Human Biology

Volume 85
Issue 1 *Special Issue on Revisiting the "Negrito" Hypothesis*

Article 9

2013

Genetic Diversity of Four Filipino Negrito Populations from Luzon: Comparison of Male and Female Effective Population Sizes and Differential Integration of Immigrants into Aeta and Agta Communities

E Heyer
EcoAnthropologie et Ethnobiologie, MNHN, CNRS, Université Paris Diderot, Paris, France, heyer@mnhn.fr

M Georges
EcoAnthropologie et Ethnobiologie, MNHN, CNRS, Université Paris Diderot, Paris, France

M Pachner
Andaman and Nicobar Association, Liestal, Switzerland

P Endicott
EcoAnthropologie et Ethnobiologie, MNHN, CNRS, Université Paris Diderot, Paris, France

Genetic Diversity of Four Filipino Negrito Populations from Luzon: Comparison of Male and Female Effective Population Sizes and Differential Integration of Immigrants into Aeta and Agta Communities

Abstract
Genetic data corresponding to four negrito populations (two Aeta and two Agta; $n = 120$) from the Luzon region of the Philippines have been analyzed. These data comprise mitochondrial DNA (mtDNA) hypervariable segment 1 haplotypes and haplogroups, Y-chromosome haplogroups and short tandem repeats (STRs), autosomal STRs, and X-chromosome STRs. The genetic diversity and structure of the populations were investigated at a local, regional, and interregional level. We found a high level of autosomal differentiation, combined with no significant reduction in diversity, consistent with long-term settlement of the Luzon region by the ancestors of the Agta and Aeta followed by reduced gene flow between these two ethnolinguistic groups. Collectively, the Aeta have a much higher ratio of female:male effective population size than do the Agta, a finding that supports phylogenetic analysis of their mtDNA and Y-chromosome haplogroups, which suggests different genetic sex-biased contributions from putative Austronesian source populations. We propose that factors of social organization that led to the reduction in Agta female effective population size may also be linked to the limited incorporation of female lineages associated with the settlement of the Philippines by Austronesian speakers; conversely, the reduction in Aeta male effective population size, relative to females, could be indicative of a limited incorporation of male lineages associated with this demographic process.

Keywords
Negritos, Sex-Specific Behaviors, Genetic Diversity

Genetic Diversity of Four Filipino Negrito Populations from Luzon: Comparison of Male and Female Effective Population Sizes and Differential Integration of Immigrants into Aeta and Agta Communities

E. HEYER,[1]* M. GEORGES,[1] M. PACHNER,[2] AND P. ENDICOTT[1]

Abstract Genetic data corresponding to four negrito populations (two Aeta and two Agta; $n = 120$) from the Luzon region of the Philippines have been analyzed. These data comprise mitochondrial DNA (mtDNA) hypervariable segment 1 haplotypes and haplogroups, Y-chromosome haplogroups and short tandem repeats (STRs), autosomal STRs, and X-chromosome STRs. The genetic diversity and structure of the populations were investigated at a local, regional, and interregional level. We found a high level of autosomal differentiation, combined with no significant reduction in diversity, consistent with long-term settlement of the Luzon region by the ancestors of the Agta and Aeta followed by reduced gene flow between these two ethnolinguistic groups. Collectively, the Aeta have a much higher ratio of female:male effective population size than do the Agta, a finding that supports phylogenetic analysis of their mtDNA and Y-chromosome haplogroups, which suggests different genetic sex-biased contributions from putative Austronesian source populations. We propose that factors of social organization that led to the reduction in Agta female effective population size may also be linked to the limited incorporation of female lineages associated with the settlement of the Philippines by Austronesian speakers; conversely, the reduction in Aeta male effective population size, relative to females, could be indicative of a limited incorporation of male lineages associated with this demographic process.

The Philippines is remarkable for the high level of diversity seen in physical traits and languages among its indigenous inhabitants; for example, currently over 170 ethnolinguistic groups are recognized (Gordon and Grimes 2005). Of these groups, 20 have been historically defined as "negrito," based on short stature, frizzy hair,

Raw data are available upon request to heyer@mnhn.fr.

[1]UMR7206, EcoAnthropologie et Ethnobiologie, MNHN, CNRS, Université Paris Diderot, Paris, France.
[2]Andaman and Nicobar Association, Liestal, Switzerland.
*Correspondence to: E. Heyer, UMR7206, EcoAnthropologie and Ethnobiologie, MNHN, CNRS, Université Paris Diderot, Paris, France. E-mail: heyer@mnhn.fr

Human Biology, February–June 2013, v. 85, no. 1–3, pp. 189–208.

KEY WORDS: NEGRITOS, SEX-SPECIFIC BEHAVIORS, GENETIC DIVERSITY.

and dark skin color (Bean 1910). Although it is now generally accepted that human settlement of Island Southeast Asia commenced during the late Pleistocene, the relative contribution of this early migration to the present-day Philippine gene pool is still unresolved. A more recent period of immigration occurred during the middle to late Holocene and was associated with the spread of the Malayo-Polynesian branch of the Austronesian language family, from Taiwan, via the Philippines, to Near and Remote Oceania (Gray et al. 2009). The Austronesian speakers are thought to have arrived at Luzon in the north of the Philippines and to have developed a system of settled agriculture. All negrito populations speak languages belonging to the Malayo-Polynesian branch of the Austronesian family, which are believed to have replaced non-Austronesian languages previously spoken in the middle to late Holocene (Reid 1994; Gray et al. 2009; Reid this issue).

A recent study showed that many of the mitochondrial DNA (mtDNA) lineages of the Philippines are shared with Taiwanese aboriginal groups (Tabbada et al. 2010). Other mtDNA lineages in the Philippines today appear to have their origins within the region of Southeast Asia and neighboring New Guinea (Tabbada et al. 2010; Gunnarsdottir et al. 2011). An extensive study of Y chromosomal DNA in the Philippines showed that all negrito populations have a substantial component of unresolved K*-M9 lineages, while a few have elements attributable to C*-RPS4Y, both of which are presumed to predate the Austronesian expansion; the percentages of these lineages was found to be particularly high among Aeta (87%, 100%) and Agta (47%) populations of northern Luzon (Delfin et al. 2011). The distribution of K*-M9 is not limited solely to groups defined as negrito, because it occurs at a frequency of 67% among the Hanunuo of southern Luzon (Delfin et al. 2011). Overall, these results suggest that both the maternal and paternal gene pools of the Philippines retain pre-Austronesian haplogroups but that the level of Austronesian genetic input is higher for mtDNA than for the Y chromosome in Luzon.

