information on those individuals in the villages who were born on plantations, there are no data on the people living on the plantations at the time of the census. Further, although there are a number of plantations distributed around the island the birth-place information does not record on which particular plantation an individual was born. These difficulties have been resolved in two ways. The first (open model) was to regard the plantation population as a single 'outside world' element so large as to be essentially unaffected either in terms of relatedness or genetic composition by any in-migration to it. In the deterministic model this means that at the equilibrium the outside world is the sole provider of ancestors and genes. The alternative treatment was to regard those born on plantations as being born in the village in which they were resident and thus to ignore the plantation population altogether (closed model). This procedure will tend to exaggerate the estimate of village endogamy but may not be unrealistic if children born on plantations tend to return to the village home of their parents.

The results of applying the migration matrix models to the Karkar data are shown in figures 3, 4 and 6. In all cases the effect of the propinquity of the village groups to one another is, as is to be expected, very marked, and the consequence of the communication system and especially of the circumferential road is clearly to be seen.

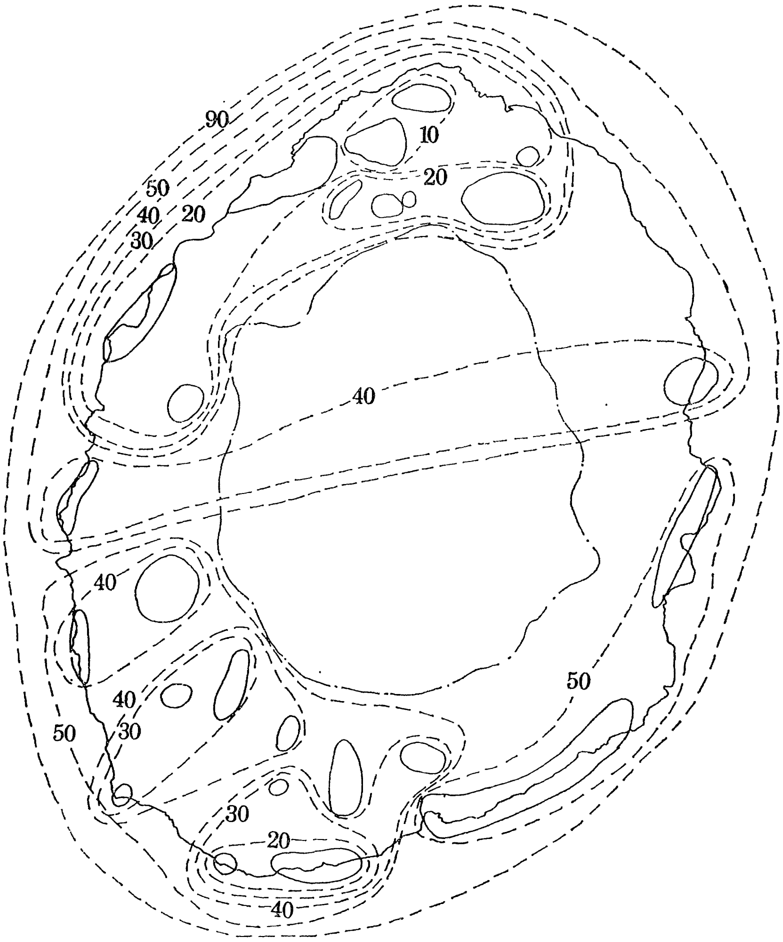

Figure 3. Patterns of developing relatedness on Karkar Island. Open model. Clusterings made on the basis of the number of generations required to establish 95% common ancestry.

Under the deterministic model, relatedness between the populations – shown in terms of 'contours' of generation times embracing groups which have come to share 95% of their ancestry in common by the specified generation in figures 3 and 4 – develops in a distinctive pattern with inland groups first tending to come together with the neighbouring coastal ones and then these joint clusters becoming related to adjacent ones in a more or less circular fashion. This follows the pattern of tracks connecting coastal and inland villages and radially arranged rivers separating adjacent groups except for the circular road.

