Original Research Article

The Origins and Genetic Distinctiveness of the Chamorros of the Marianas Islands: An mtDNA Perspective

MIGUEL G. VILAR,1,2* CHIM W CHAN,2,3,4 DANA R SANTOS,2,5 DANIEL LYNCH,6 RITA SPATHIS,2,3,4
RALPH M GARRUTO,2,4 AND J KOJI LUM2,3

Background: Archaeological and linguistic evidence suggests the Marianas Islands were settled around 3,600 years before present (ybp) from Island Southeast Asia (ISEA). Around 1,000 ybp latte stone pillars and the first evidence of rice cultivation appear in the Marianas. Both traditions are absent in the rest of prehistoric Oceania.

Objective: To examine the genetic origins and postsettlement gene flow of Chamorros of the Marianas Islands.

Methods: To infer the origins of the Chamorros we analyzed ~360 base pairs of the hypervariable-region 1 (HVS1) of mitochondrial DNA from 105 Chamorros from Guam, Rota, and Saipan, and the complete mitochondrial genome of 32 Guamanian Chamorros, and compared them to lineages from ISEA and neighboring Pacific archipelagoes from the database.

Results: Results reveal that 92% of Chamorros belong to haplogroup E, also found in ISEA but rare in Oceania. The two most numerous E lineages were identical to lineages currently found in Indonesia, while the remaining E lineages differed by only one or two mutations and all were unique to the Marianas. Seven percent of the lineages belonged to a single Chamorro-specific lineage within haplogroup B4, common to ISEA as well as Micronesia and Polynesia.

Conclusions: These patterns suggest a small founding population had reached and settled the Marianas from ISEA by 4,000 ybp, and developed unique mutations in isolation. A second migration from ISEA may have arrived around 1,000 ybp, introducing the latte pillars, rice agriculture and the homogeneous minority B4 lineage. Am. J. Hum. Biol. 00:000–000, 2012. © 2012 Wiley Periodicals, Inc.
Also unique to the Marianas are prehistoric rice and latte. The first archaeological evidence of rice are grains infused in pottery shards dating to around 1,000 ybp (Hunter-Anderson et al., 1995). Although rice cultivation was common throughout ISEA at the time, the Marianas are the only known Pacific islands where cereal plants were grown prehistorically. Around the same time, the first latte (Thompson, 1940) structures were constructed. Latte were rectangular stone columns topped by a hemisphere, some totaling more than 5 m tall, and thought to have functioned as house supports. Although similar in shape to wooden building supports found in Wallacea, stone latte are unique to the Marianas.

Another unique aspect to the Marianas is the Chamorro language. Chamorro is an Austronesian language with no mutually intelligible extant-related language. It is thought to have evolved in isolation from proto-Malayo-Polynesian (MP) (Bellwood, 1991), and currently it is only distantly related to other western MP languages spoken in the southern Philippines and Sulawesi (Fig. 1). In contrast, the languages of neighboring Caroline Islands south and east of the Marianas are part of the Oceanic subgroup of the Austronesian family with roots in eastern Melanesia and directly associated with the spread of the Lapita Cultural Complex (Kirch, 2000).

Like other Pacific islanders, early Chamorros were large and robust. Skeletal remains from late precontact Chamorros show that their humeral diameters, nuchal areas, and upper bodies in general were stouter than those of nearly all other modern human populations, and rivaled Neanderthals in humeral robustness (Heathcote, 2006). Spanish conquistadors wrote in detail about the Chamorro’s muscular physiques and defiant nakedness unique among all indigenous Pacific populations (Rogers, 1995).

During Spanish colonization the Chamorro population was decimated by disease and war, declining from an estimated 50–100,000 in the early 17th century to fewer than 1,000 by 1820 (Rogers, 1995). It was during this period of conflict and religious conversion that the Spanish emptied the island of Saipan, and relocated most Chamorros from throughout the Marianas to Guam (Bowers, 1950). Saipan was later resettled by Carolinians from the atolls to the south. Today, many Saipan islanders trace their ancestry to the Caroline Islands, while others self-identify as Chamorilians, or of mixed Chamorro and Carolinian ancestry.