The same pattern can be seen in Oceania, where the Austronesian colonization was strongly sex biased: 66% of Polynesian Y chromosomes are of Near Oceanian origin, whereas 94% of Polynesian mtDNAs are of East Asian origin (Kayser et al. 2006; Kayser 2010). Although single-locus studies may not accurately reflect the actual demographic history, a comparison of X-chromosomal and autosomal single-nucleotide polymorphism (SNP) data from Indonesia (Cox et al. 2010) reached the same conclusion of a strong sex bias to patterns of admixture, such that Asian women made a higher contribution to the resulting gene pool. It was proposed that this sex-biased genetic admixture involving Austronesian women and Near Oceanic men was the result of the matrilineal structure and matrilocal residence pattern of Austronesian societies in the past (Jordan et al. 2009; Lansing et al. 2011). An exception to this global trend can be found in an Aeta group from the Luzon region of the Philippines, where X-chromosome admixture levels were significantly lower than those for the autosomes; in contrast, the non-negrito populations studied had balanced amounts of admixture (Cox et al. 2010).

In order to properly assess sex-specific behaviors, one needs to disentangle the respective influence of changes in effective population sizes and migratory forces

(Heyer et al. 2012). This can be achieved by contrasting measures of diversity obtained from uniparentally inherited markers (mtDNA and Y chromosome) and autosomal and X-chromosome polymorphisms. Therefore, in order to better understand sex-specific behaviors among the negritos of Luzon, we conducted a study combining autosomal short tandem repeats (STRs), X-STRs, and Y-STRs, together with mtDNA hypervariable segment 1 (HVS1) and Y-chromosomal SNP data used for inferring haplogroups based on the nonrecombining region of the Y chromosome (NRY). We chose two populations belonging to the Aeta and two from the Agta ethnolinguistic groups of Luzon in order to compare intra- and intergroup variation within this northern region of the Philippines where Austronesian settlement likely commenced.

Materials and Methods

DNA Sampling. We sampled 120 unrelated individuals from four populations currently residing in the Luzon region in the Philippines. Two populations are from the western part of the island, Aeta from Aglao (coded as NEGa) and Aeta from Sta. Juliana (NEGj), and two are from the eastern part, Agta from Umiray (NEGu) and Agta from Cozo (NEGc) (Figure 1).

Saliva samples from 120 from healthy adult donors were collected with informed consent in the presence of a local representative of the National Commission on Indigenous Peoples using the Oragene® DNA self-collection kit from DNA Genotek Inc. (Ontario, Canada) and extracted following manufacturer's instructions. Samples were then stored in 1× Tris-HCl (10 mM) EDTA (1 mM) buffer at –20°C. All samples were anonymized.

DNA Analysis

Autosomal, X-, and Y STRs. For their probable selective neutrality and independence, we used tetranucleotide STRs (28 autosomal STRs and 8 X-STRs) that were previously reported (Verdu et al. 2009) and chosen from the Marshfield Foundation Mammalian Genotyping Service Screening Set 10 (http://research.marshfieldclinic.org/genetics/GeneticResearch/screeningsets.asp).

For 66 male individuals, we genotyped 12 STRs (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS393, DYS439, DYS385a, DYS385b, DYS426, DYS388, DYS392) on the nonrecombining region of the Y chromosome (NRY), chosen from the Y Haplotype Reference Database (www.yhrd.org). For comparison with Delfin et al. (2011), we retained only seven NRY STRs (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393).

Polymerase chain reaction (PCR) amplifications were performed in a final volume of 20µl containing 1× buffer, 0.25 mM of each deoxyribonucleotide triphosphate (dNTP), 0.06 U Taq DNA polymerase (Bioline, MA, USA), 0.125µM primers (forward primers fluorescently labeled with 6-FAM, VIC, NED, or PET fluorochromes; Applied Biosystems, CA, USA) and 20 ng DNA, under the

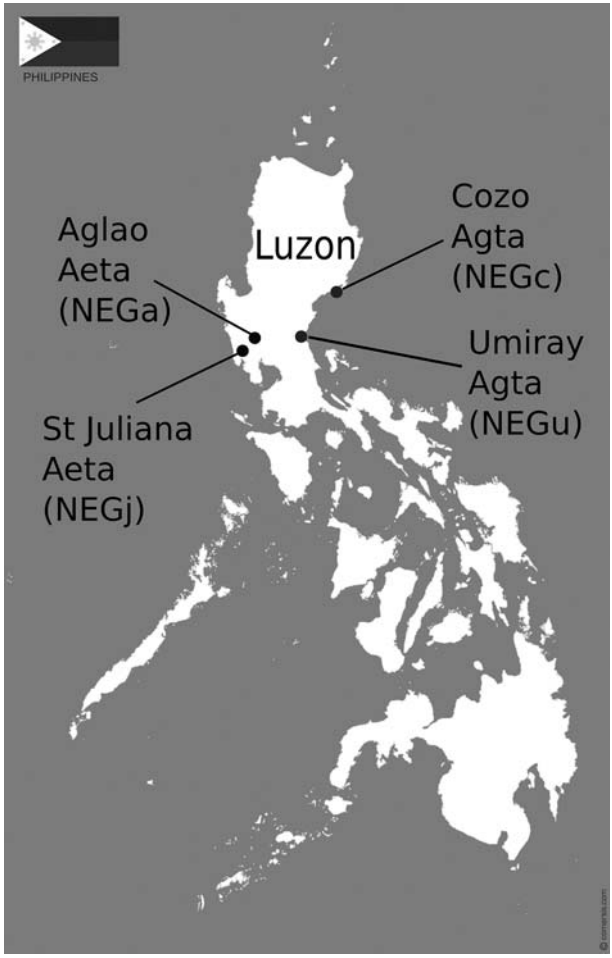


Figure 1. Map of North Philippines.

following conditions for all markers: initial denaturation at 94°C for 10 min; followed by 35 cycles of 94°C for 30 s, annealing at 55°C for 75 s, and extension at 72°C for 20 s; and a final extension step at 72°C for 10 min. PCR products were analyzed in an Applied Biosystems 3100 automated sequencer. Alleles were scored with the software GeneMarker version 1.6 (SoftGenetics LLC, PA, USA), and each allele call was manually verified. All ambiguous allelic states, as well as all alleles called only once in a single individual (singleton), were reprocessed and verified independently several times.

NRY Haplogroups. High-resolution melt curve analysis was carried out on a CFX96 Real-Time PCR Detection System (Bio-Rad, CA, USA) using 5 ng DNA in a 10μl reaction volume containing 2× SsoFast Evagreen Supermix (Bio-Rad) and

300 nM each forward and reverse primers for hierarchical binary polymorphisms M9, M38, M122, M175, M216, and M45 using previously published primers (Brion et al. 2005). This strategy was applied to those individuals not clearly assigned by comparison of Y-STR haplotypes with the reference data set, together with a subset of the remainder as controls. Binary polymorphisms typed are equivalent to those used by Delfin et al. (2011) according to the International Society of Genetic Genealogy (www.isogg.org/). Allelic differences were identified using the Bio-Rad Precision Melt Analysis software.

mtDNA HVS1. The mtDNA first hypervariable segment of the mtDNA control region (HVS1) was amplified using primers L15987 (5′-TCAAATGGGCCTGTCCTTGTA-3′) and H580 (5′-TTGAGGAGGTA AGCTACATA-3′). PCR amplifications were performed in a final volume of 20μL) containing 1× buffer, 0.25 mM of each dNTP, 0.05 U Taq DNA polymerase (Bioline Institut), 0.125μM primers and 20 ng DNA, under the following conditions: initial denaturation at 94°C for 5 min, followed by 39 cycles of 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 30 s; and a final extension step at 72°C for 10 min. The amplification products were subsequently purified with EXO-SAP (Ozyme, France) standard procedure following manufacturer's instructions. DNA sequencing was performed using primers L15925 (5′-TAATACACCAGTCTTGTAAAC-3′) and HH23 (5′-AATAGGGTGATAGACCTGTG-3′). Sequences from nucleotide positions 16,024–16,391 were considered, and the poly(C) tract was excluded from analysis. Multiple sequence alignment was performed on the sequences (Aeta vs. Agta; Aeta, Agta, and other populations) prior to their use in the estimation of summary statistics. The mtDNA haplotypes from HVS1 data were assigned to haplogroups using the nomenclature from Phylotree.org (version Build 14).

