

More Genetic Sharing among the Populations of Taiwan than Expected: A Plain tribes (Pinpu) Perspective

Marie LIN, Chien-Liang LEE, Jun-Hun LOO, Ju-Chen YEN, Chen-Chung CHU,
Hui-Ling LEE and Jean A. TREJAUT

Molecular Anthropology and Transfusion Medicine Laboratory
Mackay Memorial Hospital, Taipei, Taiwan

Background:

Geographically Taiwan is on the Pacific Rim, in a position central to continental Southeast Asia (CSEA, southeast China and Indochina) and Island Southeast Asia (ISEA, the Philippines and Indonesians). During the Glacial period the Taiwan Strait was a land bridge connecting Taiwan to continental Asia which, somehow, distantly connect to the Sunda landmass (Barker et al. 2002). Taiwan has been the refuge of many immigrants coming from the continent through land and/or water. Ancient arrivals, predating 15000 years ago in Taiwan, are possible, but so far, using mitochondrial DNA coalescence methods, three periods of settlement can be outlined:

- _The earliest (late Paleolithic) arrival in Taiwan dates up to 15000 years ago and correspond to haplogroups found only among TwA (Taiwan Aborigines) and ISEA (Island Southeast Asians): E, B5a2b and M7b3.
- _The second period (early Holocene/Neolithic) 6000-10000 years ago: B4a1a, B4a2a, B4b1a, E1a1a, M7c3 (van Oven and Kayser 2009).
- _Late "recent local expansion" of 1000 to 4000 years ago: E2b2 and a few others.

In the last 400 hundred years, Japanese, Dutch and Spanish have resided for short periods in Taiwan, before the massive arrival of peoples from southeast coast of China who were already sinicized indigenous peoples of ethnic Yueh origin. The descendant of these modern settlers, the Minnan and Hakka, referred in this abstract as the "non Taiwan Aborigines" (NTwA) constitute today the major population of Taiwan. The NTwA have always considered themselves as the descendant of northern Han Chinese who migrated to southern China to flee from massacres in the north between AD300-1300. When the NTwA first settled in the southwestern plains of Taiwan, they cohabitated with plain tribe people (matriarchal society of Taiwan Aborigines) and from this cohabitation resulted admixture, sinicization and disappearance of the plain tribe cultures.

Further studies on mitochondrial DNA (mtDNA) (Trejaut et al. 2005), on autosomal Human Leukocyte Antigen (HLA) (Lin et al. 2001), and 58 of paternal non recombining Y SNP (NRY) of hundred NTwA (Trejaut and Lin in preparation), showed strong genetic affinity between NTwA and CSEA, and showed that about half

of the NTwA shared their genetic origin with TwA or with ISEA. Also, it was shown about 85% of NTwA could carry at least one gene (either HLA, NRY or mtDNA) of TwA or ISEA origin (Lin et al 2008). Similarly, a study performed on two plain tribes, the Siraya (294 individuals) and the Pazeh (67 individuals), showed that 95% of Siraya individuals carried at least one gene (HLA, NRY or mtDNA) of TwA or ISEA origin, and likewise, the Pazeh showed 89% individual with at least one gene of TwA or ISEA origin (Lin et al. 2009). In this report we try to further analyze the extant genetic make up of the Taiwan plain tribes.

Number of individuals used for the mtDNA/HLA and NRY SNP analysis:

Minnan and Hakka (n=246 for mtDNA/HLA and 94 for NRY SNP); Fujian (149 and 53); Plain tribes (Siraya: 354 and 221, Pazeh: 61 and 39); TwA (Atayal: 108 and 52, Toroko: 53 and 20, Saisiat: 64 and 24, Thao: 26 and 16, Tsou: 60 and 41, Bunun: 89 and 56, Ami: 98 and 39, Paiwan: 55 and 25, Rukai: 51 and 29, Puyuma: 52 and 23, Yami: 79 and 30); Philippines (375 and 122); Indonesia (427 and 246); Thailand (78 mtDNA only) and Hanoi (57 mtDNA only).

Methods: mtDNA sequence based haplogrouping; NRY SNP haplogrouping; HLA-A, B, DRB1 sequence based allele typing.

