PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Molecular genetic evidence for the human settlement of the Pacific: analysis of mitochondrial DNA, Y chromosome and HLA markers

E. Hagelberg, M. Kayser, M. Nagy, L. Roewer, H. Zimdahl, M. Krawczak, P. Lió and W. Schiefenhöve

Phil. Trans. R. Soc. Lond. B 1999 **354**, 141-152
doi: 10.1098/rstb.1999.0367

THE ROYAL SOCIETY

Molecular genetic evidence for the human settlement of the Pacific: analysis of mitochondrial DNA, Y chromosome and HLA markers

E. Hagelberg[1]†, M. Kayser[2]‡, M. Nagy[2], L. Roewer[2], H. Zimdahl[3], M. Krawczak[4], P. Lió[1] and W. Schiefenhövel[5]

[1]*Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK*
[2]*Institut für Rechtsmedizin, Humboldt-Universität Berlin, Hannoversche Strasse 6, D-10115 Berlin, Germany*
[3]*Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 73, D-14169 Berlin, Germany*
[4]*Institute of Medical Genetics, University of Wales, Heath Park, Cardiff CF4 4XN, UK*
[5]*Max-Planck-Institut für Verhaltensphysiologie, Von-der-Tann Strasse 3, D-82346 Andechs, Germany*

Present-day Pacific islanders are thought to be the descendants of Neolithic agriculturalists who expanded from island South-east Asia several thousand years ago. They speak languages belonging to the Austronesian language family, spoken today in an area spanning half of the circumference of the world, from Madagascar to Easter Island, and from Taiwan to New Zealand. To investigate the genetic affinities of the Austronesian-speaking peoples, we analysed mitochondrial DNA, HLA and Y-chromosome polymorphisms in individuals from eight geographical locations in Asia and the Pacific (China, Taiwan, Java, New Guinea highlands, New Guinea coast, Trobriand Islands, New Britain and Western Samoa). Our results show that the demographic expansion of the Austronesians has left a genetic footprint. However, there is no simple correlation between languages and genes in the Pacific.

Keywords: human evolution; Pacific; mitochrondrial DNA; Y chromosome; HLA

1. INTRODUCTION

(a) *The colonization of the Pacific*

The human colonization of Polynesia has attracted the attention of scholars of different disciplines, including archaeology, anthropology, genetics and linguistics (for review see Hill & Serjeantson 1989; Bellwood *et al.* 1995). The Polynesian languages belong to the Austronesian language family, spoken today in island South-east Asia (including Taiwan, Sumatra, Borneo and the Philippines), parts of the north and south-east coast of New Guinea, parts of island Melanesia, all of Polynesia and even Madagascar (figure 1). Archaeologists and linguists suggest that the widespread distribution of Austronesian languages in island South-east Asia and the Pacific was the result of a demographic expansion that followed the development of agriculture in southern China about 8000 years ago. The Polynesians were the last offshoot of the Austronesian expansion, which culminated in the colonization of the most remote regions of the Pacific, Hawaii, Easter Island and New Zealand, in the course of the last 1500 years (Bellwood 1989, 1995).

†Present address: Department of Biochemistry, University of Otago, PO Box 56, Dunedin, New Zealand (erika.h@sanger.otago.ac.nz).
‡Present address: Max-Planck-Institut für Evolutionäre Anthropologie, Inselstrasse 22, D-04103 Leipzig, Germany.

The spread of the Polynesians has been the subject of numerous genetic studies using classical and molecular markers. These studies indicate that the Polynesians are genetically homogeneous and more closely related to South-east Asians than to Australian aborigines and New Guineans (Cavalli-Sforza *et al.* 1994). Molecular analysis of globin gene polymorphisms supports the view of the South-east Asian origin of the Polynesians, but reveals relatively high levels of admixture between present-day populations of Polynesia and island Melanesia (O'Shaughnessy *et al.* 1990). In contrast, studies of the maternally inherited mitochondrial DNA (mtDNA) indicate that the expansion of proto-Polynesians into the Pacific was recent and rapid, with little genetic contribution of Melanesian genes to the Polynesian gene pool. The mtDNA data favour the so-called 'express train to Polynesia' hypothesis described by Diamond (1988) and do not reveal significant admixture between Melanesians and the Polynesian newcomers (for review, see Hagelberg 1997). It is important to remember, however, that mtDNA is just a single genetic locus, and that the question of the mode and tempo of the proto-Polynesian expansion has by no means been answered conclusively.

In contrast to the linguistic and genetic homogeneity of the Polynesians, the people of Papua New Guinea (PNG) and island Melanesia are heterogeneous, reflecting the ancient settlement of the western Pacific (Stoneking *et al.*

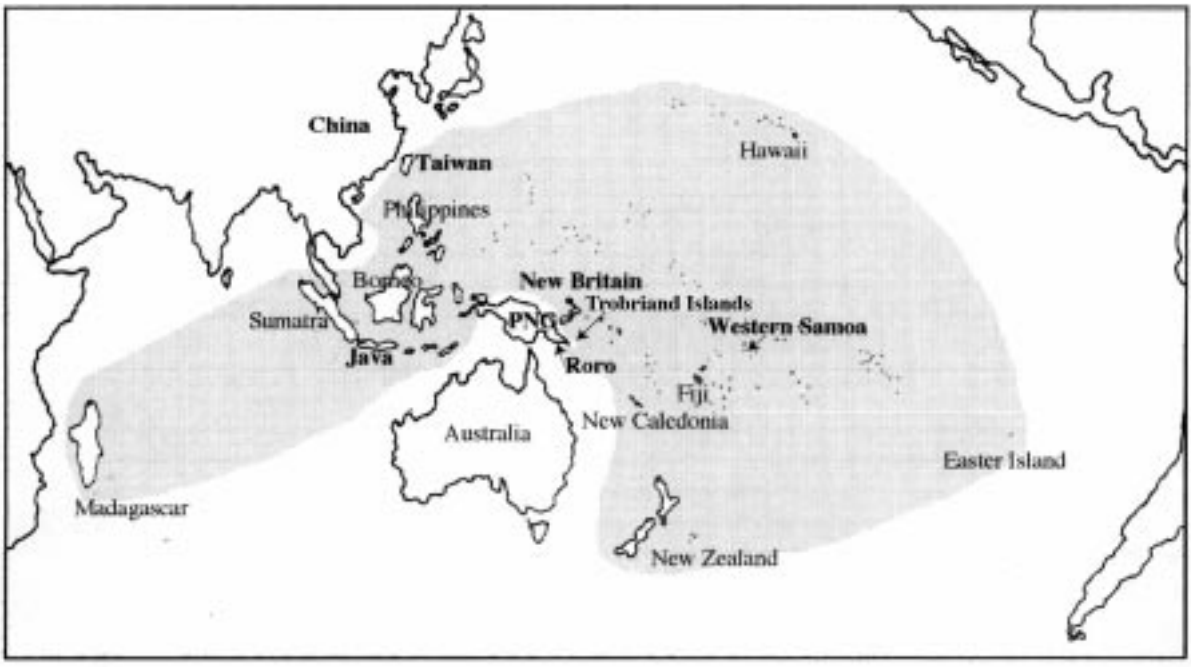

Figure 1. Map of the Pacific, indicating the spread of the Austronesian languages. Polynesia is the large triangular area bounded by Hawaii, New Zealand and Easter Island. The geographical origin of the eight population samples in our study is indicated by bold typeface. (Adapted from Bellwood *et al*. (1995).)

1990). It is thought that modern humans reached PNG about 60 000 years ago, and archaic humans possibly much earlier (Morwood *et al*. 1998). The peoples of New Guinea speak several hundred different languages (known as non-Austronesian or Papuan languages), which evolved as a result of the isolation of small tribal groups in the remote parts of this huge and rugged island. The subsequent arrival of the Austronesian seafarers is thought to have displaced Papuan speakers from some coastal and many island areas of Melanesia, areas inhabited today by people who speak Austronesian languages although they are phenotypically 'Melanesian'.

To investigate the genetic affinities of the Austronesian-speaking peoples, we analysed mitochondrial DNA, HLA and Y-chromosome polymorphisms in present-day individuals from eight geographical locations in Asia and the Pacific. The locations chosen were China, Taiwan, Java, New Guinea highlands, New Guinea coast, Trobriand Islands, New Britain and Western Samoa. The Han Chinese were taken to represent an Asian mainland population, while the Taiwan aboriginals and Javanese are supposedly the descendants of early Austronesian settlers. The New Guinea highlanders are derived from a much earlier expansion of people and speak languages unrelated to the Austronesians. In contrast, the Roro from the south coast of New Guinea, the Trobriand Islanders and the Tolai of New Britain are all Austronesian-speaking inhabitants of Melanesia. Finally, the Western Samoans are Polynesians, the last branch of the Austronesian expansion. A total of 290 individuals were surveyed for polymorphisms in the first segment of the hypervariable mtDNA non-coding region. In addition, DNA of 228 males from the eight populations was subjected to Y-chromosome microsatellite analysis, to obtain information on the evolutionary history of the male lineages. We also examined nucleotide sequence polymorphisms at several HLA loci in 346 individuals.