Statistical Analysis. Arlequin software version 3.0 (Excoffier et al. 2005) was used to calculate genetic diversity indices: average gene diversity (H), average number of alleles, Garza-Williamson index (Garza and Williamson 2001) for autosomal STRs, haplotype diversity for NRY STRs, average gene diversity, mean pairwise differences (P_i), number of mutations (S), number of haplotypes (k), Tajima's D tests, and Fu's test (F_S) for mtDNA HVS1. SPAGeDi version 1.2 (Hardy and Vekemans 2002) was used to calculate X-STR average gene diversity and variance of allele size.

Regarding population genetic distances, we computed the pairwise levels of genetic differentiation (F_{ST}; Weir and Cockerham 1984) for all pairs of populations and all four genetic systems separately, using Arlequin 3.0 for autosomal, Y-chromosome, and mitochondrial HVS1 data, and SPAGeDi 1.2 for the haplodiploid X-chromosome markers. Ten thousand permutations of individuals among populations were performed to obtain p-values. Negative and nonsignificant values of F_{ST} were set to zero. For population comparisons, F_{ST} values were computed for populations chosen from previously published works, except for X-STRs where

no data were available. The three matrices of pairwise F_{ST} values thus obtained (autosomal STRs, NRY STRs, mtDNA HVS1) were used to perform separately three two-dimensional metric multidimensional scaling (MDS) analyses using the software package R (R Development Core Team 2012). To evaluate the accuracy of the two-dimensional MDS projection of the original F_{ST} matrices, we computed the Spearman correlation coefficient ρ between Euclidian distances between each pair of population on the MDS plot and their genetic distances.

Results

Genetic Diversity. The average genetic diversity (H) across the 28 autosomal loci ranged from 0.65 to 0.67, with values very similar for all four populations (Table 1). Comparing to data for 52 worldwide populations, using the same 28 STRs, the values fall midway between those for Yoruba of Africa (0.746) and Amerindians (0.47), and are very close to the mean genetic diversity worldwide of 0.693 ± 0.082 (Rosenberg et al. 2002). In a regional context, this is within the range of values found in other Southeast Asian groups, intermediate between Melanesia (0.61) and Papua New Guinea (0.69). The average number of alleles is lower than in Melanesia (6.5) but higher than in Papua New Guinea (4.36). These results highlight the fact that these four negrito populations do not show a reduction of intrapopulation genetic diversity: their genetic diversity is in the range of other human populations.

The Garza-Williamson index measures whether a population has experienced a genetic bottleneck. Values for the four negrito populations are in the same range as those for other populations worldwide (Table S1). Since the Garza-Williamson index is supposed to be much smaller in populations having experienced a genetic bottleneck, we can exclude any severe reduction in effective population size for these four negrito populations of the Philippines.

The average haplotype diversity based on the 12 Y-STRs is high, ranging from 0.83 in the Aeta population NEGa to 0.96 in the Agta population NEGu. Restricting the data to the seven STRs used for comparison with previous studies, we found values ranging from 0.89 to 0.94. The lowest value was found in the Aeta from Aglao NEGa, and the highest, in the Agta from Umiray (NEGu). These values are within the range of NRY diversity (0.59–0.98) previously reported in negrito and non-negrito populations (Delfin et al. 2011).

While the mitochondrial sequence diversity values are in general fairly high (Table 2), the NEGu stand out as having a lower haplotypic diversity, a reduced number of haplotypes ($k = 6$), but relatively high mean pairwise differences (P_i). Haplotype diversity is lower than that previously reported for non-negrito populations (0.97) from the Philippines (Tabbada et al. 2010). All four groups show nonsignificant Tajima's D and Fu's F_S, consistent with stationary population sizes.

Population Differentiation. Regarding autosomal STRs, the average F_{ST}

Table 1. Genetic Diversity Indices Based on STRs

GENETIC SYSTEM	AETA		AGTA	
	AGLAO (NEGa)	STA. JULIANA (NEGj)	COZO (NEGc)	UMIRAY (NEGu)
Autosomal STRs (28 STRs)				
Sample size	30	30	30	30
Average gene diversity (*H*)	0.65	0.66	0.66	0.67
Standard deviation	0.33	0.33	0.33	0.33
Average number of alleles	4.68	4.96	4.86	5.11
X-STRs (8 STRs)				
Sample size	30	30	30	30
Average gene diversity	0.66	0.70	0.63	0.58
Standard deviation	0.09	0.09	0.09	0.14
Variance of allele size	19.68	24.89	18.21	21.20
NRY STRs				
Sample size	13	20	17	15
Twelve STRs				
Haplotype diversity	0.83	0.91	0.92	0.96
Standard deviation	0.08	0.05	0.04	0.04
Seven STRs				
Haplotype diversity	0.90	0.92	0.94	0.92
Standard deviation	0.05	0.05	0.04	0.05

Indices are categorized by population, location, and code. Average gene diversity was computed over all loci.

Table 2. Mitochondrial DNA Diversity Indices

STATISTIC	AETA		AGTA	
	AGLAO (NEGa)	STA. JULIANA (NEGj)	COZO (NEGc)	UMIRAY (NEGu)
Sample size	29	28	30	29
Average gene diversity	0.87	0.87	0.82	0.67
Standard deviation	0.04	0.04	0.04	0.06
p_i	5.37	4.44	4.58	4.20
S	19	23	19	16
k	9	11	9	6
D	0.39	–0.89	–0.16	0.10
F_S	1.35	–0.92	0.83	3.01

Indices are categorized by population, location, and code. Abbreviations: P_i, mean number of pairwise differences; *S*, number of polymorphic sites; *k*, number of haplotype; *D*, Tajima's *D* (*p*-value > 0.05); F_S, Fu's F_S (*p*-value > 0.05).

among the four populations is 0.059 (Table 3). They cluster in two groups such that the Aeta NEGa and NEGj group together, as do the Agta NEGc and NEGu. When the genetic distances of the Aeta and Agta populations are compared with a worldwide sample from the Centre d'Etude du Polymorphisme Humain (CEPH)