Results and Discussion:

Analysis showed that 56% of mtDNA haplotypes were shared between plain tribes, TwA and ISEA (B4a1a, B4c1b, B5a2, E1a, E1a1a, F1a1a, F1a3, F1a4, F3b2b, M7b1, M7b3 and M7c3).

Only 8% of mtDNA haplotypes were shared between plain tribes and TwA (B4a2a, B4b1a, D5b3, E2b1, F1a1a'tw, F4b, M7c3a and N9a3).

Also, 28% haplotypes were shared between plain tribes and CSEA (B4, B4a, B4a1, B4b1, B4c1b, B5a1, C6, D, D4, D4a, F1a, F2, F3a, M7, M7b, M10a2, R9b1 and R9e).

Finally the plain tribes possess haplotypes seen in Continental Asia (CA) and northeast Asia. These are: A4, A5b, C, D4b2, D4e, M8a and N9a.

E2a, a lineage never seen in CA, was restricted to Siraya and Pazeh (each had 2 individuals) and ISEA (~1-3%), but was not seen among TwA. As shown with E2a (Soares et al. 2008, Lin et al 2008), exclusive presence of ISEA haplotypes among plain tribe indicate a northward population movement (from ISEA to Taiwan), most likely to the western coast among plain tribe people.

Y SNP (paternal lineages) analysis showed 55% of Siraya and 70% of Pazeh men sharing paternal lineages O1a*, O1a1*, O1a2, O3*, O3a* and O3a3* with TwA and ISEA, and except for O1a* and O1a2, were also seen in Fujian (which we here

classify as CSEA). Among other Y haplogroups, O2*, O2a*, O3a4*, O3a3c* and O3a3c1 comprised 40% of the Siraya and 28% of Pazeh gene pool and were also commonly seen all over CSEA and ISEA. Male movement between CSEA, ISEA and Taiwan will be better understood when high definition NRY-SNP determination (and associated Y-STR) of other Asian populations become available.

Analysis of the HLA-A-B-DRB1 haplotype distribution among the plain tribes, TwA, Fujian, Philippines and Indonesians revealed that minimal sharing of HLA-A-B-DRB1 haplotypes between Siraya or Pazeh and TwA are 56% and 16% respectively; with Fujian 18% and 14% respectively; with the Philippines 9% and 3%; and with Indonesians 7% and 7%. The most common haplotype in Siraya was A*2402-B*4002-DRB1*1401 (12.8%), it was also seen in Yami (22% where it may represent a founding effect), Paiwan (7.8%), Tsou (4.9%), Atayal (2.7%) and the Philippines (1%). This haplotype is not seen in CSEA or CA and suggest restricted insular sharing or common ancestry between Siraya, TwA and ISEA.

The second most common haplotype in Siraya (A*1101-B*1525-DRB1*1602) (4.7%) was also found in Yami (25.3%), Thao (11.7%) and Pazeh (8.2%). Similarly, this haplotype was not seen in CA and appeared more representative of the plain tribes than the former. The third haplotype was A*3303-B*5801-DRB1*0301 (3.3%). It was one of the most common haplotypes in NTwA, CSEA and ISEA, and is possibly a marker of past continental gene flow toward insular Asia.

A glimpse on the genetic ancestry of plain tribe and TwA using ancient DNA (aDNA):

1. HVI and HVII DNA sequences of an ancient human bone from a supposed TwA burial site in the southwestern plains of Taiwan (estimated to > 400 years) was typed as mtDNA haplogroup F2a. In our data base, F2a haplotype is found in Fujian (4%) where one individual was identical to the aDNA specimen ((HVI sequence). F2a is also found in Hakka (3%), Xinmem (11%, an island near Fujian), Siraya (1%) and Vietnam (1.5%). Interestingly, F2a is not found among TwA and therefore may indicate a CSEA origin.

2. aDNA HVI sequence of bone dating 5 to 800 years old from South Science Park of the southwestern plain of Taiwan showed haplotype M33c. An identical haplotype was found in an extant NTwA family (brother and sister) from Kaohsiung who have several generations of Siraya ancestry. To date, no other lineages identical to this haplotype has ever been seen among TwA or CA.

3. Two aDNA HVI sequences of ancient bones dating 4500 years old from South Science Park typed as haplogroup D4 (presence of coding mutation 14668) and M7c3 respectively. D4 haplogroups is found in CSEA and is not seen among TwA.