**EARLIER BIOLOGICAL STUDIES**

Earlier genetic studies (Plato and Cruz, 1967) as well as recent molecular studies of Micronesian microsatellites and mitochondrial DNA (mtDNA) (Lum et al., 1998; Lum and Cann, 2000) have shown the Chamorros as quite distinct from their neighboring populations. More than 85% of Chamorros belong to mtDNA haplogroups E1 and E2 (Lynch et al., 2008; Reiff et al., 2011) relatively common (15–20%) in ISEA (Hill et al., 2007; Soares et al., 2008; Trejaut et al., 2005), but rare (<5%) in other Pacific archipelagoes (Friedlaender et al., 2007; Lum and Cann, 2000; Vilar et al., 2008). Most remaining Chamorro lineages (~10%) belong to a unique lineage of haplogroup B4a1a1a (Lynch et al., 2008; Reiff et al., 2011), a group common in ISEA and Melanesia (10–20%) (Hill et al., 2007; Soares et al., 2011; Trejaut et al., 2005; Vilar et al., 2008), and the most common group (>85%) in Central and Eastern Micronesia (Lum and Cann, 2000) and Polynesia (Sykes et al., 1995). Lineage B4a1a1a, also referred to as the Polynesian Motif, is the most common lineage of the B4 haplogroup in Oceania, and has been previously associated with the Lapita expansion (Lum and Cann, 1998, 2000; Merriwether et al., 1999; Vilar et al., 2008).

Our genetic inquiry into Chamorro origins addresses whether they are direct descendents of a single mid-Holocene migration from ISEA around 4,000 ybp, with latte structures and rice developing in situ, or descendents of two migrations from ISEA, the second which occurred around 1,000 ybp introducing rice and latte. Additionally,
we address whether the Chamorro gene pool shows any evidence of gene flow from the culturally and linguistically distinct neighboring archipelagoes.

**PARTICIPANTS AND METHODS**

To infer the origins of the Chamorros, we analyzed the mtDNA hypervariable-region 1 (hvs1) sequences of 105 self-identified Chamorro volunteers (Guam \( n = 85 \), Rota \( n = 6 \), Saipan \( n = 14 \)) and 17 Saipan islanders of Carolinian maternal ancestry from samples collected in the Marianas Islands (Table 1). In addition, we re-sequenced the hvs1 sequences from 210 individuals from neighboring Palau, Yap, and the Caroline Islands adding an additional 200 base pairs (bp) to previously published data (Lum and Cann, 2000).

DNA was extracted from: (a) plucked hair samples using the method of Higuchi (1989), (b) blood and tissue using the silica extraction protocol of Boom et al. (1990), (Lum and Cann, 2000; Lynch et al., 2008), or (c) blood using Qiagen spin-column (Qiagen Valencia, CA) following manufacturer’s protocol. A 405bp segment of the mtDNA hvs1
of the control region was amplified by Polymerase Chain Reaction (PCR) using primers L15996 and H16401 as described previously (Lum and Cann, 2000; Vigilant et al., 1989). The PCR product was purified using a Manu 03050 Filter Plate (Millipore Corporation, Billerica, MA) and sequenced in both directions with the Big Dye Terminator Kit version 3.1 in an Applied Biosystems 3730xl capillary DNA analyzer (Applied Biosystems, Foster City, CA). In total, 360bp of unambiguous mtDNA hvs1 sequence were determined for each participant and used for phylogenetic analysis. The 122 sequences were compared to the 210 sequences from neighboring archipelagoes (Lum and Cann, 2000), as well as to other sequences previously published on Genbank (http://www.ncbi.nlm.nih.gov). All sequences were aligned using ClustalX (Thompson et al., 1997). A median-joining network of the 105 Chamorro individuals (Fig. 2) was constructed using Network 4.500 (fluxus-engineering.com), and times of coalescence were estimated using a mutation rate of 1 substitution every 16,667 years (Soares et al., 2009) for the extended version of hvs1 (bp 16,050–16,400).