Table S1. Genetic Diversity Indices of CEPH Populations (Rosenberg et al. 2002) Based on the 28 Autosomal STRs Used in the Present Study

REGION/POPULATION CEPH CODE	AVERAGE GENE DIVERSITY (H)	STANDARD DEVIATION	AVERAGE NUMBER OF ALLELES	GARZA–WILLIAMSON INDEX
Americas				
Surui-Brazil	0.47	0.20	3.04	0.28
Colom	0.65	0.12	4.14	0.29
Maya	0.69	0.08	4.93	0.29
Pima	0.61	0.14	4.21	0.28
East Asia				
Cambo	0.70	0.12	4.64	0.28
Dai	0.70	0.09	4.50	0.27
Daur	0.69	0.08	4.54	0.28
Han-Nchina	0.71	0.09	4.64	0.28
Han	0.71	0.06	5.86	0.27
Hezhe-China	0.69	0.09	4.36	0.29
Lahu	0.67	0.13	4.46	0.27
Miao	0.67	0.15	4.50	0.27
Mongo-China	0.71	0.11	4.68	0.27
Naxi	0.67	0.10	4.50	0.28
Oroqe-China	0.73	0.09	4.68	0.28
She	0.69	0.10	4.54	0.26
Tu	0.71	0.12	4.61	0.28
Tujia-China	0.70	0.12	4.79	0.27
Xibo	0.71	0.07	4.50	0.27
Yi	0.70	0.08	4.64	0.27
Japan	0.71	0.08	5.82	0.28
Yakut-Siber	0.69	0.09	5.29	0.28
Central-South Asia				
Uygur-China	0.72	0.07	4.79	0.27
Baloch	0.71	0.06	5.50	0.27
Brahui	0.71	0.06	5.61	0.27
Burusho	0.71	0.06	5.43	0.27
Hazar	0.71	0.07	5.39	0.28
Kalash	0.70	0.07	4.82	0.28
Makrani	0.72	0.06	5.64	0.27
Pathan	0.72	0.06	5.71	0.27
Sindhi	0.72	0.06	5.61	0.26

populations (excluding the Americas) based on 28 autosomal STRs (Figure 2a), the negrito populations are situated on their own in the MDS plot. The two groups (Aeta and Agta) also show a high level of genetic differentiation between themselves; the genetic distance is as high as that between each group and mainland Asiatic populations.

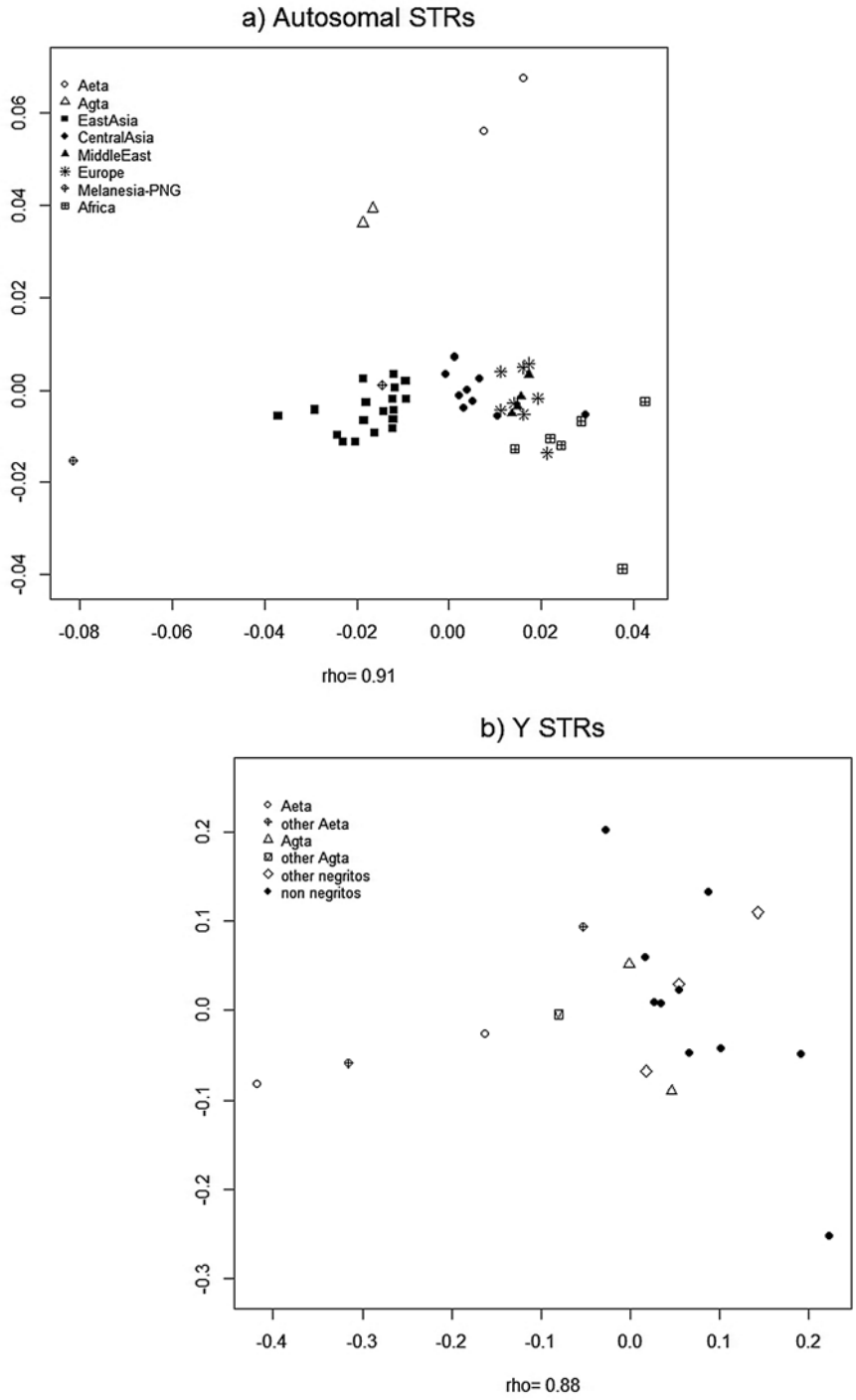
The global Y-chromosome F_{ST} is 0.250 (the global Y-chromosome estimated genetic distance (R_{ST}) is 0.292). Some Aeta populations clearly differentiate from other Philippine groups: they are far away from the cluster that includes most of the non-negrito populations (Figure 2b). Conversely, Agta populations cluster within