(b) *Analysis of different genetic components*

MtDNA is an excellent tool for the study of human evolutionary history and the genetic affinities of peoples of different regions of the world because it has a number of useful features, including simple organization, maternal mode of inheritance and a relatively fast rate of evolution (Cann *et al*. 1987; Vigilant *et al*. 1991; Wallace 1995). Previous studies of mtDNA variation in the Pacific have shown that a particular mutation, consisting of a deletion of nine base pairs between the mtDNA cytochrome oxidase II and lysyl tRNA genes, is virtually fixed in Polynesia (Wrischnik *et al*. 1987; Hertzberg *et al*. 1989). In the Pacific, the nine-base-pair (9-bp) deletion is associated with a number of specific base substitutions in the hypervariable non-coding region of the mitochondrial genome, compared to the published reference sequence (Anderson *et al*. 1981). These substitutions, known as the Polynesian haplotype or Polynesian motif, have been detected in ancient human skeletal remains and living populations of the Pacific (Hagelberg & Clegg 1993; Hagelberg *et al*. 1994; Lum *et al*. 1994; Melton *et al*. 1995; Redd *et al*. 1995; Sykes *et al*. 1995) and Madagascar (Soodyal *et al*. 1995). It was suggested that the 9-bp deletion originated in mainland Asia. A substitution at position 16217 arose on the background of the 9-bp deletion, followed by a substitution at position 16261. It was proposed that the full Polynesian motif (16189; 16217; 16247; 16261) arose in east Indonesia, and a group of proto-Polynesians bearing the motif started to expand into the Pacific about 5500 years ago (Melton *et al*. 1995; Redd *et al*. 1995). This is consistent with archaeological and linguistic evidence for the spread of Neolithic

agriculturalists from mainland South-east Asia to island Southeast Asia and a subsequent migration of Polynesians east into the Pacific.

However, these conclusions are based on the maternally inherited mtDNA locus. How about the paternal lineages? In contrast to the ease of study and apparently simple interpretation of the mtDNA data, analyses of Y-chromosome polymorphisms are more cumbersome and this area of study has been slower to develop than mtDNA research. There is increasing interest in the use of the non-recombining portion of the Y chromosome for inferring evolutionary events (for review, see Jobling & Tyler-Smith 1995), but few data are available on Y-chromosome polymorphisms in Oceania, showing mainly the low degree of information inherent in some commonly used Y-chromosomal restriction fragment length polymorphisms (RFLPs) (Spurdle *et al*. 1994). A previously described Y-chromosome Alu insertion (YAP) (Hammer 1994) has up to now proved uninformative for Pacific populations (Spurdle *et al*. 1994; Hammer *et al*. 1997, 1998). A number of polymorphic Y-chromosome point mutations have been detected in populations outside Oceania (Seielstad *et al*. 1994; Hammer 1995; Whitfield *et al*. 1995; Underhill *et al*. 1996; Veitia *et al*. 1997; Zerjal *et al*. 1997), but as haplotype studies have shown that several of the currently known biallelic markers are on a YAP+ background, they will probably not be useful for the study of Pacific populations. A transition of C to T at position 9138 of the SRY region was observed in a male of the Bougainville Islands (Whitfield *et al*. 1995), but is thought to be a very recent or private mutation since it has not been observed in other males from the Pacific region (Hammer *et al*. 1998; Hurles *et al*. 1998; M. Kayser, unpublished observations).

This situation is due to change with the discovery of four informative Y-chromosome point mutations which are polymorphic in Oceania (Underhill *et al*. 1997). Previous to this, the only appropriate markers for the study of male lineages in the Pacific were the fast-evolving microsatellite polymorphisms on the Y chromosome (Roewer *et al*. 1996; Kayser *et al*. 1997). A worldwide study on the frequency of these polymorphisms in human populations, including Asia and Oceania, was published recently (Deka *et al*. 1996). For the present study, we selected seven microsatellites to define a haplotype, based on the following criteria: locus definition (DNA sequence and mapping data), comparable mutation rates at different loci, high genetic diversity values and availability of global haplotype data for comparison. The microsatellite loci chosen—DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393—were previously used in a worldwide study of Y-chromosome variation (de Knijff *et al*. 1997; Kayser *et al*. 1999).

Finally, we performed analyses on four different HLA class II loci: DRB1, DQA1, DQB1 and DPB1. HLAs, or human leukocyte antigens (sometimes referred to as human lymphocyte antigens), are proteins on the surface of white blood cells, which are involved in the immune response and are important in organ transplantation. Not surprisingly, they are highly polymorphic and have yielded valuable information about human population affinities. Although traditionally analysed by serological methods, they are increasingly studied using molecular techniques. A number of studies have been carried out on Pacific populations based on serological and molecular HLA analysis methods (for review, see Serjeantson 1989; Gao *et al*. 1992*a*; Serjeantson & Gao 1995). The results support the most widely accepted model of the colonization of Polynesia, indicating shared features between coastal PNG and island Melanesia, but no genetic overlap between Melanesia and Polynesia. The main conclusion of the HLA studies is that the early Austronesian voyagers left a genetic footprint in coastal and island Melanesia, but did not carry Melanesian genes into Polynesia. This supports the 'fast train to Polynesia' hypothesis.

2. MATERIALS AND METHODS

(a) *DNA samples*

Our sample consisted of 364 individuals from eight different locations in Asia and the Pacific. These included Han Chinese, Javanese (rural schoolchildren from an area south of Jakarta), Taiwanese (aboriginals from the four tribes Ami, Atayal, Bunun and Paiwan), PNG highlanders (from four different tribes), Roro people from the south coast of PNG, Trobriand islanders living off the east coast of PNG, Tolais of the Melanesian archipelago of New Britain and Polynesians from Western Samoa. Blood samples were taken from the individuals involved in this study with full informed consent. DNA was extracted from blood by conventional phenol–chloroform extraction methods. MtDNA analysis was carried out on 290 individuals, 228 males were subjected to Y-chromosome microsatellite analysis and 346 individuals to HLA analysis.

(b) *Mitochondrial DNA analysis*

DNA samples were polymerase chain reaction (PCR) amplified in 20-ml reaction volumes, using primers HVR1 (5′-CTAACCTGAATCGGAGGACAAC-3′) and HVR4 (5′-GCATACCGCCAAAAGATAAAA-3′) specific for a 1239-bp fragment of human mtDNA, under the conditions described previously (Hagelberg & Clegg 1991). The PCR fragments were subjected to agarose gel electrophoresis and visualized under UV light. A 1 µl aliquot of each PCR product was reamplified with internal primers and subjected to automated DNA sequencing on a LI-COR 4200 sequencing system (MWG-Biotech, Milton Keynes, UK). We obtained approximately 360-bp sequence between nucleotide positions 16041 and 16400 of the mtDNA reference sequence for most individuals. Samples exhibiting a T to C transition at position 16189 had a homopolymeric tract of usually 12 C residues that prevented sequencing past this point. In these cases, two partial single sequences that joined at the C tract were obtained. The DNA sequences were aligned manually. For the analysis of the intergenic 9-bp deletion, PCR amplification was carried out as described previously (Hagelberg & Clegg 1993; Thomas *et al*. 1998).

(c) *Y-chromosome microsatellite analysis*

The Y-chromosomal microsatellites (also known as short tandem repeats or STRs) DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393, were analysed as described previously (Kayser *et al*. 1997). Since the DYS389II PCR product contains the DYS389I product, the number of repeats observed at the DYS389I locus was subtracted from those at DYS389II to avoid counting the DYS389I repeats twice.