Additionally, complete mitochondrial genomes were sequenced for 32 Chamorro individuals (Reiff et al., 2011). Twenty-four overlapping fragments of ~800 bp were amplified using primers and conditions described by Rieder et al. (1998). PCR products were purified and sequenced and polymorphisms identified as described earlier for hvs1. The complete genome haplotypes (Supporting Information Fig. S1) were compared with other mitochondrial genome haplotypes deposited in Genbank. Network phylogenies (Fig. 3) were constructed for the whole genomes of haplogroup E lineages and compared to other previously published complete genomes. No time estimates were used for the complete genome sequences since the samples were originally selected for a biomedical study (Reiff et al., 2011) and may not be representative of the whole Chamorro diversity.

Approval for this study was obtained from the Committee for Compliance on Experimentation with Human Subjects of the University of Hawai‘i at Manoa, the Human Subjects Review Committee of the Federated States of Micronesia, and the traditional leaders of Palau (Rubeuk Belau) and Yap State (Council of Pilung), as well as the Binghamton University Human Subjects Research Review Committee.

**RESULTS**

**HVS1 results**

The 122 individuals yielded 19 unique hvs1 haplotypes, while the 105 Chamorros yielded 14 hvs1 haplotypes (GenBank accession numbers: JX880090-JX880103). Of the 105 Chamorro lineages shown in the Median-joining Network (Fig. 2), 92% ($n=97$) shared mtDNA hvs1 transitions at base pairs (bp) 16,223, 16,362, and 16,390 (Table 1), characteristic of haplogroups E1 and E2. None of the 17 individuals of Carolinian ancestry had these mutations.

Approximately 65% ($n=68$) of Chamorro lineages had an additional transition at bp 16,051, a polymorphism that defines haplogroup E2. The network cluster for this subgroup (Fig. 2) was composed of one high-frequency central node ($n=49$) and six low frequency branch tips ($n<6$) characteristic of a young expanding population. The time of coalescent for haplogroup E2 was estimated to be $5,641 \pm 2,807$ years (fluxus-engineering.com) using a mutation rate of one substitution every 16,677 years for the extended hvs1 region (Soares et al., 2009).

Approximately 28% ($n=29$) of Chamorro lineages belonged to E1, a haplogroup distinguished from E2 by the 16051 transition. The E1 cluster was also characteristic of a young expanding population. The central node was the second most common lineage in our analysis ($n=25$). This network had four low-frequency branches ($n=1$) radiating from the central node. The time of coalescence for haplogroup E1 was $2,875 \pm 1,286$ years (fluxus-engineering.com) using the mutation rate listed earlier.

Haplogroup B4 accounted for only 8% ($n=8$) of Chamorro lineages, but 100% ($n=17$) of individuals of Carolinian ancestry from Saipan (Table 1). Of the 25 haplogroup B4 lineages, lineage B4a1a1a (sharing mutations at 16,189, 16,217, 16,247, and 16,261) made up 88%
(n = 22) of the haplogroup B lineages, most of which were from Carolinians in Saipan. All seven haplogroup B4a1a1a lineages from Guam and Rota had an additional transition at bp 16114 (C16114T) unique to those two islands. All B4a1a1a lineages from Saipan, as well as those in the rest of Western Micronesia, ISEA, and throughout GenBank, lacked the 16,114 transition.

**Complete mtDNA genome results**

To determine the sub-grouping of the lineages and similarity to other populations, 32 Chamorro complete mitochondrial genomes were also analyzed (Reiff et al., 2011, GenBank accession numbers: HQ700839-HQ700870) (Supporting Information Fig. S1). Eighteen of the sequences were from E2 individuals (the largest group), 12 were from E1, and two were B4. The 18 E2 mitochondrial genome sequences, including fourteen individuals from the central hvs1 node, generated seven unique E2a haplotypes not found among the published haplogroup E2 complete genome sequences from ISEA or Melanesia (Friedlaender et al., 2007; Gunnarsdottir et al., 2011; Soares et al., 2008; Tabbada et al., 2010). The twelve E1 mitochondrial genomes yielded nine distinct E1a2 haplotypes that were also unique and at least three mutations away from the seven closest E1a2 haplotypes found in Indonesia, the Philippines, and the Bismarcks (Friedlaender et al., 2007; Soares et al., 2008; Tabbada et al., 2010) (Fig. 3). Our two complete B4a1a1a genomes were identical to one another (Reiff et al., 2011), but also unique in the Pacific, differing by two transitions (G15257A and C16114T) from the presumed ancestral haplotype (Polynesian Motif), and at least three mutations different from individuals from other Oceanic, Indonesian, and even Malagasy populations (Friedlaender et al., 2007; Rafaindrazaka et al., 2010; Soares et al., 2011).