Other factors than topography, however, are highly influential, especially the tribal division, which presents an extremely strong barrier to movement and converts the circular distribution of population into, in essence, two linear sequences. From a modelling standpoint this is unfortunate, but demonstrates the importance of taking social factors into account in human population genetics. It is the main cause for the relative isolation of population groups 10 and 11 in the northern Waskia division and of groups 12 and 24 in the Takia which is especially to be seen when no 'outside world' plantation effect is recognized (figure 4). These populations represent as it were the end of each line.

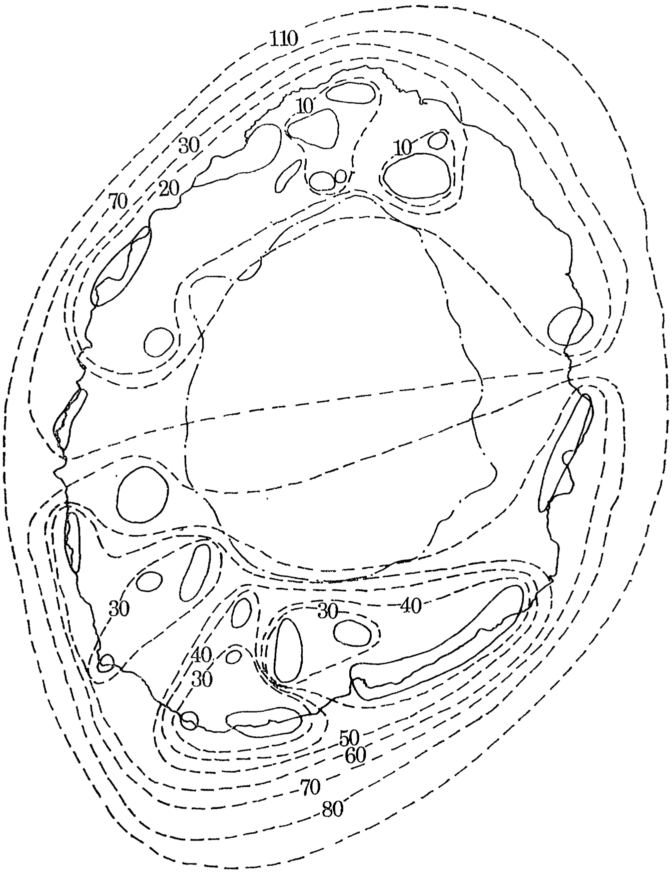

Figure 4. Pattern of developing relatedness on Karkar Island. Closed model. Clusterings made on the basis of the number of generations required to establish 95% common ancestry.

Another factor of consequence would seem to be the occurrence of an extensive area of rain-forest, and consequent low population density, on the eastern side of the island. The reason that the population groups here, 11, 23 and 24, are slow in developing relatedness with others in the open-system model is partly due to the high proportion of people in these populations who were born on plantations. (One might predict that the villages actually grow up out of plantation workers.) The nature of the open model is such that populations with a high 'outside world' contribution can only become related with others shortly before the whole system becomes indistinguishable from the 'outside world'. In this connexion it may be noted that the open model divides the Takia into a large group which becomes related to the Waskia before it joins up with the small eastern group. However, examination of the migration matrix itself shows little exchange between villages on the eastern side of the island, and in the closed system homogeneity between the two tribal groups would appear mainly to develop through flow on the west.

The open model also produces some other anomalies, especially the initial grouping of population units 10 and 11 on opposite sides of the island. Admittedly this develops late and only after the other populations of the Waskia division have come together. It probably derives from the fact that no account could be taken of the particular plantations from which individuals came so that people born in many different places on the island are grouped together as a single unit. The anomaly and smaller ones like it disappear in the closed system, notwithstanding the relative isolation of groups 10 and 11.

Exchange rates are usually higher in the Waskia division than in the Takia, and because of the form of population distribution the road would appear to play a more influential role. By the 30th generation, all the groups in the north except the two terminal populations have become fully related to each other in the open model, and only one of these is excluded in the closed model. By this time, under either model, clustering is only just beginning among the Takia.

Overall relatedness develops more rapidly in the open model than the closed one but it still requires a very long time for the system to run to equilibrium. It takes 90 generations for all 24 village clusters to come to share 95 % of their ancestry in common and 100 generations for these also to become homogeneous with the outside world. In the closed model the 24 populations have not reached equilibrium by 100 generations but the two tribal groupings arise as rapidly as in the open one, and the difference in overall time is modest.