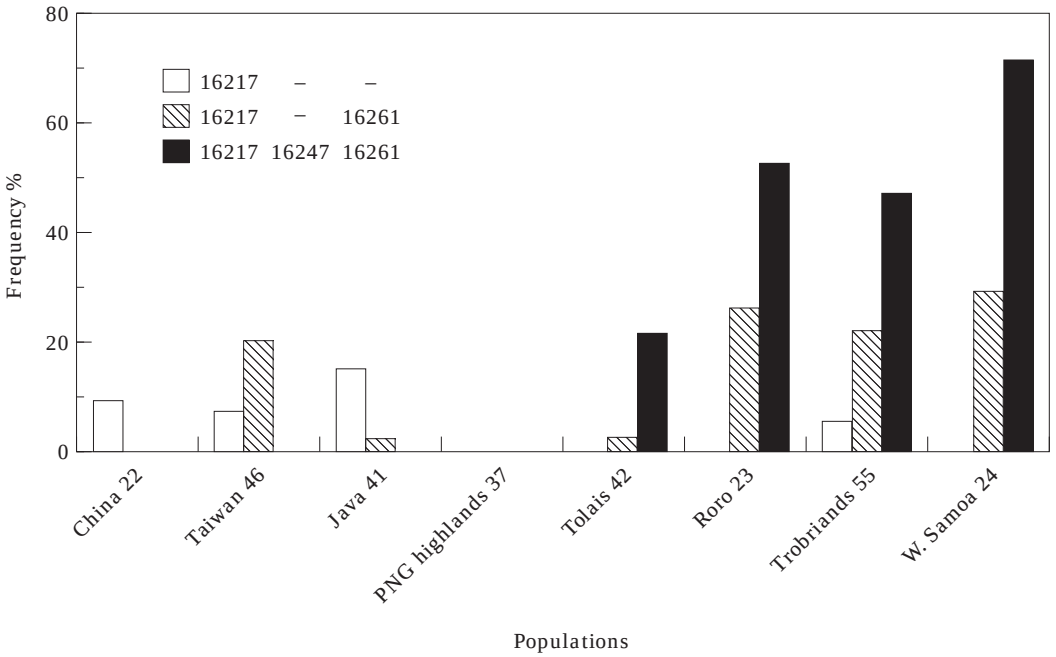

Figure 2. Diagram of the frequency of the Polynesian mtDNA motif (16217/16247/16261) and ancestral mtDNA types (16217 and 16217/16261) in the eight populations of this study. The number next to the population name refers to the number of individuals in the sample.

For the sequence analysis of the DYS390 locus, PCR was done using one primer labelled with indodicarbocyanin-(Cy5)-phosphoramidite and the other labelled at the 5′ end with biotin. Single-stranded sequencing templates were prepared using magnetic beads (Dynal, Hamburg, Germany), following the manufacturer's instructions. Cycle sequencing reactions were carried out using the Sequitherm Cycle Sequencing Kit (Epicentre Technologies, Madison, USA), and sequencing was done on an A.L.F.*express* automated DNA sequencer at 1500 V, 38 mA, using a 6% Long Ranger Gel (FMC Bioproducts, Rockland, USA).

(d) *HLA analysis*

We analysed nucleotide sequence polymorphisms at the loci DRB1, DQA1, DQB1 and DPB1, using sequence-specific oligonucleotide typing (SSO), sequence-specific PCR (SSP) and sequence-based typing, as described previously (Nagy *et al.* 1997), and using the same thermocycler and DNA sequencer described in the previous section. The allele frequencies were obtained by direct counting. Hardy–Weinberg analysis was performed on the Roro, Tolai and Trobriand population. The haplotype frequencies and linkage disequilibrium values for the two locus haplotypes (HLA DRB1, HLA DQB1) were calculated using 2×2 contingency tables.

(e) *Phylogenetic analysis*

The mtDNA and Y-chromosome haplotype discriminance was calculated for each population simply by dividing the number of haplotypes by the number of individuals in the population sample. The mtDNA and Y-chromosome microsatellite haplotype diversity was calculated to obtain an estimate of the variation in each population, using the following equation (Nei 1987):

$$h = \frac{(1 - \sum x^2)n}{n - 1},$$

where h is the haplotype diversity, n the sample size and x the frequency of each haplotype.

An analysis of molecular variance, AMOVA (Excoffier *et al.* 1992), was applied to Y-chromosomal microsatellite haplotypes as described previously (Roewer *et al.* 1996). Φ_{st} values were calculated for each population pair and significance levels obtained by comparing the observed Φ_{st} values to the empirical null distribution from 10 000 randomizations. A neighbour-joining tree (Saitou & Nei 1987) based on the observed Φ_{st} values was drawn using the NEIGHBOR and DRAWTREE programs of the PHYLIP package (Felsenstein 1989). Phylogenetic analysis of the mtDNA haplotypes was performed using the neighbour-joining method, using the programs described above. For the HLA data, the Nei genetic distance calculation was performed using version 3.5c of the PHYLIP package (Felsenstein 1995).

3. RESULTS

(a) *Mitochondrial DNA*

We obtained DNA sequence information for the first hypervariable segment of the mtDNA noncoding control region of 290 individuals. We typically obtained 360 bp of double-stranded sequence between nucleotide positions 16041 and 16400 of the reference sequence (Anderson *et al.* 1981), with 116 variable sites, 244 fixed sites and three insertions in the 360 base pairs sequenced. There were 147 different mtDNA types among the 290 individuals studied. The most abundant mtDNA type observed was the Polynesian motif, characterized by the 9-bp deletion and base transitions at positions 16189 (T to C), 16217 (T to C), 16247 (A to G) and 16261 (C to T) of the mtDNA reference sequence (Anderson *et al.* 1981). This mtDNA type (CCGT) was observed in Western Samoans, Trobrianders, Roro people and Tolais. None of the Taiwanese exhibited the Polynesian motif, but 20% of the sample of 46 individuals had the substitutions 16189, 16217 and 16261 (CCAT), an mtDNA type which is abundant in Polynesia and thought to be ancestral to the Polynesian

Table 1. *Haplotype discriminance and haplotype diversity values for mitochondrial DNA sequence data and seven-locus Y-chromosome microsatellites*

(For the mtDNA data, the frequency of the 9-bp deletion and the Polynesian motif and related haplotypes (16217/16247/16261 and 16217/16261) is also given. The haplotype diversity of the Y-chromosome microsatellites is markedly higher in Polynesia than the mtDNA diversity.)

	mitochondrial DNA sequences						Y-chromosome microsatellites			
population	*n*	haplotype number	haplotype discrimi-nance	haplotype diversity	9-bp deletion %	Polynesian mtDNA haplo-types (CCAT and CCGT) %	*n*	haplotype number	haplotype discrimi-nance	haplotype diversity
China	22	21	0.96	1.00	16	0	36	34	0.94	1.00
Taiwan	46	28	0.61	0.97	33	20	27	22	0.82	0.95
Java	41	33	0.81	0.99	29	2	53	38	0.72	0.94
PNG highlands	37	31	0.84	0.99	0	0	15	11	0.73	0.85
Tolais	42	17	0.41	0.83	24	23	16	14	0.88	0.98
PNG coast/Roro	23	8	0.35	0.72	78	78	12	7	0.58	0.86
Trobriands	55	14	0.26	0.81	75	69	59	40	0.68	0.98
Western Samoa	24	4	0.17	0.59	100	100	10	8	0.80	0.93

motif (Melton *et al*. 1995; Redd *et al*. 1995). These two mtDNA types, and haplotypes deviating by one substitution from these types, were present in 100% of the Samoans, 78% of the Roro, 69% of the Trobrianders and 23% of the Tolais. The 9-bp deletion and control region polymorphisms at positions 16189 and 16217, also ancestral to the Polynesian motif, were present in Chinese, Javanese, Taiwan aboriginals and Trobrianders, but absent in PNG highlanders (figure 2).

The lowest mtDNA haplotype diversity was observed in Western Samoa, with only four haplotypes among 24 individuals and a high degree of haplotype sharing. In contrast, there was high haplotype diversity among the Chinese, Javanese and New Guinea highlanders, consistent with the greater antiquity of these populations (table 1).

A neighbour-joining tree of the 147 different mtDNA haplotypes is shown in figure 3. Haplogroup I, characterized by the 9-bp deletion (and including the Polynesian motif) forms a large distinct cluster, with its deepest branches in Java, Taiwan and China, and the youngest haplotypes in Samoa and island Melanesia (Tolais, Trobriands and Roro). However, the deepest branches of the entire tree belong to non-9-bp deleted lineages, mainly from Java, Taiwan and China. Most of the PNG haplotypes fell into two discrete clusters, containing the haplogroup II and haplogroup III mtDNA types previously observed by us in New Guinea and island Melanesia (Hagelberg *et al*. 1999).

(b) *Y-chromosome microsatellites*

We obtained full haplotype information on seven Y-chromosome microsatellites for 228 males of the eight populations in our study. A total of 161 different seven-locus haplotypes was observed, 150 of which were population specific (not shared by individuals belonging to the different populations in our sample) and 127 occurring in one individual only. The haplotype diversity was generally high, being highest in the Chinese sample (over 0.99), and in the Trobrianders and Tolai (0.98 respectively). The haplotype diversity was lowest in the New Guinea highlanders (0.85) and the Roro (0.86), indicating more genetically homogeneous populations. Curiously, the haplotype diversity in the PNG highlanders was significantly lower than in the West Samoans, a much younger population, although sample numbers were small in both cases (table 1). Nevertheless, haplotype diversity was above 0.85 in all eight populations, reflecting a large number of male founders or, alternatively, an ancient settlement by few founding male lineages which have had time to diversify.