**DISCUSSION**

To infer the putative origins of the Chamorros and deduce the source populations and gene flow associated with the modern Chamorro gene pool we analyzed the mtDNA hvs1 sequences of 105 Chamorro individuals. And with a higher resolution analysis we also analyzed the complete mtDNA genome of a subgroup of 32 individuals (Reiff et al., 2011).

**Haplogroups E1 and E2**

The Chamorro hvs1 analysis showed very limited diversity. Three lineages accounted for more than 65% of all samples. Two of these lineages, the central nodes for the haplogroups E1 and E2 clusters have been found in low frequencies in the Philippines, and the Sulawesi and Maluku islands of Indonesia (Hill et al., 2007). However, none of the branch tips for either haplogroup E cluster (Fig. 2) were found outside of the Marianas. This pattern is indicative of a founder effect, where two founding Haplogroup E lineages (our central nodes) arrived in the Marianas from ISEA with the archipelago's first settlers ~4,000 ybp. Over the course of 3,500 years of isolation the two ancestral lineages acquired the mutations that gave rise to the branch tips (Fig. 2) unique to the Marianas.

The complete genome analysis revealed a similar pattern. All complete genome haplotypes were unique to the Marianas, but closely related (fewer than five mutations different) to haplotypes from ISEA, specifically: Sulawesi, Sumba, the Moluccas, Borneo and the Philippines (Fig. 3). The most common E2a complete genome haplotype in our data set (n = 12) was also the inferred ancestral lineage to all E2a haplotypes in the Marianas as well as to those in Genbank from ISEA. Further, given the lineage age estimates from Soares et al. (2008), this pattern of Chamorro diversity suggests that the migration from ISEA to the Marianas may have occurred as early as 5,000 ybp. Since that time, Chamorros and populations in ISEA independently acquired distinct mutations found in the derived lineages throughout ISEA as well as in the Marianas. This time frame is consistent with the dates obtained from indirect evidence of human settlement as suggested by charcoal remains and changes in pollen composition. The complete genome E1a2 Chamorro lineages included two individuals which were the ancestral lineage. This lineage has now been found among Chamorros, but not among ISEA individuals previously published (Fig. 3). The occurrence of the ancestral form of E1a2 and lineages recently derived (three or fewer mutations away) from the ancestral form among the Chamorro data set, and likewise recently derived lineages throughout several islands of Eastern Indonesia (Soares et al., 2008), suggests a strong link between the two geographic regions that also dates to approximately 4,000–5,000 ybp.

**Haplogroup B4**

The third high-frequency lineage was the Chamorro unique B4 lineage, or the B4a1a1a1a lineage with the additional C16114T transition. This was the only B4a1a1a1a lineage found on Guam and Rota. Moreover, C16114T was absent from 17 related haplogroup B4a individuals of Carolinian ancestry from Saipan, as well as the 236 haplogroup B4a individuals from Palau, Yap, and the outer islands of Yap. That specific transition was also absent from the B4a1a1a1a lineages from ISEA, Melanesia, Micronesia, and Polynesia published in GenBank. Although the presumed hvs1 ancestral form (Polynesian Motif) has been reported throughout Micronesia, Melanesia, Polynesia (Lum and Cann, 2000; Sykes et al., 1995; Vilar et al., 2008) as well as in low frequency in Sulawesi, the Moluccas (Hill et al., 2007; Soares et al., 2011) and the Bird's Head peninsula of West New Guinea (Cascione et al., 2006), the lineage inclusive of C16114T mutation was unique not only to Chamorros, but specifically to those of Guam and Rota. The genetic homogeneity and uniqueness of this lineage to Guam and Rota, suggests it may be a recent introduction to the islands. One scenario is that a second migration associated with the introduction of rice agriculture and latte, both practices found nowhere else in the Pacific, arrived ~1,000 ybp introducing this minority lineage. However, given the high frequency of the ancestral type (Polynesian Motif) throughout Yap and the Caroline Islands and the proximity of these archipelagoes to the Marianas, we cannot rule out the possibility that this B4 lineage arrived from the neighboring Micronesian archipelagoes sometime in the last 2,000 years (Lum and Cann, 2000) and acquired mutation C16114T in the Marianas, while losing the ancestral Polynesian Motif form.