In view of the small population sizes and the many generations it would take for overall relatedness to develop under the sole influence of deterministic migration, one might well expect considerable genetic variety to be present on the island, were it originally colonized by a single founder group. If it is assumed that no form of selection is operating (a highly controversial assumption which it would be inappropriate to discuss here) populations of such small size as the Karkar villages would be expected to drift apart at the low levels of inter-population exchanges observed. The pattern by which this develops, and at equilibrium, is predicted by the Bodmer & Cavalli-Sforza migration matrix model.

Figure 5 shows for both the open system and the closed one the expected changes in between-population variance in gene frequency (given certain assumptions intrinsic to the model). The data employed in the analysis are the same as those used in the deterministic model and because it relates to clustered villages, in which population size has been artificially increased and its variance artificially decreased, the drift effect has been underestimated. Recognizing this, it is evident that an equilibrium distribution of between-population gene frequency variance is approached by 30 generations in the open situation. In the closed one the variance is not only higher but in the nature of the model continuously increases. It is therefore the open pattern which is presented in figure 6, where the geographical clustering represents the predicted patterns of co-variance of gene frequencies between the population units after 30 generations.

It is evident that the distribution of clusters is very similar to those obtained from the deterministic models with propinquity, the circumferential road and the tribal division the main contributing factors to the pattern. It perhaps needs emphasizing again that the correspondence in results between the two types of model is enhanced by the initial village clustering which reduced the variation in the estimates of effective population sizes (table 1) as compared with that in the 59 natural village units. The influence of the migration is therefore correspondingly increased. However, it seems most unlikely that there would have been any substantial change in the broad geographical pattern had it been possible to run a 60 × 60 matrix.

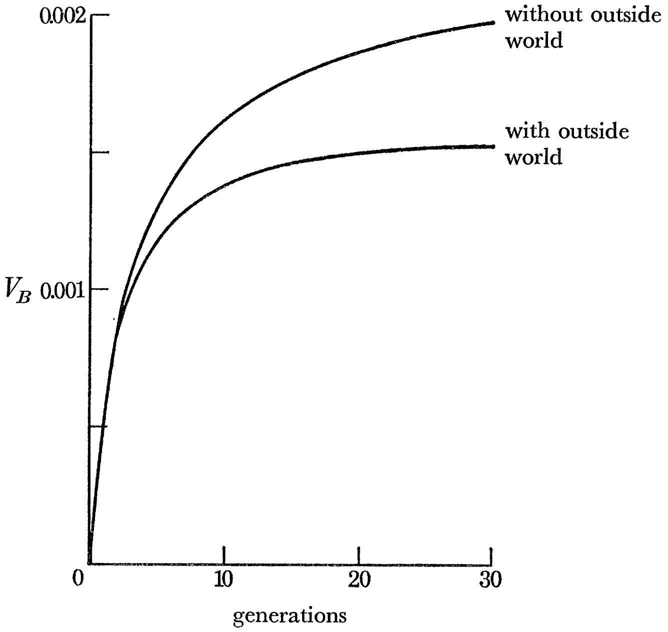

FIGURE 5. Predicted development of between-population variances in gene frequencies on Karkar Island in open and closed stochastic model.

The stochastic model also shows the special position of the populations at each end of the two tribal chains, though in this case groups 11 and 12 appear not as aberrant as in the deterministic situation, and the small village of Gnor (11) in fact, may be expected to have quite similar gene frequencies to those of the northern group of inland villages with which it is associated tribally but from which it is rather isolated geographically. On the other hand it would appear that populations 10 and 24 are likely to show abnormally low similarities with even neighbours of the same tribe. In the case of population 24 this would seem to be almost entirely due to its isolation in the relatively undeveloped eastern side of the island but in the case of 10 its position as something of a bridge between the two tribal groups would appear to be important. Indeed in the closed stochastic model population 10 at 30 generations clusters with the Takia groups before it does so with its fellow Waskias.

Unfortunately genetic information is not yet available to test these various predictions about the patterns of variety on the island. Hopefully it will be forthcoming in the future. While in our view it would be desirable to have as many of the population groups sampled as possible, so as to test the assumptions of the models, and particularly the selection issue, it is recognized that this would be a major undertaking and may not be practically possible. In this case it would appear that highest priority should be given to establishing the genetic composition of the terminal populations in each tribal group. The value of modelling is as great for establishing such practical priorities as it is for more general issues.