A pairwise AMOVA approach based on the seven locus microsatellite haplotypes was used to calculate the genetic distance measure Φ_{st} (Roewer *et al*. 1996; de Knijff *et al*. 1997). The Φ_{st} values reflect the proportion of the molecular variance explicable in terms of the population differences between the frequency of the Y-chromosome microsatellite haplotypes and can therefore be used as an index of population differentiation. One might assume that the smaller the Φ_{st} value, the closer the relationship between two populations. For all population pairwise comparisons, the interpopulation variance was lower than the intrapopulation variance. The differences in the Φ_{st} values between all populations were significant, with the exception of PNG and Java. The Φ_{st} values between the Chinese, Taiwanese, Javanese, Tolais and Trobrianders were low, mostly below 0.1, and indicated a marked distance between these populations and the New Guineans, Roro people and Samoans (table 2). A similar relationship was observed in the number of haplotypes shared by the different populations. Although the total number of shared haplotypes was low, most were shared between Chinese, Taiwanese, Javanese, Tolais and Trobrianders. The unrooted neighbour-joining tree based on the Φ_{st} values (figure 4) reflects this relationship, with the Roro, New Guineans and Samoans placed at the end of the longest branches.

(c) *HLA analysis*

Table 3 lists the most common HLA DRB1 and DQB1 haplotypes in the populations studied. The results of the HLA analyses showed reduced genetic diversity in island Melanesia. The number of different HLA alleles in the

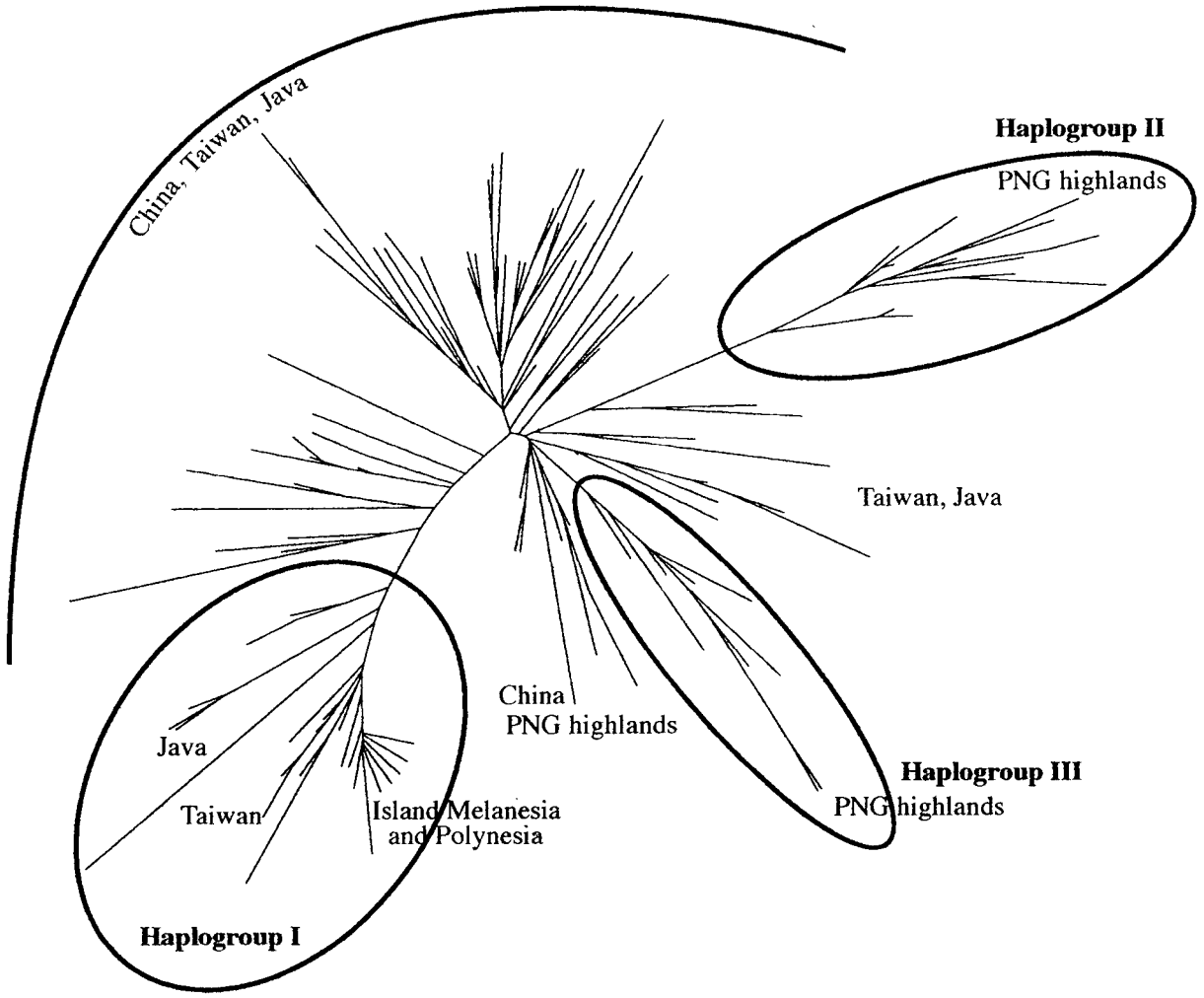

Figure 3. Neighbour-joining tree of the 147 different mtDNA haplotypes observed among 290 individuals. The Polynesian motif and ancestral mtDNA types (with the 9-bp deletion) form group I. The majority of the New Guinea highland lineages cluster in group II and group III.

Table 2. *Y-chromosome microsatellite haplotype analysis of eight populations* ($n = 228$)

(Φ_{st} values generated by the analysis of molecular variance (AMOVA) are shown above the diagonal. The number of haplotypes shared between different populations is shown below the diagonal.)

	China	Taiwan	Java	Tolai	Trobriands	PNG highlands	Roro	Western Samoa
China	—	0.054	0.063	0.071	0.063	0.175	0.232	0.283
Taiwan	2	—	0.078	0.126	0.062	0.239	0.333	0.374
Java	1	3	—	0.064	0.063	0.052	0.135	0.297
Tolai	0	0	0	—	0.09	0.132	0.156	0.286
Trobriands	0	2	2	1	—	0.115	0.224	0.263
PNG highlands	0	0	3	0	0	—	0.215	0.44
Roro	0	0	0	0	0	0	—	0.422
Western Samoa	1	0	0	0	0	0	0	—

three populations, Trobrianders, Roro and Tolai, was small, with only 9–14 different alleles at the DRB1 locus (compared with 31 in a German sample), and four to seven different alleles at the DPB1 locus (compared with 17 in Germans). The most salient feature of the HLA analysis was the remarkably high frequency of the previously described DPB1 allele 0501 in the Trobriand islanders, Samoans, Roro and Taiwanese aboriginals (figure 5). This was in contrast to the relatively low frequency of the allele in the Austronesian-speaking Tolais of New Britain.

There was no significant deviation of the genotype frequencies from Hardy–Weinberg equilibrium (goodness-of-fit test for the most frequent genotypes). Because genetic data of complete families were not available, class II haplotypes were inferred from the analysis of linkage disequilibrium patterns of alleles in homozygous and heterozygous

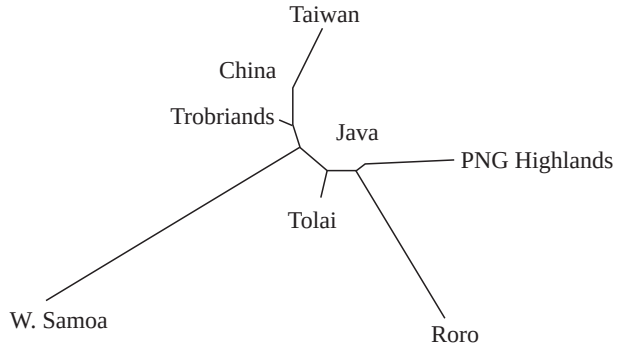

Figure 4. Neighbour-joining tree of eight population samples based on the Φ_{st} values generated from seven-locus Y-chromosome microsatellite haplotypes.

individuals. The DRB1-DQA1-DQB1 haplotypes of the Trobrianders and Roro were the same as those commonly observed in other Asian populations. The most common haplotype in the Trobriands (33%) DRB1*0803-DQA1*0103-DQB1*0601, is also abundant in Koreans and Japanese (Tsuji *et al*. 1992; Hashimoto *et al*. 1994). The haplotypes, DRB1*15021-DQB1*0502 and DRB1*0901-DQB1*0601, observed in the Trobriands and the Roro, have not been detected in any other populations. The most abundant haplotype in the Tolais, DRB1*1101-DQB1*0301, was the same as the abundant 'Melanesian haplotype' observed previously in New Britain, New Caledonia and Fiji (Gao & Serjeantson 1992).