REGION/POPULATION CEPH CODE	AVERAGE GENE DIVERSITY (H)	STANDARD DEVIATION	AVERAGE NUMBER OF ALLELES	GARZA–WILLIAMSON INDEX
Europe				
Basque	0.71	0.06	5.46	0.27
French	0.70	0.05	5.32	0.27
Italy	0.70	0.05	4.86	0.27
Sardinia	0.70	0.04	5.43	0.26
Tuscan	0.71	0.06	4.14	0.28
Orcade	0.69	0.07	4.93	0.28
Russia	0.72	0.04	5.43	0.27
Adyge	0.70	0.05	5.07	0.28
Middle East				
Druze	0.72	0.06	6.18	0.27
Palestinian	0.71	0.05	6.29	0.28
Bedouin	0.72	0.05	6.25	0.28
Mozabite	0.72	0.07	6.50	0.28
Oceania				
Melanesia Bouganvilliers	0.61	0.13	4.36	0.29
Papuan New Guinea	0.69	0.07	4.82	0.28
Africa				
Biaka	0.73	0.06	6.50	0.28
Mbuti	0.72	0.06	5.79	0.28
Bantu	0.74	0.05	5.14	0.27
San	0.71	0.12	4.36	0.26
Yoruba	0.75	0.06	6.46	0.29
Mandenka	0.74	0.05	6.54	0.28
Philippines				
NEGa	0.66	0.11	4.68	0.28
NEGj	0.67	0.08	4.96	0.28
NEGc	0.66	0.08	4.86	0.28
NEGu	0.67	0.08	5.11	0.27

CEPH, Centre d'Etude du Polymorphisme Humain.

the Philippine non-negrito Y-chromosome diversity. This again shows the high level of genetic differentiation among the Aeta and Agta. Regarding mtDNA, the two Aeta populations are closer (F_{ST} between NEGa and NEGj = 0.065); the two Agta populations also cluster together, with the F_{ST} between them equal to 0.099. Overall, the negrito populations are more scattered than non-negritos, confirming the greater genetic differentiation among them (Figure 2c; Table 3).

Comparing F_{ST} X and F_{ST} Autosomes: Sex-Specific Effective Population Size. In the absence of any sex-specific behavior, the X-chromosome effective population size should be three-fourths of the autosomes. The F_{ST} between two populations depends on the inverse of the product of the effective population size multiplied by


a) Autosomal STRs
Aeta
Agta
EastAsia
CentralAsia
MiddleEast
Europe
Melanesia-PNG
Africa
0.06
0.04
0.02
0.00
-0.02
-0.04
-0.08
-0.06
-0.04
-0.02
0.00
0.02
0.04
rho= 0.91
b) Y STRs
Aeta
other Aeta
Agta
other Agta
other negritos
non negritos
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.3
-0.2
-0.1
0.0
0.1
0.2
rho= 0.88

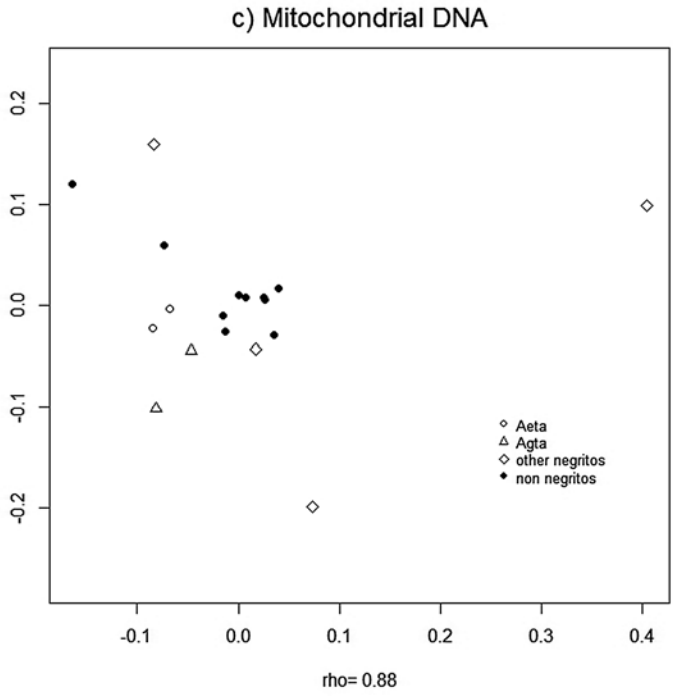


Figure 2 (*opposite and above*). MDS plots. (a) Autosomal genetic distances based on 28 STRs for the negrito populations under study and a worldwide sample of CEPH populations (except Americas) listed in Table S1. ρ = 0.91. PNG, Papua New Guinea. (b) Y-chromosome genetic distance based on seven STRs for the negrito populations under study and populations from the Philippines included in Delfin et al. (2011). ρ = 0.88. (c) mtDNA genetic distance based on HVS1 sequences for the negrito populations under study and populations from the regions listed in Table 3.

the migration rate. The smaller the effective population size, the more drift in each population, and therefore, differences between the populations increase. The higher the effective population size, the less drift occurs in each population, and the less divergent they are. Since the copy number of the X chromosome for females is twice that for males, the ratio with the autosomes is influenced more by the female than by the male effective population size N_e^m. Therefore, in the absence of any sex-specific behavior, we should observe a higher X-chromosome F_{ST} (F_{ST}^X) between populations than autosomal F_{ST} (F_{ST}^A). When the opposite occurs, then the female effective population size (N_e^f) is larger than that of the male.

There is no direct estimator available for these sex-specific effective population sizes, but F_{ST} is a product of the effective population size and the sex-specific migration rate. It has been previously shown that when $F_{ST}^A > F_{ST}^{X-}$, the effective

Table 3. Populations Used for Mitochondrial MDS

REFERENCE	POPULATION CODE[a]	*POPULATION*
Gunnarsdottir et al. 2011	MAM	Negritos from Mamanwa
	MAN	Non-negritos from Manobo
	SUR	Surigaonon non-negritos
Hill et al. 2007	BAN	Banjamarsin from Borneo
	KK	Kota Kianabulu from Borneo
	FIL	Filipinos from Philippin
	AMI (TAIW)	Ami from Taiwan
	ATA (TAIW)	Atayal from Taiwan
	BUN (TAIW)	Bunum from Taiwan
	PAI (TAIW)	Paiwan from Taiwan
Loo et al. 2011	BD	Ivatan from Batanes Archipelago
	YF	Yami (Tao) from Orchid Island
Scholes et al. 2011	BTK	Batak negrito from Palawan
	TB (noBTK)	Tagbanua non-negrito from Palawan
	CAG (noBTK)	Cagayanin non-negrito
	CUY (noBTK)	Cuyonin non-negrito
Tabbada et al. 2010	PHIL	Urban population from Luzon, Cebu city in Visayas, Zamboanga in Mindanao
Thangaraj et al. 2003	GA	Negritos from Great Andaman
	ONJE	Negritos from Onge

[a]In parentheses are our recodings when we grouped some populations.

Table 4. Genetic Differentiation (F_{ST})

GENETIC SYSTEM	AMONG THE TWO AETA AETA POPULATIONS	AMONG THE TWO AGTA POPULATIONS	ALL
Autosomal STRs (28 STRs)	0.021	0.019	0.059
X-STRs (8 STRs)	0.006	0.062	0.037
NRY STRs	0.221	0.118	0.250
mtDNA	0.064	0.099	
F_{ST}^{X}:F_{ST}^{A} ratio	0.26	3.31	0.64

number of females is significantly higher than that of males, whatever the pattern of sex-specific dispersal (Segurel et al. 2008). Verdu et al. (2013) formally demonstrate that for F_{ST}^{X}:F_{ST}^{A} ratio < 1, male effective population size is smaller than that for females, regardless of any sex-specific migration (Verdu et al. 2013). Conversely, when F_{ST}^{X}:$F_{ST}^{A} > 2$, male effective population size is bigger than female effective population size.