Most B4 lineages from Carolinians in Saipan are exact matches to lineages found in the Caroline Islands (Lum and Cann, 2000) thus suggesting high gene flow between
the two populations or complete replacement of one population by another. This pattern supports the historic accounts of the forced removal of Chamorros from Saipan and subsequent repopulation by Carolinians.

Last, the occurrence of a single haplogroup B4b1 lineage in Guam may be the result of gene flow from Japan or elsewhere in East Asia. Further investigation into the individual’s genealogy may be needed to clarify the exact origin.

CONCLUSIONS

Our data suggests that the Marianas were settled between 5,000 and 3,500 years ago directly from ISEA, likely Wallacea (Sulawesi and the Moluccas) where E1a2 and E2a lineages have been previously identified and may have originated (Hill et al., 2007; Soares et al., 2008). Both haplogroups E1a2 and E2a likely arrived differentiated during that first migratory settlement. Then the population remained genetically isolated from ISEA while the lineages developed mutations in isolation now unique to the Marianas. The unique B4a1a1a1 lineage from Guam likely descendants of recent movement back from Guam and Rota, while the B4a1a1a1 lineages in Saipan (those of Carolinians from the outer islands of Yap, and none with their neighbors in Guam and Rota or even with Chamorros on the same island of Saipan) are likely descendants of recent movement back from Guam and Rota, while the B4a1a1a1 lineages in Saipan (those of Carolinian ancestry) are a product of historic resettlement by Carolinians. This genetic connection between Saipan and the Caroline islands is the only evidence of mtDNA gene flow between the Marianas and their neighboring archipelagoes.

Our on-going Y-Chromosome DNA and autosomal STR analyses on the Chamorros may reveal greater gene flow from Spanish or Mexican populations. No haplotypes commonly associated with these populations were found in our mtDNA analysis. We do expect to see male-biased gene flow from Spain, Mexico and the Philippines, which could have resulted from the imposed male-dominated colonization, the role of the Marianas as a stopping-point in trans-Pacific trade, and the forced immigration of Filipino workers during colonial times.

ACKNOWLEDGMENTS

The authors thank the National Institute of Neurological Disorders and Stroke of the National Institutes of Health in Bethesda Maryland, the Binghamton University Serum Archive Facility and Graduate Program in Biomedical Anthropology, and all individuals from Micronesia who donated samples and anxiously await the results of our study.

LITERATURE CITED


Cascione I, Attimonelli M, Sukryani Y, Noer SA, Marzuki S, Tommaseo-Ponzetta M. 2006. La variabilità mitocondriale delle popolazioni del Bird’s Head. Received through personal communication.


Green RC, Davidson J. 1969.Archaeology in Western Samoa. Vol II. Bulle-


Heathcote G. 2006. Taotao Tagga: Glimpses of his life history, recorded in his skeleton. Anthropology Resource Center, University of Guam, Mangilao, GU.


Hunter-Anderson RL, Thompson GB, Moore DR. 1995. Rice as a prehis
toric valuable in the Mariana Islands, Micronesia. Asian Perspectives 34:69–89.


Solheim WG. 2006. Archaeology and culture in Southeast Asia: unraveling the Nusantao University of the Philippines Press, Quezon City, Philippines.

American Journal of Human Biology