The unrooted neighbour-joining tree of the DRB1 and DPB1 loci showed that the Trobrianders, Roro, Taiwanese and Samoans were the closest populations to each other, and most distant to the PNG highlanders and the Tolais (figure 6). The Javanese fell out from all the other populations in the study, with no overlap in the haplotypes observed. The frequency of the typical Asian DPB1 allele 0501 in our Java sample was low (14%) and comparable to the level observed in the PNG highlanders (13%).

4. DISCUSSION

The aim of this study was to investigate the genetic affinities of the Austronesian-speaking peoples of the Pacific, using information from three different genetic systems, namely mtDNA, Y-chromosome microsatellites and HLA polymorphisms. Our results are an extension of a previous study on HLA alone by several of us (Nagy *et al*. 1997), where we addressed the question, 'Are the Trobriand islanders emigrants of South-east Asia?' This study revealed a remarkably high frequency of the previously described 'Asian' DPB1 0501 allele in the Trobrianders (98%), evidence of the ultimate Asian ancestry of these people. The frequency of this allele was 47% in Chinese, 70% in Taiwanese, 71% in the Roro and 70% in Samoans, but only 13% in New Guinea highlanders and 14% in the Javanese (compare this with a frequency of 1% in white French, 2% in white Germans, 1% in African Americans and 1% in black South Africans). The reduced levels of HLA diversity in the Austronesian-speaking Trobrianders, Roro and Tolais could be explained by a genetic bottleneck in the founding of these populations, followed by a population expansion. However, there was little overlap between the HLA haplotypes of the Trobrianders and Roro (and Taiwanese and Samoans) on one hand, and the Tolais on the other. Although these are all Austronesian-speaking populations, it is clear that they have different settlement histories. Interestingly, the most common HLA haplotype in the Trobriand islanders (33% of DRB1*0803-DQA1*0103-DQB1*0601), also abundant in Japanese and Koreans, is common in Australian aborigines from Cape York (29%) (Gao *et al*. 1992*a*,*b*).

In general, the mtDNA and HLA data gave a similar picture. The HLA data (figure 6), revealed close similarities between the Taiwanese, Trobrianders, Roro and Samoans, which supports the view of a Taiwanese origin of the Austronesian-speaking peoples of the Pacific. The mtDNA data showed a very high frequency of the Polynesian motif and/or variant lacking the transition at 16247 in Samoans (100%), Roro (78%) and Trobrianders (69%), and a moderately high frequency in Tolais (23%) and Taiwanese (20%). The mtDNA data support the view that the proto-Polynesians expanded from an ancestral population in Taiwan. In our study, the Chinese and Javanese with the 9-bp deletion lacked the Polynesian motif and their position in the phylogenetic tree (figure 3) suggests that they derived from an older population expansion than the proto-Polynesians. The full Polynesian motif has been previously detected in east Indonesia (Melton *et al*. 1995; Redd *et al*. 1995), but we observed only the variant lacking 16247 in a single individual of Java, further west in Indonesia. The Polynesian motif might have originated in east Indonesia, as previously suggested (Redd *et al*. 1995), but it could also have been introduced into east Indonesia from Polynesia or coastal PNG. The direction of the currents (east to west) favours this trajectory (for example, William Bligh and the other *Bounty* survivors sailed straight from the central Pacific through the Torres Straight to Timor in east Indonesia with relative ease in an open boat). Based on the available evidence, the place of origin of the Polynesian motif is still a matter of conjecture, although the ancestral form (lacking the 16247 transition) is likely to have originated or expanded in Taiwan.

Linguistic evidence suggests that Austronesian language speakers expanded early into Taiwan. This expansion was characterized by settled cereal and tuber cultivation and a knowledge of seafaring. About 4000 years ago, the Austronesians expanded further south to Borneo, Sulawesi and the Moluccas, eventually settling parts of coastal New Guinea and island Melanesia and moving out much later to colonize the whole of Polynesia (Bellwood 1989). However, well before the arrival of the Austronesians, the western Pacific was occupied by foragers descended from Palaeolithic peoples. In the course of some of the earliest migrations by anatomically modern humans, people moved out of mainland Asia into island South-east Asia, Australia and Papua New Guinea. By 30 000 years ago, people were able to cross quite respectable sea distances and to move east into the Pacific to reach the Solomons in island Melanesia (Wickler & Spriggs 1988). Modern-day Australian aborigines and New Guinea highlanders have genetic similarities (Roberts-Thomson *et al*. 1996), but their language and culture have diverged. New Guineans adopted horticulture and expanded and diversified in the highlands, many

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 3. *Most abundant HLA DRB1-DQB1 haplotypes in the eight populations in this study*

population origin	DRBl	DQBl	%
China	0901	03032	28
Taiwan	0803	06011	9
Java	1202	0301	50
PNG highlands	1408	05	27
New Britain, Tolai	1101	0301	54
PNG coast, Roro	1602	0502	23
Trobriand Islands	08032	0601	33
Western Samoa	0901	03	45

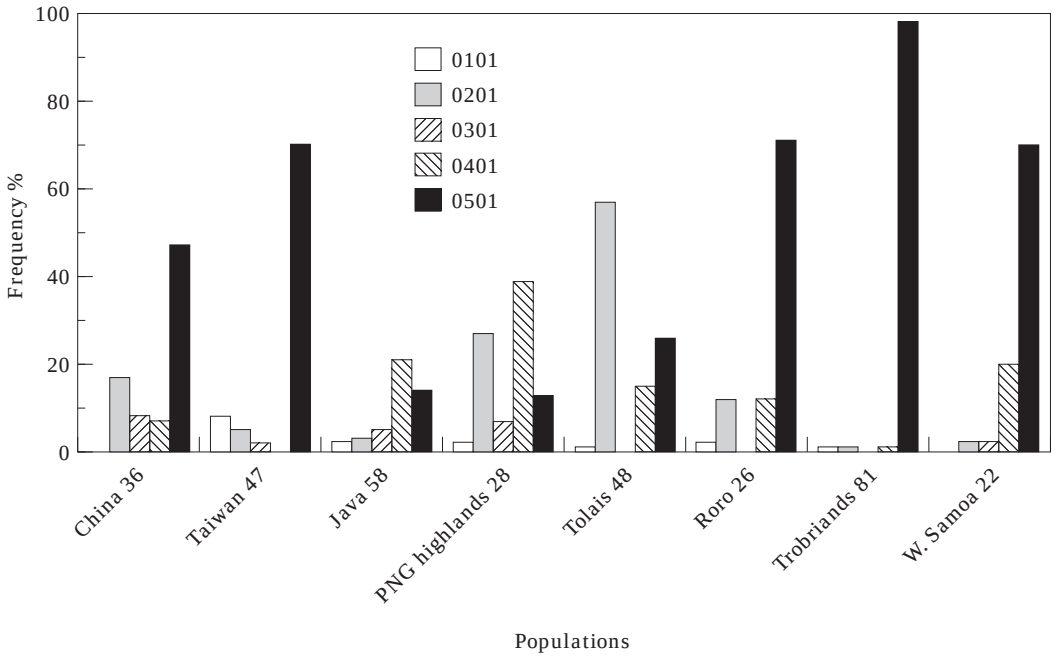

Figure 5. Frequency of several HLA-DPB1 alleles in the eight populations of our study. Allele 0501 is present at very high frequencies in several Asian and Pacific populations. The number next to the population name refers to the number of individuals in the sample.

groups remaining isolated until this century. Foragers exhibiting Melanesian physical characteristics are still found today in parts of island South-east Asia, such as the Philippines (Bellwood 1995, 1996).

The development of agriculture gave the Austronesians a demographic advantage and they quickly made incursions into many of the areas previously settled by the earlier Australoid and Papuan settlers. Their occupation of South-east Asia and coastal and island regions of Melanesia is evident from a rich archaeological record. In island Melanesia, the Austronesian expansion is linked to a number of developments known by archaeologists as the 'Lapita cultural complex' and characterized by trade over vast distances, manufacture of ornate pottery, tuber cultivation and the introduction of domesticated animals (pigs, chickens and dogs). Most prehistorians adhere to the view that the Lapita culture derived from the Neolithic expansion in South-east Asia, and that the Lapita people were the first colonizers of the previously inhabited region of the central Pacific and eastern Polynesia (Spriggs 1997).

The HLA and mtDNA data certainly support the idea of close genetic affinities between the Austronesians in Taiwan, the south coast of New Guinea, the Trobriand Islands and Samoa. It is evident that the Austronesian newcomers displaced earlier settlers from many parts of island Melanesia and coastal PNG. However, the genetic influence of the Austronesians on the Tolais of New Britain was less marked, as these people show a reduced frequency of Polynesian-like mtDNA types (23%) and the 'Asian' HLA DPBl 0501 allele (26%), and their most common HLA haplotype is the 'Melanesian' type shared with New Caledonians and Fijians (Serjeantson & Gao 1995). The Tolais of New Britain are phenotypically 'Melanesian', with fairly dark skin and frizzy hair, sometimes almost blonde as in some highland Papuan groups. However, the Tolai language is Austronesian. As mentioned above, these parts of island Melanesia were first settled well before the arrival of the Austronesians and the social organization of the Tolais has many elements which are typical of Papuan peoples (Epstein 1991). Austronesian languages might have been adopted by these Papuan peoples of island Melanesia to facilitate trade with the Austronesian newcomers.