Collectively, among the data for the four Philippine populations reported here, F_{ST}^{X} is smaller than F_{ST}^{A} (F_{ST}^{X}:$F_{ST}^{A} = 0.64$; Table 4). Thus, at the global level, F_{ST}^{X}:$F_{ST}^{A} < 1$, indicating that, overall for these four negrito populations, male effective population size is smaller than that of females. When considering the pairs of linguistically related populations separately, however, F_{ST}^{X}:$F_{ST}^{A} = 0.26$ between the

two Aeta populations but F_{ST}^{X}:F_{ST}^{A} = 3.31 between the two Agta groups. The high Agta ratio indicates a higher N_e^m relative to N_e^f. This demonstration of sex-biased diversity, where the male effective population size is higher than that for females, is quite rare and has only been previously reported for the Pygmies of Central Africa (Verdu et al. 2013).

Y-Chromosome Haplogroups. Table 5 reports the main grouping of the Aeta and Agta Y-chromosome haplogroups identified by a combination of Y-STR haplotypes and the high-resolution melt curve SNP analysis. The most common position in the phylogeny is that of K*-M9. Our study has not yet excluded the positions defined by various other markers within this part of the current phylogeny, but as these account for less than 2% in other Philippines populations (Delfin et al. 2011), here we approximate 51% of the Y chromosomes typed to K*-M9 (xM175xM45). This compares with 33% in other Philippine negrito populations and just 11% in the general population. Of the other group of Y chromosomes proposed to be indigenous and preceding the Austronesian expansion, C*-M216 (xM9) here displays frequencies similar to those previously found in both negrito and non-negrito populations (Delfin et al. 2011).

The O-M175 lineages, considered to be likely candidates for introduction by Austronesian speakers, are present in both Agta populations but only one Aeta population (NEGj). The total of O-M175 (xM122) and O-M122 lineages combined, at approximately 31%, is substantially less than that observed in other negrito groups of the Philippines, which average nearly twice this figure and constitute more than 80% of non-negrito lineages (Delfin et al. 2011). The distribution across

Table 5. Phylogenetic Position of Y Chromosomes from Males in the Four Philippine Negrito Populations of the Present Study, Inferred from SNP and STR Data

	AETA		AGTA		PERCENTAGE		
					PRESENT STUDY:	DELFIN ET AL. 2011	
PHYLOGENETIC POSITION OF Y CHROMOSOMES	NEGa	NEGj	NEGc	NEGu	NEGRITO	NEGRITO	NON-NEGRITO
C-M216 (xM9)	0	5	0	2	10.77	8.90	7.10
K-M9 (xM175xM45)	**12**	**9**	**5**	**7**	**50.77**	**32.80**	**11.00**
O-M175 (xM122)	0	2	8	2	18.46	37.3	69.90
O-M122	0	3	4	2	13.85	21.10	11.40
P-M45	0	1	0	2	4.62	0.00	0.50
Other	1	0	0	0	1.54	0.00	0.00
Total individuals	13	20	17	15	64	180	210

Data are expressed as numbers per populations or as total percentages. Data from Delfin et al. (2011) are provided for comparison. The only category shared across all four populations in the present study (K*-M9xM175xM45) is shown in boldface. All O-M175 lineages from Delfin et al. (2011) that were not O-M122 have been collapsed into the O-M175 (xM122) category. Likewise, lineages identified by marker O-M124 have been collapsed into the K-M9xM175xM45 category. The latter constitute only 0.6% of individuals among negritos and 3% of non-negritos in Delfin et al. (2011).

language groups in the present study is pronounced, such that the Agta groups have approximately 80% of the O-M175 lineages among the individuals reported here.

mtDNA Haplogroups. The mtDNA haplogroup results are presented in Table 6. The Agta and Aeta populations have only nine haplogroups among them, and their distribution is quite stochastic, with some present in one population only (E1b1, F1a4, R9*). The two most common mtDNA haplogroups, B4b1 and P9, are present in all four negrito populations and constitute 45% of all lineages found. P9 (previously called P8) has been found at very low frequency among the general population of Luzon and Visaya provinces, and together with P10, these have been proposed as autochthonous lineages (Tabbada et al. 2010). B4b1 overall represents more than 20% of the negrito maternal lineages, compared with approximately 7% of the general population of the Philippines.

Those mtDNA lineages proposed to be associated with the Out-of-Taiwan mtDNA Austronesian expansion (Tabbada et al. 2010) found among the Aeta and Agta are B4a1a, D4, M7b3, M7c3c, and Y2. These constitute approximately 39% of the Aeta mtDNA lineages and approximately 17% of the Agta's. This difference, however, is mostly due to the presence of B4a1a among the Aeta, the frequency of which is almost double that found in the general populations of Luzon (Tabbada et al. 2010). Of the others, M7c3c is the only proposed Out-of-Taiwan lineage found within both Aeta and Agta groups studied here, and its overall frequency is very similar to other negrito and non-negrito populations of the Philippines. E1a1a, which constitutes 11% of the Philippine general population and is another possible mtDNA lineage associated with the Austronesian expansion (Tabbada et al. 2010), is notable by its absence from all four negrito groups in the present study.

Discussion

High Levels of Autosomal Genetic Diversity. The MDS plot (Figure 2a) demonstrates the high level of genetic differentiation among the four negrito populations, such that there is more genetic distance between them than any two Eurasian populations. When compared with Pygmy populations from western Central Africa, the Pygmies show a much smaller differentiation (F_{ST} = 0.019; Verdu et al. 2009) than that found when comparing an Aeta with an Agta population (F_{ST} > 0.07). Two nonexclusive mechanisms could lead to such high autosomal genetic differentiation among Aeta and Agta populations: (1) a low effective population size combined with reduced gene flow, which can quickly generate a high level of genetic differentiation; and (2) a long-term settlement in the region combined with a subsequent reduction in gene flow between the populations.