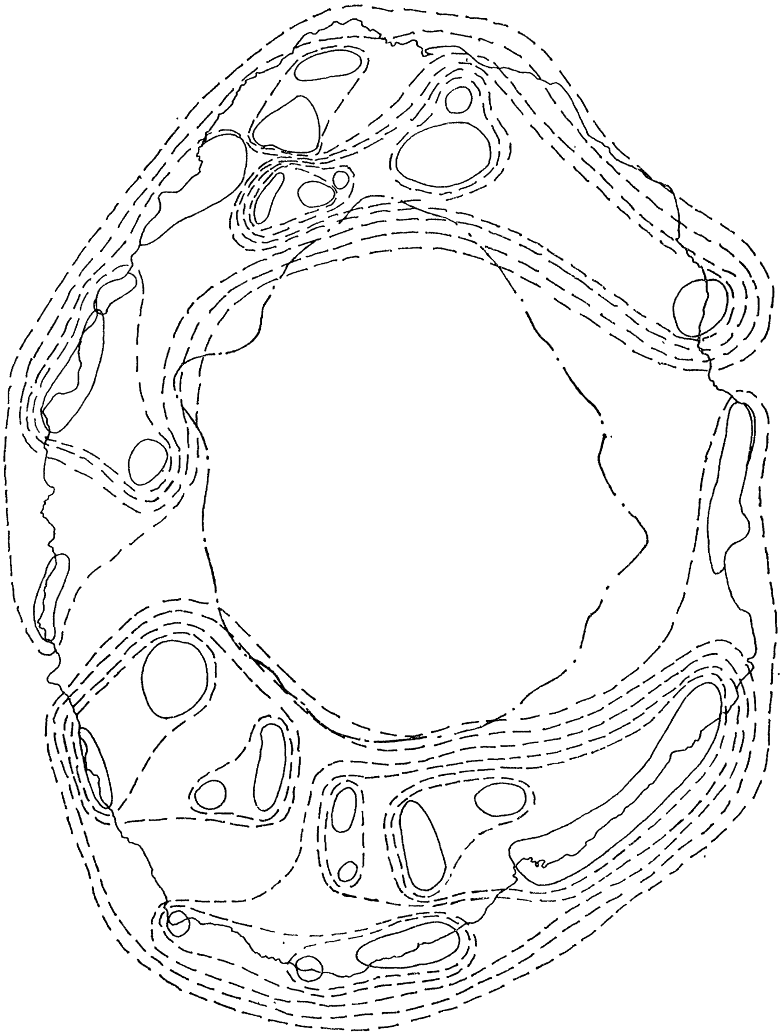

Figure 6. Pattern of predicted variation in gene frequencies on Karkar Island. Clusters made on the basis of the covariation in gene frequencies between populations at 30 generations in the open model.

We are grateful to Dr R. W. Hornabrook of the Institute of Human Biology in New Guinea and Mr W. Craig and Mrs G. Keig of the Australian National University for making available to us the demographic data on which this analysis is based. We would also like to acknowledge with thanks the help of Mrs P. J. Carrivick of the Department of Biomathematics, Oxford, who is supported on a Science Research Council Grant, for help in the computations.

References (Harrison, Hiorns & Boyce)

Bodmer, W. F. & Cavalli-Sforza, L. L. 1968 A migration matrix model for the study of random genetic drift. *Genetics* **59**, 565–592.

Cavalli-Sforza, L. L. 1973 Analytic review: Some current problems of human population genetics. *Am. J. Hum. Genet.* **25**, 82–104.

Harrison, G. A. & Boyce, A. J. 1972 Migration, exchange and the genetic structure of populations. In *The structure of human populations* (ed. G. A. Harrison and A. J. Boyce). Oxford: Clarendon Press.

Hiorns, R. W., Harrison, G. A., Boyce, A. J. & Küchemann, C. F. 1969 A mathematical analysis of the effects of movement on the relatedness between populations. *Ann. Hum. Genet.* **32**, 237–251.