Haplogroup M7c3 is found in TwA, in ISEA and is less common in CSEA.

4. aDNA HVI sequences and the presence of coding region mutation 6026 allowed the determination of haplogroup C4. This type was found among four different specimens of ancient bones dating 2000 to 4000 years old from the east coast of Taiwan (Hualien). An identical sequence has also been found in a present day NTWA of Tainan. C4 is usually seen among Continental Asians and have never been seen in the extant TwA population.

Conclusion:

Autosomal recombining HLA and non-recombining haploid systems (mtDNA and NRY) of Taiwan present-day populations revealed that plain tribes shared their genes with TwA, ISEA and CSEA, and appear more distant from CA.

Using ancient bone remains from Taiwan plain tribe burial sites, and studying non-recombining maternal mtDNA system, it was shown that the CSEA and CA genetic heritage of the Taiwan plain tribes was not only acquired during the past 400 years from people of the southeast coast of China (current belief), but may have been acquired as far back as 4000 years ago from ancient prehistoric sailing agriculturist migrants (the broad genetic structure in those days was similar to that seen in CSEA today). Older bone remains will be needed to demonstrate older migration, but most importantly these preliminary results will need further analysis in the future. Caution will have to be taken as it is usual to label haplogroups that are not seen among TwA as genetic markers of recent continental origin, here, on the contrary, our results show that the presence of these same haplogroups have arrived much earlier and may actually help to reconstruct the genetic make up of descendants of prehistoric migrants from continental Asia a few thousand years ago.

The Mitochondrial DNA analysis revealed two genetic "Aboriginalities" within the plain tribes, one already well defined set of haplogroups shared with mountain tribe Aborigines (TwA) (and ISEA), and the other set of haplogroups shared between plain tribes Aborigines and CSEA. The arrival time of the latter set of haplogroups could be traced back to several thousand years prior the commonly accepted period of 400 years.

We showed that the genetic make up of the plain tribes today, most likely, contains an older Taiwan heritage than previously thought. This study also support our previous findings describing that 95% of Siraya carried at least one gene of TwA origin (or gene also seen in TwA and ISEA), and Pazeh 89%. Finally, when confirmed these resulted, may have strong implication in drawing new hypothesis about the historical demography of present day Taiwanese.

References:

- Lin M, Trejaut JA, Loo LJ, Chu CC, Lee CL, He CL, Lee HL and Yen JC. (2009) Genetic study of Siraya, Pazeh and Ketangalan. Plain tribes never die. In Proceeding of the Austronesian Classification and their dispersal: Anthropology, archeology, genetic and linguistic studies. Academia Sinica. Linguistic Intitute 2009, January 17th.1-29.
- Lin M, Trejaut JA, Loo JH, He CL, Lee CL, Chu CC, Lee HL, Lin CS.(2008) Genetic profile of non-aboriginal Taiwanese. www.HUGO-AP2008.ph P54.
- Barker G, Barton H, Beavitt P, Bird M, Daly P, Doherty C, Gilbertson D, Hunt C, Krigbaum J, Lewis H (2002) Prehistoric foragers and farmers in southeast Asia: renewed investigations at Niah Cave, Sarawak. Proc Prehist Soc 68:147-164
- Lin M, Chu C, Chang S, Lee H, Loo J, Akaza T, Juji T, Ohashi J, Tokunaga K (2001) The origin of Minnan and Hakka, the so-called "Taiwanese", inferred by HLA study. Tissue Antigens 57:192-199
- Soares P, Trejaut JA, Loo J-H, Hill C, Mormina M, Lee CL, Chen YM, Hudjashov G, Forster P, Macaulay V, Bulbeck D, Oppenheimer S, Lin M, Richards MB (2008) Climate Change and Postglacial Human Dispersals in Southeast Asia. Mol Biol Evol 25:1209-1218
- Trejaut JA, Kivisild T, Loo JH, Lee CL, He CL, Hsu CJ, Li ZY, Lin M (2005) Traces of archaic mitochondrial lineages persist in Austronesian speaking Formosan populations. PloS 3:e247 (241-211)
- van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. <http://www.phylotree.org>. Hum Mutat 30:E386-E394