In contrast, the Trobrianders and Roro people have close genetic affinities to both Taiwanese and, more markedly, Samoans. In fact, our Trobriander and Roro samples have 69% and 78% of Polynesian mtDNA types. The Trobriand Islands are an archipelago in the Solomon Sea,

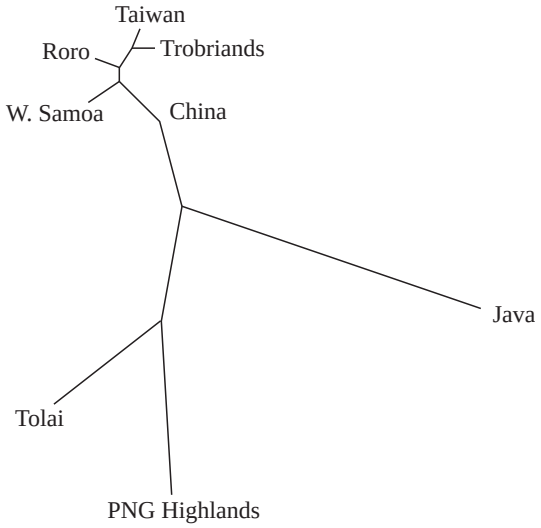

Figure 6. Neighbour-joining tree of the eight population samples typed for the HLA loci DRB1 and DPB1. The Taiwanese, Roro, Trobrianders and Samoans cluster closely together, but apart from the Austronesian-speaking Tolais and Javanese, and the non-Austronesian New Guinea highlanders.

north of the easternmost tip of Papua New Guinea. The indigenous inhabitants of this archipelago number about 25 000. The Trobrianders are typical Austronesian people: they speak an Austronesian language, have matrilinear rules of descent, a stratified society with powerful male chiefs who inherit their position through the maternal lineage and a competitive social organization. They are skilled navigators and build single outrigger dugout canoes and have a wide sphere of inter-island contacts and ritualized exchange of prestige goods, known as *kula*. Whereas the Trobriand culture is distinctively Austronesian and the language is also Austronesian, the physical appearance of the people suggests that they are the descendants of both Papuan and Austronesian peoples. A Papuan language is still spoken in Rossel, a small island south of the Trobriands, which suggests that the people here might represent the original Papuan population, superseded elsewhere by one or several waves of Austronesian immigrants (Liep 1991).

The Roro are one of the distinct ethnic groups in the Central Province of Papua New Guinea. They live on Yule Island and on the coast of the mainland near Bereina, approximately 200 km from Port Moresby. The Roro speak an Austronesian language similar to the language of the Motu, the native population of the Port Moresby region. These languages are similar to the present-day languages of the Indonesian archipelago and quite different to the Austronesian languages of the Trobriands. Although the country west of here is inhabited exclusively by Papuan speakers, the coast towards the east and around the tip of New Guinea to the north coast, as well as many islands off the east and north coasts are inhabited by Austronesian peoples. One could speak of an Austronesian continuum stretching from the Moluccas along the north coast of New Guinea to the Bismarck Archipelago and around the south-eastern tip New Guinea.

The Roro are coastal people but do not have developed navigational skills like the Trobrianders. The dominant phenotype appears to be Papuan, with traits denoting admixture with Austronesians. Lines of descent are patrilineal and political power is held by men, like in other Papuan groups. The modern type of Austronesian spoken by the Motu and Roro indicates a relatively recent arrival, supported by the archaeological evidence. Papuan-looking people as far west as the Fly River use *Piper methysticum*, the Polynesian *kava*, to make an intoxicating drink. This would suggest cultural contacts with Polynesians from the east.

Our genetic evidence indicates that despite the Papuan appearance of the Roro, they have a predominance of Austronesian mtDNA types, notably the Polynesian motif. Our mtDNA data are highly suggestive of a recent migration of Polynesian maternal lineages to island Melanesia and the south coast of PNG. Rather than being the descendants of Austronesian language speaking proto-Polynesian colonizers, the remarkable affinity between Trobrianders, Roro and Polynesian people argues for a recent migration of people east from Polynesia into island Melanesia and the south coast of PNG.

As regards the Y-chromosome microsatellite analysis based on seven different microsatellite loci, the most salient feature was the high degree of haplotype diversity in all the populations studied. The highest haplotype diversity was in the Han Chinese sample (over 0.99), not surprisingly, but it was astonishingly high in the relatively endogamous Trobrianders and the Tolai (0.98 respectively). The lowest haplotype diversity was in the New Guinea highlanders, suggesting settlement by a small number of founding males and comparative isolation. The sample of males from Samoa was small, only ten individuals, but eight different haplotypes were observed. The high diversity of the Y-chromosome haplotypes in Samoa contrasts with the low mtDNA diversity and can be explained by possible differences in the migration patterns of males and females. Our Y-chromosome results indicate a comparatively close relationship between the male lineages in China, Taiwan, Java and in the Austronesian-speaking Tolais and Trobriand islanders. The apparently large genetic distance between these populations and the Roro, PNG highlanders and Samoans can be explained partly by the small sample size of Roro and Samoan males. However, the large difference in Φ_{st} values between the Samoans and other populations and their large separation in the tree (figure 4), is probably caused by the high frequency (seven out of ten men) of very short alleles at microsatellite locus DYS390, created by a large-scale mutation characterized by a deletion in the repetitive sequence (see below). This mutation would have led to a drastic modification of the haplotype and an extreme shift in Φ_{st} value.

The relatively constant high haplotype diversity in the Y-chromosome microsatellites contrasts with the gradual reduction in mtDNA diversity which is observed from west to east in the Pacific (table 1). The chief deviation from this trend was the population sample from the New Guinea highlands, which exhibited high mtDNA diversity and relatively low Y-chromosome microsatellite diversity. However, the mtDNA phylogenetic tree indicated that the PNG highlanders expanded recently from a small number of founders, which is consistent with archaeological data suggesting an intensification of horticulture

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

in the highlands in the last few thousand years (Bayliss-Smith 1996).

One of the problems associated with population genetic analyses using Y-chromosome microsatellites is the high rate of mutation, with empirical values of 0.2% to 0.32% (Heyer et al. 1997; Kayser *et al*. 1997), which might prevent events older than 5400–8000 years, or even less, to be traced (de Knijff *et al*. 1997). Ideally, microsatellites should be combined with other Y-chromosomal markers with a slower mutation rate. One of these markers is a specific deletion at the DYS390 locus, mentioned above, which results in extremely short alleles (19: 195 bp, 20: 199 bp and 21: 203 bp, rarely 22: 207 bp) and which was recently identified by us (Forster *et al*. 1998). So far, the deletion has only been observed in PNG, island Melanesia and Polynesia in a survey of over 100 alleles of different lengths in 14 world populations (M. Kayser and P. de Knijff, unpublished observations). Of the populations included here, the specific sequence motif was detected in the Trobriand Islands (9%), the Tolai of New Britain (19%), the Roro south coast of PNG (17%) and in Western Samoa (70%) but not in China, Java, Taiwan or PNG highlands (although it was observed in 6% of a different sample of PNG highlanders and in 25% of a sample of north coast New Guineans; Forster *et al*. 1998). The mutation seems to have a single origin and since it appears to be specific to Pacific populations, it will be a useful Y-chromosome marker in future population studies (Kayser *et al*. 1999). It is not yet clear whether the mutation arose in an Austronesian population and was carried into the Pacific by proto-Polynesians or whether it originated in Polynesia and was carried back to island Melanesia in a recent back migration. The low frequency of the mutation in the PNG highlands (Forster *et al*. 1998) could be due to admixture with coastal populations, where the frequecy of the deletion is much higher. Additional studies are needed, including more populations and individuals, to reconstruct the origin of the DYS390 marker. It would be particularly desirable to carry out analyses on single nucleotide polymorphisms to help trace the history of these populations. Four biallelic markers (M4, M5, M16, M21) became available recently (Underhill *et al*. 1997) and a study incorporating these markers is currently in progress, but more will be needed to increase the resolution of Y-chromosome analyses.