However, the first hypothesis can be dispensed with because the genetic diversity estimates show no reduction in diversity compared with populations worldwide. These results are consistent with Migliano et al. (this issue), who find that Aeta, Batak, and Agta negrito populations of the Philippines have genomic

Table 6. Mitochondrial DNA Haplogroups from the Four Negrito Populations in the Present Study

HAPLOGROUP BY POSSIBLE ORIGIN	AETA		AGTA		PRESENT STUDY:	PERCENTAGE TABBADA ET AL. 2010			TREJAUT ET AL.
	NEGA	NEGJ	NEGC	NEGU	NEGRITO	LUZON	VISAYAS	MINDANAO	2005: TAIWAN
Taiwan									
B4a1a	3	9	0	0	10.3	11.86	9.82	12.86	8.70
D4	0	0	0	1	0.9	1.13	0.89	2.86	4.60
M7b3	3	1	0	0	3.4	2.82	5.36	2.86	8.23
M7c3c	3	0	8	1	10.3	12.99	7.14	12.86	7.87
Y2	2	1	0	0	2.6	4.52	2.68	7.14	1.37
Island Southeast Asia									
B4b1	4	6	1	**13**	20.7	7.34	7.14	7.14	6.04
B5b	0	5	1	0	5.2	10.17	4.46	4.29	0.00
F1a3a	9	4	0	0	11.2	3.39	2.86	1.43	0.00
F1a4a	0	0	2	0	1.7	3.39	6.25	2.86	0.00
E1b1	0	1	0	0	0.9	3.09	0.00	1.43	0.00
R9*/F*	0	0	6	0	5.2	0.00	0.00	0.00	0.00
Philippines									
P10	0	0	1	3	3.5	0.52	0.89	1.43	0.00
P9	**5**	**1**	**11**	**11**	24.1	1.13	0.89	0.00	0.00
Number of Individuals	29	28	30	29	116	177	112	70	640

The three categories of haplogroups are grouped according to previous attempts to define their possible origins. Haplogroups shared among all four negrito populations are shown in boldface. Data from Tabbada et al. (2010) and Trejaut et al. (2005) are included for comparison.

heterozygosity values similar to those of other populations in East Asia. Our results, therefore, are consistent with a long-term settlement and high effective population size, coupled with low levels of gene flow between these populations.

Differences in Sex-Specific Effective Population Sizes. A second point of considerable interest is the ratio of male to female effective population sizes, which is elevated among the Agta and reduced among the Aeta. Several sex-specific behaviors can reduce sex-specific effective population size (see Heyer et al. 2012 for a review). For example, if polygyny is practiced, it can reduce the male effective population size, whereas, conversely, polyandry should reduce the female effective population size. Both of these examples of polygamy need to be extremely strong in order to have a significant impact on the sex-specific effective population size. The cultural transmission of reproductive success, however, has the potential to significantly reduce the effective population size in a sex-biased manner.

Incorporation of Putative Austronesian Haploid Lineages. According to proposals for Out-of-Taiwan mtDNA lineages, the Aeta may have incorporated more female immigrants from Austronesian settlers than did the Agta, consistent with the higher female effective population size for the Aeta. Conversely, there is an elevated frequency of mtDNA lineages of clear-cut indigenous origin observed among the Agta; P9 and P10 contribute approximately 44% of the Agta mtDNA lineages, compared with only about 11% in the Aeta. These two rare lineages link the negritos to the maternal prehistory of New Guinea and Australia as part of macrohaplogroup P (Hudjashov et al. 2007), although it is not possible to discern whether they represent a signal of the pioneering settlement of the region or a subsequent back-migration from New Guinea (Tabbada et al. 2010).

When comparing Aeta with Agta for Y-chromosomal haplogroups thought to be associated with the Austronesian dispersal, haplogroup O lineages constitute approximately 50% of the Agta paternal lineages, compared with about 14% among the Aeta. These results are consistent with a previous study (Delfin et al. 2011) and suggest a differential incorporation of male immigrants of putative Austronesian descent into the ancestors of the Agta. This skewing observed in patterns of paternal lineages is consistent with the elevated ratio of male:female effective population sizes among the Agta, estimated from the autosomal and X-chromosomal STR data.

A previous estimate for admixture based on SNPs found that the Asian contribution to another group of Aeta was approximately 17% smaller for the X chromosome than for the autosomes, while balanced among the general Philippine population (Cox et al. 2010). Although our own nuclear DNA data indicate different sex-specific behaviors between negrito populations, we only compare these groups to the general population using haploid loci. The apparent "discrepancy" for the Aeta with admixture data from the nuclear genome, therefore, may be due to the limited utility of mtDNA to infer demographic history, or differences in sex-specific behavior within other groups of Aeta in the past.

Pre-Austronesian Haploid Lineages. The mtDNA haplogroups B4b1, B5b, E1b1, F1a3, F1a4, and R9* constitute ~33–57% of the Agta and Aeta maternal lineages but are not unique to the Philippines and are considered to have distributions arising prior to the dispersal of Austronesian speakers from Taiwan (Hill et al. 2007; Tabbada et al. 2010). Of these, B4b1 merits particular attention because of its ubiquitous presence among the four negrito populations, across the East Asian mainland and Japan (Trejaut et al. 2005). The relative prevalence of B4b1 among the negritos of Luzon suggests a presence in the Philippines prior to the Austronesian expansion approximately 4–6 kya, which could also predate the long-term separation of the Aeta and Agta indicated by the differentiation in autosomal genetic diversity.

The frequency of Y-chromosomal haplogroups across the negrito populations is also consistent with the preservation of pre-Austronesian lineages. At more than 50%, K*-M9 is elevated compared with the general populations of the Philippines. Some K*-M9 STR haplotypes in the Philippine negrito populations are similar to those found in northern Australia (Delfin et al. 2011). Whether all these individuals belong to the same, as yet undefined, subclade of K, however, is not clear. The lack of congruence between a substantial proportion of the K*-M9 STR haplotypes and other Philippines individuals from this phylogenetic position in our results is consistent with the high levels of differentiation between the four negrito groups. C*-M216 is present in a minority of negrito individuals (~11%), at levels equivalent to the non-negrito population, but is also presumed to be pre-Austronesian in its time depth (Delfin et al. 2011).

Exploring Possible Correlations with Sociocultural Factors. An overall limited incorporation of maternal and paternal lineages into Aeta and Agta populations from incoming Austronesian settlers during the middle to late Holocene is consistent with the signals emerging from studies of their languages. Although the negritos all speak Austronesian languages, these are commonly isolates or first-order groups within their subfamily and maintain features of the Austronesian parent language lost in those of their non-negrito neighbors (Reid this issue). This combination of features can be interpreted as evidence for an initial social interaction between ancestral negrito groups in Luzon and Austronesian settlers, during which time the Austronesian languages were adopted, followed by a withdrawal of the negritos into relative social and physical isolation (Reid this issue).

The social factors leading to the differential incorporation of haploid markers into the Agta and Aeta are likely linked to different sex-specific patterns of behavior in the past, reflected in the ratios of F_{ST} for autosomal and X-chromosomal STRs. The negrito peoples today have a bilateral descent system (Headland 1987; Turner this issue), but little is known about their social structure in the past. Our findings, however, allow for other kinship systems with sex-specific descent rules to explain the genetic evidence for contrasting sex-specific behaviors among the four negrito populations of the present study as it has been shown to exist in Central Asia (Chaix et al. 2007). One potential alternative—but not mutually exclusive—phenomenon

to explain these results is the transmission of reproductive success enhanced by sociocultural factors, which leads to advantages for "local" females among the Agta and for "local" males among the Aeta. The female transmission of reproductive success based on social factors has been detected in other hunter-gatherer populations (Blum et al. 2006).

Conclusion

Significant levels of genetic differentiation for autosomal, X-linked, and haploid loci, combined with the maintenance of intrapopulation diversity, are consistent with the long-term settlement of the Luzon region by the ancestors of the four negrito populations, followed by reduced gene flow between themselves but no significant reduction in effective population sizes. The overall high levels of Y-chromosomal and mtDNA lineages of indigenous origin retained by the Agta and Aeta suggest limited incorporation of mid-Holocene Austronesian settlers and are consistent with interpretations from linguistic studies for relative isolation following the initial contact period. The ratios of male:female effective population sizes among the various populations, however, have been strongly skewed in a sex-specific manner by sociocultural factors in the past. These factors are likely responsible for the differential incorporation of male and female lineages argued to have accompanied the expansion of the Austronesian language family approximately 6–4 kya, into Agta and Aeta populations, respectively. These findings underscore the advantages of multilocus studies to investigate past demographic processes operating and their effect on shaping the genetic diversity of the present among the negrito populations of Luzon.

Acknowledgments We thank all participants to the study, Rosella Cuenco-Pachner for help in the sampling, and two reviewers for improving the manuscript.

Received 19 November 2012; revision accepted for publication 14 February 2013.

Literature Cited

Bean, R. B. 1910. Types of negritos in the Philippines. *Am. Anthropol.* 12:220–236.

Blum, M. G., E. Heyer, O. Francois et al. 2006. Matrilineal fertility inheritance detected in hunter-gatherer populations using the imbalance of gene genealogies. *PLoS Genet.* 2:e122.

Brion, M., J. J. Sanchez, K. Balogh et al. 2005. Introduction of a single nucleotide polymorphism-based "major Y-chromosome haplogroup typing kit" suitable for predicting the geographical origin of male lineages. *Electrophoresis* 26:4,411–4,420.

Chaix, R., L. Quintana-Murci, T. Hegay et al. 2007. From social to genetic structures in central Asia. *Curr. Biol.* 17:43–48.

Cox, M. P., T. M. Karafet, J. S. Lansing et al. 2010. Autosomal and X-linked single nucleotide polymorphisms reveal a steep Asian-Melanesian ancestry cline in eastern Indonesia and a sex bias in admixture rates. *Proc. Biol. Sci.* 277:1,589–1,596.

Delfin, F., J. M. Salvador, G. C. Calacal et al. 2011. The Y-chromosome landscape of the Philippines: Extensive heterogeneity and varying genetic affinities of negrito and non-negrito groups. *Eur. J. Hum. Genet.* 19:224–230.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1:47–50.

Garza, J. C., and E. G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. *Mol. Ecol.* 10:305–318.

Gordon, R. G., and B. F. Grimes. 2005. *Ethnologue: Languages of the World*. Dallas, Texas: SIL International.

Gray, R. D., A. J. Drummond, and S. J. Greenhill. 2009. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323:479–483.

Gunnarsdottir, E. D., M. Li, M. Bauchet et al. 2011. High-throughput sequencing of complete human mtDNA genomes from the Philippines. *Genome Research* 21:1–11.

Hardy, O. J., and X. Vekemans. 2002. SPAGeDi: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* 2:618–620.

Headland, T. N. 1987. Kinship and social behavior among Agta negrito hunter-gatherers. *Ethnology* 26:261–280.

Heyer, E., R. Chaix, S. Pavard et al. 2012. Sex-specific demographic behaviours that shape human genomic variation. *Mol. Ecol.* 21:597–612.

Hill, C., P. Soares, M. Mormina et al. 2007. A mitochondrial stratigraphy for Island Southeast Asia. *Am. J. Hum. Genet.* 80:29–43.

Hudjashov, G., T. Kivisild, P. A. Underhill et al. 2007. Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. *Proc. Natl. Acad. Sci. USA* 104:8,726–8,730.

Jordan, F. M., R. D. Gray, S. J. Greenhill et al. 2009. Matrilocal residence is ancestral in Austronesian societies. *Proc. Biol. Sci.* 276:1,957–1,964.

Kayser, M. 2010. The human genetic history of Oceania: Near and remote views of dispersal. *Curr. Biol.* 20:R194–201.

Kayser, M., S. Brauer, R. Cordaux et al. 2006. Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. *Mol. Biol. Evol.* 23:2,234–2,244.

Lansing, S. J., M. P. Cox, T. A. de Vet et al. 2011. An ongoing Austronesian expansion in Island Southeast Asia. *J. Anthropol. Archaeol.* 30:262–272.

Loo, J. H., J. A. Trejaut, J. C. Yen et al. 2011. Genetic affinities between the Yami tribe people of Orchid Island and the Philippine Islanders of the Batanes archipelago. *BMC Genet.* 12:21.

Migliano, A. B., I. G. Romero, M. Mespalu et al. 2013. Evolution of the pygmy phenotype: Evidence of positive selection from genome-wide scans in African, Asian, and Melanesian pygmies. *Hum. Biol.* 85:251–284.

R Development Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Reid, L. A. 1994. Unraveling the linguistic histories of Philippine negritos. In *Language Contact and Change in the Austronesian World*, T. Dutton and D. T. Tyron, eds. Berlin, Germany: Mouton de Gruyter, 443–475.

Reid, L. A. 2013. Who are the Philippine negritos? Evidence from language. *Hum. Biol.* 85: 329–358.

Rosenberg, N. A., J. K. Pritchard, J. L. Weber et al. 2002. Genetic structure of human populations. *Science* 298:2,381–2,385.

Scholes, C., K. Siddle, A. Ducourneau et al. 2011. Genetic diversity and evidence for population admixture in Batak negritos from Palawan. *Am. J. Phys. Anthropol.* 146:62–72.

Segurel, L., B. Martinez-Cruz, L. Quintana-Murci et al. 2008. Sex-specific genetic structure and social organization in Central Asia: Insights from a multi-locus study. *PLoS Genet.* 4:e1000200.

Tabbada, K. A., J. Trejaut, J. H. Loo et al. 2010. Philippine mitochondrial DNA diversity: A populated viaduct between Taiwan and Indonesia? *Mol. Biol. Evol.* 27:21–31.

Thangaraj, K., L. Singh, A. G. Reddy et al. 2003. Genetic affinities of the Andaman Islanders, a vanishing human population. *Curr. Biol.* 13:86–93.

Trejaut, J. A., T. Kivisild, J. H. Loo et al. 2005. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biol.* 3:e247.

Turner, J. W. 2013. Kinship matters: Structures of alliance, indigenous foragers, and the Austronesian diaspora. *Hum. Biol.* 85:359–382.

Verdu, P., F. Austerlitz, A. Estoup et al. 2009. Origins and genetic diversity of pygmy hunter-gatherers from western Central Africa. *Curr. Biol.* 19:312–318.

Verdu, P., N. S. Becker, A. Froment et al. 2013. Sociocultural behavior, sex-biased admixture, and effective population sizes in Central African Pygmies and non-Pygmies. *Mol. Biol. Evol.* 30:918–937.

Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* 38:1,358–1,370.