In conclusion, our genetic study of eight populations of Asia and the Pacific supports archaeological and linguistic evidence for the expansion of the Austronesian-language speakers in Taiwan, from an ultimate source population in mainland Asia. However, our study revealed only low genetic affinities between the Austronesian-language speakers in Java and the Austronesian Roro, Trobrianders and Samoans, who carry Polynesian mtDNA types and Asian HLA markers at very high frequencies. Likewise, although the Tolais of New Britain are Austronesian speakers, they show relatively low genetic affinities to the Taiwanese and to other Austronesians in coastal and island Melanesia. It is likely that the demographic expansion of the Austronesians had a greater genetic impact in some regions than in others, depending on the population size and technological development of the earlier settlers. Our Y-chromosome data suggest a large inflow of male lineages into island and coastal Melanesia, and into Polynesia, and indicate that the New Guinea highlanders expanded from a low number of founders and have remained in comparative isolation from outside migrants. In addition, we propose the occurrence of considerable back migration from Polynesia to island Melanesia and coastal New Guinea in recent times.

This work was carried out with the support of Thor Heyerdahl and the Kon-Tiki Museum, Oslo. The Y-chromosome research was funded by Deutsche Forschungsgemeinschaft grants to L.R. (Ro1040/2-2); W.S. was funded by the Chiang Ching-Kuo Foundation of Taiwan. We thank J. B. Clegg and C. Tyler-Smith for DNA samples, A. Aichler and M. Schulte for collecting Tolai blood samples, P. de Knijff for help in the statistical analyses, and the Roro and Trobriand people for their cooperation in this study.

REFERENCES

Anderson, S. (and 13 others) 1981 Sequence organisation of the human mitochondrial genome. *Nature* **290**, 457–465.

Bayliss-Smith, T. 1996 People–plant interactions in the New Guinea highlands: agricultural hearthland or horticultural backwater? In *The origins and spread of agriculture and pastoralism in Eurasia* (ed. D. R. Harris), pp. 499–523. London: UCL Press.

Bellwood, P. S. 1989 The colonization of the Pacific: some current hypotheses. In *The colonization of the Pacific: a genetic trail* (ed. A. V. S. Hill & S. W. Serjeantson), pp. 1–59. Oxford University Press.

Bellwood, P. 1995 Austronesian prehistory in Southeast Asia: homeland, expansion and transformation. In *The Austronesians* (ed. P. Bellwood, J. J. Fox & D. Tryon), pp. 96–111. Canberra: Australia National University.

Bellwood, P. 1996 The origins and spread of agriculture in the Indo-Pacific region; gradualism and diffusion or revolution and colonization? In *The origins and spread of agriculture and pastoralism in Eurasia* (ed. D. R. Harris), pp. 465–498. London: UCL Press.

Bellwood, P., Fox, J. J. & Tryon, D. (eds) 1995 *The Austronesians*. Canberra: Australia National University.

Cann, R. L., Stoneking, M. & Wilson, A. C. 1987 Mitochondrial DNA and human evolution. *Nature* **325**, 31–36.

Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. 1994 *The history and geography of human genes*. Princeton University Press.

Deka, R., Jin, L., Shriver, M. D., Yu, L. M., Saha, N., Barrantes, R., Chakraborty, R. & Ferrell, R. E. 1996 Dispersion of human Y chromosome haplotypes based on five microsatellites in global populations. *Genome Res.* **6**, 1177–1184.

De Knijff, P. (and 28 others) 1997 Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int. J. Leg. Med.* **110**, 134–149.

Diamond, J. 1988 Express train to Polynesia. *Nature* **336**, 307–308.

Epstein, A. L. 1991 Tolai. In *Encyclopedia of world cultures*, vol. II (ed. T. E. Hays), pp. 333–336. Boston, MA: Oceania, Hall & Co.

Excoffier, L., Smouse, P. E. & Quattro, J. M. 1992 Analysis of molecular variance inferred from metric distances among DNA haplogroups: applications to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.

Felsenstein, J. 1989 PHYLIP. Phylogeny inference package (version 3.2). *Cladistics* **5**, 164–166.

Felsenstein, J. 1995 PHYLIP. Phylogeny inference package (Version 3.57c).

Forster, P., Kayser, M., Meyer, E., Roewer, L., Pfeiffer, H., Benkmann, H. & Brinkmann, B. 1998 Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. *Molec. Biol. Evol.* **15**, 1108–1114.

Gao, X. & Serjeantson, S. W. 1992 Analysis of 2600 HLA-DR, DQ haplotypes in Asia-Oceania. In *HLA 1991*, vol. 2 (ed. K.

Tsuji, M. Aizawa & T. Sazazuki), pp. 229–232. Oxford University Press.
Gao, X., Bhatia, K., Trent, R. J. & Serjeantson, S. W. 1992*a* HLA-DR, DQ nucleotide sequence polymorphism in five Melanesian populations. *Tissue Antigens* **40**, 31–37.
Gao, X., Veale, A. & Serjeantson, S. W. 1992*b* HLA class II diversity in Australian aborigines: unusual HLA-DRB1 alleles. *Immunogenetics* **36**, 333.
Hagelberg, E. 1997 Ancient and modern mitochondrial DNA sequences and the colonization of the Pacific. *Electrophoresis* **18**, 1529–1533.
Hagelberg, E. & Clegg, J. B. 1991 Isolation and characterisation of DNA from archaeological bone. *Proc. R. Soc. Lond.* B **244**, 45–50.
Hagelberg, E. & Clegg, J. B. 1993 Genetic polymorphisms in prehistoric Pacific islanders determined by analysis of ancient bone DNA. *Proc. R. Soc. Lond.* B **252**, 163–170.
Hagelberg, E., Quevedo, S., Turbon, D. & Clegg, J. B. 1994 DNA from prehistoric Easter Islanders. *Nature* **369**, 25–26.
Hagelberg, E., Goldman, N., Lió, P., Whelan, S., Schiefenhövel, W., Clegg, J. B. & Bowden, D. K. 1999 Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proc. R. Soc. Lond.* B **266**. (In the press.)
Hammer, M. F. 1994 A recent insertion of an Alu element on the Y chromosome is a useful marker for human population studies. *Molec. Biol. Evol.* **11**, 749–761.
Hammer, M. F. 1995 A recent common ancestry for human Y chromosomes. *Nature* **378**, 376–378.
Hammer, M. F., Spurdle, A. B., Karafet, T., Bonner, M. R., Wood, E. I., Novelletto, A., Malaspina, N., Mitchell, R. J., Horai, S., Jenkins, T. & Zegura, S. L. 1997 The geographic distribution of human Y chromosome variation. *Genetics* **145**, 787–805.
Hammer, M. F., Karafet, T., Rasanayagam, A., Wood, E. T., Altheide, T. K., Jenkins, T., Griffiths, R. C., Templeton, A. R. & Zegura, S. L. 1998 Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. *Molec. Biol. Evol.* **15**, 427–441.
Hashimoto, M., Kinoshita, T., Yamasaki, M., Tanaka, H., Imanishi, T., Ihara, H., Ichikawa, Y. & Fukunishi, T. 1994 Gene frequencies and haplotypic associations within the HLA region in 916 unrelated Japanese individuals. *Tissue Antigens* **44**, 166–173.
Hertzberg, M., Mickleson, K. P. N., Serjeantson, S. W., Prior, J. F. & Trent, R. J. 1989 An Asian-specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. *Am. J. Hum. Genet.* **44**, 504–510.
Heyer, E., Puymirat, J., Dieltjes, P., Bakker, E. & de Knijff, P. 1997 Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigreess. *Hum. Molec. Genet.* **6**, 799–803.
Hill, A. V. S. & Serjeantson, S. W. (eds) 1989 *The colonization of the Pacific: a genetic trail*. Oxford University Press.
Hurles, M. E., Irvine, C., Taylor, P. G., dos Santos, F. R., Macaulay, V., Loughlin, J., Jobling, M. A. & Sykes, B. C. 1998 European Y chromosomes in the Polynesian gene pool: a contrast to the structure revealed by mitochondria. *Am. J. Hum. Genet.* (In the press.)
Jobling, M. A. & Tyler-Smith, C. 1995 Fathers and sons: the Y chromosome and human evolution. *Trends Genet.* **11**, 449–456.
Kayser, M. (and 28 others) 1997 Evaluation of Y chromosomal STRs: a multicenter study. *Int. J. Leg. Med.* **110**, 125–133.
Liep, J. 1991 Rossel. In *Encyclopedia of world cultures*, vol. II (ed. T. E. Hays), pp. 277–280. Boston, MA: Oceania, Hall & Co.
Lum, J. K., Rickards, O., Ching, C. & Cann, R. L. 1994 Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. *Hum. Biol.* **66**, 567–590.
Melton, T., Peterson, R., Redd, A. J., Saha, N., Sofro, A. S. M., Martinson, J. & Stoneking, M. 1995 Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am. J. Hum. Genet.* **57**, 403–414.
Morwood, M. J., O'Sullivan, P. B., Aziz, F. & Raza, A. 1998 Fission-track ages of stone tools and fossils of the east Indonesian island of Flores. *Nature* **392**, 173–176.
Nagy, M., Zimdahl, H., Krüger, C., Anders, P., Kayser, M., Roewer, L. & Schiefenhövel, W. 1997 Are the Trobrianders emigrants of South-east Asia? In *Genetic diversity of HLA. functional and medical implications* (ed. D. Charron), pp. 185–188. Sevres, France: EDK.
Nei, M. 1987 *Molecular evolutionary genetics*. New York: Columbia University Press.
O'Shaughnessy, D. F., Hill, A. V. S., Bowden, D. K., Weatherall, D. J. & Clegg, J. B. 1990 Globin genes in Micronesia: origins and affinities of Pacific island peoples. *Am. J. Hum. Genet.* **46**, 144–155.
Redd, A. J., Takesaki, N., Sherry, S. T., McGarvey, S. T., Sofro, A. S. M. & Stoneking, M. 1995 Evolutionary history of the COII/tRNALys intergenic 9-bp deletion in human mitochondrial DNAs from the Pacific. *Molec. Biol. Evol.* **12**, 604–615.
Roberts-Thomson, J. M., Martinson, J. J., Norwich, J. T., Harding, R. M., Clegg, J. B. & Boettcher, B. 1996 An ancient common origin of aboriginal Australians and New Guinea highlanders is supported by a-globin haplotype analysis. *Am. J. Hum. Genet.* **58**, 1017–1024.
Roewer, L., Kayser, M., Dieltjes, P., Nagy, M., Bakker, E., Krawczak, M. & de Knijff, P. 1996 Analysis of molecular variance (amova) of Y-chromosome specific microsatellites in two closely related human populations. *Hum. Molec. Genet.* **7**, 1029–1033.
Saitou, N. & Nei, M. 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molec. Biol. Evol.* **4**, 406–425.
Seielstad, M. T., Herbert, J. M., Lin, A. A., Ibrahim, M., Vollrath, D. & Cavalli-Sforza, L. L. 1994 Construction of human Y chromosomal haplotypes using a new polymorphic A to G transition. *Hum. Molec. Genet.* **3**, 2159–2161.
Serjeantson, S. W. 1989 HLA genes and antigenes. In *The colonization of the Pacific: a genetic trail* (ed. A. V. S. Hill & S. W. Serjeantson), pp. 120–173. Oxford University Press.
Serjeantson, S. W. & Gao, X. 1995 *Homo sapiens* is an evolving species: origins of the Austronesians. In *The Austronesians* (ed. P. Bellwood, J. J. Fox & D. Tryon), pp. 165–179. Canberra: Australia National University.
Soodyall, H., Jenkins, T. & Stoneking, M. 1995 'Polynesian' mtDNA in the Malagasy. *Nature Genet.* **10**, 377–378.
Spriggs, M. 1997 *The island Melanesians*. Oxford: Blackwell.
Spurdle, A. B., Woodfield, D. G., Hammer, M. F. & Jenkins, T. 1994 The genetic affinity of polynesians: evidence from Y chromosome polymorphisms. *Ann. Hum. Genet.* **58**, 251–263.
Stoneking, M., Jorde, L. B., Bhatia, K. & Wilson, A. C. 1990 Geographic variation in human mitochondrial DNA from Papua New Guinea. *Genetics* **124**, 717–733.
Sykes, B., Leiboff, A., Low-Beer, J., Tetzner, S. & Richards, M. 1995 The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. *Am. J. Hum. Genet.* **57**, 1463–1475.
Thomas, M. G., Cook, C. E., Miller, K. W. P., Waring, M. J. & Hagelberg, E. 1998 Molecular instability in the *COII-tRNALys* intergenic region of the human mitochondrial genome: multiple origins of the 9-bp deletion and heteroplasmy for expanded repeats. *Phil. Trans. R. Soc. Lond.* B **353**, 955–965.
Tsuji, K., Aizawa, M. & Sasazuki, T. 1992 *HLA 1991*. Oxford University Press.
Underhill, P. A., Jin, L., Zemans, R., Oefner, P. J. & Cavalli-Sforza, L. L. 1996 A pre-Columbian Y chromosome specific transition and its implication for human evolutionary history. *Proc. Natn. Acad. Sci. USA* **93**, 196–200.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Underhill, P. A. (and 8 others) 1997 Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography (DHPLC). *Genome Res.* **7**, 996–1005.
Veitia, R. (and 11 others) 1997 Mutations and sequence variants in the testis-determining region of the Y chromosome in individuals with a 46, XY female phenotype. *Hum. Genet.* **99**, 648–652.
Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. 1991 African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503–1507.
Wallace, D. C. 1995 Mitochondrial DNA variation in human evolution, degenerative disease, and aging. *Am. J. Hum. Genet.* **57**, 201–223.
Whitfield, L. S., Sulston, J. E. & Goodfellow, P. N. 1995 Sequence variation of the human Y chromosome. *Nature* **378**, 379–380.
Wickler, S. & Spriggs, M. 1988 Pleistocene human occupation of the Solomon Islands, Melanesia. *Antiquity* **62**, 703–706.
Wrischnik, L. A., Higuchi, R. G., Stoneking, M., Erlich, H. A., Arnheim, N. & Wilson, A. C. 1987 Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. *Nucl. Acids Res.* **15**, 529–542.
Zerjal, T. (and 16 others) 1997 Genetic relationships of Asians and northern Europeans revealed by Y-chromosomal DNA analysis. *Am. J. Hum. Genet.* **60**, 1174–1183.

Discussion

N. Bradman (*University College London, UK*). Would not the analysis of Y-chromosome data be improved by including biallelic markers in addition to the microsatellites you have selected?

E. Hagelberg. Your question is completely right. However, when the present study was started, no biallelic markers were available for the Pacific region. The only available markers were microsatellites. The first biallelic markers for the Pacific, M4, M5, M16 and M21, were described by Underhill and colleagues in their 1997 paper, and were found to be polymorphic in individuals from Australia, New Guinea and the Solomons.

The analysis of biallelic markers in combination with microsatellites would be extremely useful because it would provide the opportunity to look at different time frames: older events using the biallelic markers, and more recent ones using microsatellites. Ideally, one would define haplogroups using the biallelic markers, and refine and date the groups with the microsatellite data. Unfortunately, this was not possible within the time frame of the present project, although we are planning a future study including biallelic markers on the same population samples described here.

B. Sykes (*University of Oxford, UK*). It is certainly an interesting idea that Polynesia was colonized first via Micronesia. However, if that were the case, how can one explain the presence throughout Polynesia of about 4% of Papuan highlander haplotypes, characterized by the A–C transversion at 16265? It is hard to see how this could have been so widely distributed, and not diluted by other Melanesian haplotypes, if they arose through more recent admixture.

E. Hagelberg. On the basis of the mitochondrial DNA data alone, it is impossible to determine the route of the settlement of the eastern Pacific. The very high frequency of the Polynesian motif in parts of the New Guinea coast, parts of island Melanesia, and in Polynesia, indicates that the people living in these areas have already gone through a genetic bottleneck which has drastically reduced the number of mtDNA lineages. It is generally assumed that the bottleneck would have occurred somewhere in remote Oceania, where the sea distances are great, rather than in the coastal or island regions of Melanesia. Common sense would argue against a genetic bottleneck in island Melanesia, unless some type of selective pressure is invoked. Where do the Polynesian genes come from? Mark Stoneking suggests that the Polynesian motif arose in east Indonesia, but why would it amplify dramatically in coastal New Guinea? My observations would suggest that the Polynesian motif arrived recently in coastal and island Melanesia, probably as a result of recent back migrations of Polynesians. The occurrence of back migrations is supported by ethnographic and linguistic evidence. High levels of recent back migrations would probably erase much of the genetic signature of the early migrants, so my feeling is that the question of the route of Polynesian colonization still remains undecided if we look at the genetic data alone. At this point we cannot rule out Micronesia simply on genetic grounds.

Regarding the occurrence of the A–C transversion at position 16265 in a small number of Polynesians, my study of mtDNA variation in the Melanesian archipelago of Vanuatu indicates that this particular marker is very abundant in island Melanesia. It might have been carried into the eastern Pacific by the original settlers, but there is nothing in the available mtDNA data that would rule out a recent (post-European) introduction of this mtDNA type into Polynesia. As you showed in your 1995 survey, a number of other non-Polynesian mtDNA types was present in the populations samples you examined, including mtDNA types previously observed in Britain, Turkey, Portugal, Germany, the Basque country, and South America. Although there have been a number of different mtDNA studies on Pacific populations, the populations samples are still small. Your figure of 4% of Papuan highlander haplotypes in Polynesia is an average, skewed by their high frequency of the haplotypes in the Cook Islands. In fact, the haplotype is only present in a handful of people in other archipelagos, together with a small proportion of other non-Polynesian haplotypes. There is no way of telling whether the non-Polynesian mtDNA types were pre- or post-European introductions. We must try to avoid selecting the data to fit a particular model.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES