DNA

MDPI

Review

Origin of the Bunun Indigenous peoples of Taiwan. A Review of published material using Y chromosome and Mitochondrial DNA gene systems

Jean A. Trejaut †

Alumni Community, City Campus, University of Technology of Sydney, 15 Broadway, Ultimo, NSW 2007, Australia
† Retired

Supplementary Text S1

Materials and Methods

1. Samples

All samples used in this review were published material. They comprise, 89 unrelated healthy Bunun indigenous people collected between 2000 and 2005 in the Xinyi Township of the Nantou county, North of the Yushan National park in Taiwan, (forth referred to as the northern Bunun group, Figure 1) [1,2], 50 self-declared Bunun individuals collected in the Sanmin Township, Southwest of Taiwan (Herein referred to as the Central Bunun) [3] (Table 1) and 96 miscellaneous Bunun (Table 1) [4,5]. While most Bunun groups today use the Isbukun dialect, neither linguistic characteristics nor signs of origin in the Mountain range were previously published for the Sanmin [3,6,7] or the miscellaneous Bunun groups. All Bunun groups were analyzed for mtDNA, only the Northern Bunun group was analyzed for mtDNA and Y chromosome. This data was compared to other groups in Taiwan (Table 1) that were previously screened for the mtDNA and Y-chromosome diversity [1,2,8]. The northern Austronesian Speakers of Taiwan (AN_Tw) comprised Saisiyat (n = 87), Atayal (n =157), and Taroko [8]. The central AN_Tw includes the Bunun (n = 235), the Thao (n = 30) and the Tsou (n = 108) [9], and the southern AN_Tw group was composed of the Rukai (n = 99), the Paiwan (n = 105), the Puyuma (n = 91), the Amis (n = 148) and the Yami islanders (n = 123) [3,10]. Non-AN_Tw comprised Taiwanese of Han descent (TwH), namely the Minnan (n = 50) and Hakka (n = 45). Finally, we added samples from neighbor populations, East China coast (Fujian and Matsu; n= 198) [8,11], Mainland Southeast Asia (MSEA), Vietnam (n = 603) [12], Thailand (n = 560) [13] and Malaysia (n = 86) [6,7], individuals from the Philippines (n = 260), western and eastern Indonesia (n= 326 + 72) [1,2,10,14]. Finally, the Nias island (n= 426) in western Indonesia, Lao in MSEA, and Madagascar (n= 229) [13,15,16] were used to determine the origin of some haplogroups.

Y chromosome reference data not shown in Table 1 but used in Tables 5 and 6 for groups in East and Southeast Asia were also obtained from the literature. They included Pazeh and Siraya from Taiwan's western plains [1,17]. The Malayo-Polynesian groups included the Philippines, the Ivatan archipelago, Oceania (Tuvalu) [18], Western Indonesia (Nias, Java, Sumatra), and Eastern Indonesia (including Bali). Mainland Southeast Asia comprised Malaysia [1,19,20], the Cham from Vietnam, Thailand, Lao, and Hainan (with the Cun and Hlai indigenous groups) [1,17,21]. South China [1,17], North China, Japan, Tibet, and Mongolia were obtained from Xue et al [22]. Finally, the Madagascar data set [15,23] was included with East Africa used as an out-group for the Treemix plot.

2. Data analysis

All sequences analyzed herein were obtained from the literature data (Table 1). The mtDNA polymorphism is represented by sequencing obtained from the D-loop HVS-I control region (nucleotide positions nps 16,006 - 16,397), and the coding regions (nps 8000 - 9000, and nps 9959 - 10,900). Complete mtDNA genome for most Taiwan specimen were obtained from the data set of Ko et al. and Trejaut et al. [2,3,9]. Mitochondrial haplogroups were assigned according to Build 17 of Phylotree [24] using Haplogrep-2 available online: https://haplogrep.i-med.ac.at/app [25].
Genotyping of Non-recombining Y chromosome (NRY) was obtained from [1] using 81 Y-SNPs (Table 1) and further analyzed using 17 Y-chromosome short tandem repeats (Y-STRs): DYS19, DYS385I, DYS385II, DYS389I, DYS389II, DYSS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, and Y GATA-H4.

3. Statistical analysis

Population expansion: Population expansion of the Bunun was estimated using Tajima's D [26], Fu's Fs [27], and p-values were generated using 1,000 simulations under a model of selective neutrality as implemented in Arlequin 3.5.2.2 [28]. Further, mtDNA mismatch frequency plots were constructed (Arlequin 3.5.2.2) to determine whether Bunun exhibited evidence of spatial range expansion or a stationary population history [26].

Multidimensional scaling (MDS): MDS plots were constructed using haplogroup frequencies with the Paleontological statistics version 3.25 (PAST) [29].

Phylogenetic analysis: To infer the phylogenetic relationships between Bunun and other groups from Taiwan, East Asia, Peninsular Asia, and ISEA we adjusted all mtDNA haplogroup assignments between studies and constructed a maximum likelihood tree using TreeMix 1.13 [30]. Treemix was run five times with the number of migration events "-m" set to 15. The Y-SNP-STR phylogenetic relationships were analyzed using the same approach. Median-joining networks: Median-joining networks using Y-STR haplotypes of individuals in the background of haplogroups O1b1a1a1a1a1-M88 and O1a2-M50 were constructed separately using the NETWORK 4510 program [31]. Only DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, and DYS393 were included in the Network analyses as some of the literature data used to compare this analysis did not include all STR markers. The STR loci were weighted according to the average of their relative variability within the haplogroup subclades [18] and run with the reduced median-joining method before running median joining (Figure 4). Based on the Y-STR diversity of the haplogroups O1a2-M50 and O1b1a1a1a1a1-M88 the divergence time of the Bunun haplogroups, as determined by the Networks, was estimated using the ϱ method of Zhivotovsky with a 25 years generation time and a mean per-locus and generation mutation rate of 6.9×10−4 [32].

References

1. Trejaut, J.A.; Poloni, E.S.; Yen, J.-C.; Lai, Y.-H.; Loo, J.-H.; Lee, C.-L.; He, C.-L.; Lin, M. Taiwan Y-chromosomal DNA variation and its relationship with Island Southeast Asia. *BMC Genet.* **2014**, *15*, 77. https://doi.org/10.1186/1471-2156-15-77.
2. Trejaut, J.A.; Kivisild, T.; Loo, J.H.; Lee, C.L.; He, C.L.; Hsu, C.J.; Li, Z.Y.; Lin, M. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biol.* **2005**, *3*, e247. https://doi.org/10.1371/journal.pbio.0030247
3. Ko, A.M.; Chen, C.Y.; Fu, Q.; Delfin, F.; Li, M.; Chiu, H.L.; Stoneking, M.; Ko, Y.C. Early Austronesians: into and out of Taiwan. *Am. J. Hum. Genet*. **2014**, *94*, 426–36
4. Tajima, A.; Sun, C.-S.; Pan, I.-H.; Ishida, T.; Saitou, N.; Horai, S. Mitochondrial DNA polymorphisms in nine aboriginal groups of Taiwan: Implications for the population history of aboriginal Taiwanese. *Hum. Genet.* **2003**, *113*, 24–33
5. Melton, T., Clifford, S., Martinson, J., Batzer, M., Stoneking, M. Genetic evidence for the proto-Austronesian homeland in Asia: mtDNA and nuclear DNA variation in Taiwanese aboriginal tribes. *Am. J. Hum. Genet*. **1998**, *63*, 1807–23

6. Hill, C.; Soares, P.; Mormina, M.; Macaulay, V.; Meehan, W.; Blackburn, J.; Clarke, D.; Raja, J.M.; Ismail, P.; Bulbeck, D.; et al. Phylogeography and Ethnogenesis of Aboriginal Southeast Asians. *Mol. Biol. Evol*. **2006**; *23*, 2480–91.
7. Jinam, T.A.; Hong, L.C.; Phipps, M.E.; Stoneking, M.; Ameen, M.; Edo, J.; HUGO Pan-Asian SNP Consortium; Saitou, N. Evolutionary history of continental southeast Asians: "early train" hypothesis based on genetic analysis of mitochondrial and autosomal DNA data. *Mol. Biol. Evol*. **2012,** *29*, 3513–27.
8. Chen, L.-R.; Trejaut, J.A.; Lai, Y.-H.; Chen, Z.-S.; Huang, J.-Y.; Lin, M.; Loo, J.-H. Mitochondrial DNA Polymorphisms of the Saisiyat Indigenous Group of Taiwan. Search for a Negrito Signature. *Edelweiss J. Biomed. Res. Rev.* **2019**, *12***,** 12–18
9. Trejaut, J.A.; Muyard, F.; Lai, Y.-H.; Chen, L.-R.; Chen, Z.-S.; Loo, J.-H.; Huang, J.-Y.; Lin, M. Genetic diversity of the Thao people of Taiwan using Y-chromosome, mitochondrial DNA and HLA gene systems. *BMC Evol. Biol.* **2019**, *19*, 64. https://doi.org/10.1186/s12862-019-1389-0.
10. Loo, J.H.; Trejaut, J.A.; Yen, J.C.; Chen, Z.S.; Lee, C.L.; Lin, M. Genetic affinities between the Yami tribe people of Orchid Island and the Philippine Islanders of the Batanes archipelago. *BMC Genet*. **2011**, *12*, 21.
11. Luo, X.Q.; Du, P.X.; Wang, L.X.; Zhou, B.Y.; Li, Y.C.; Zheng, H.X.; Wei, L.H.; Liu, J.J.; Sun, C.; Meng, H.L.; et al. Uniparental Genetic Analyses Reveal the Major Origin of Fujian Tanka from Ancient Indigenous Daic Populations. *Hum. Biol*. **2020**, *91*, 257–77.
13. Kutanan, W.; Kampuansai, J.; Srikummool, M.; Kangwanpong, D.; Ghirotto, S.; Brunelli, A.; Stoneking, M. Complete mitochondrial genomes of Thai and Lao populations indicate an ancient origin of Austroasiatic groups and demic diffusion in the spread of Tai-Kadai languages. *Hum Genet*. **2017**, *136*, 85–98.
14. Tabbada, K.A.; Trejaut, J.; Loo, J.H.; Chen, Y.M.; Lin, M.; Mirazon-Lahr, M.; Kisivild, T.; De Ungria, M.C.A. Philippine mitochondrial DNA diversity: a populated viaduct between Taiwan and Indonesia? *Mol. Biol. Evol*. **2010**, *27*, 21–31.
15. Razafindrazaka, H.; Ricaut, F.X.; Cox, M.P.; Mormina, M.; Dugoujon, J.M.; Randriamarolaza, L.P.; Guitard, E.; Tonasso, L.; Ludes, B.; Crubézy, E. Complete mitochondrial DNA sequences provide new insights into the Polynesian motif and the peopling of Madagascar. *Eur. J. Hum. Genet*. **2010**, *18*, 575–81.
16. van Oven, M.; Hammerle, J.M.; van Schoor, M.; Kushnick, G.; Pennekamp, P.; Zega, I.; Lao, O.; Brown, L.; Kennerknecht, I.; Kayser, M. Unexpected island effects at an extreme: reduced Y chromosome and mitochondrial DNA diversity in Nias. *Mol. Biol. Evol*. **2011**, *28*, 1349–61.
17. Li, D.; Li, H.; Ou, C.; Lu, Y.; Sun, Y.; Yang, B.; Qin, Z.; Zhou, Z.; Li, S.; Jin, L. Paternal genetic structure of Hainan aborigines isolated at the entrance to East Asia. *PLoS One*. **2008**, *3*, e2168.
18. Delfin, F.; Salvador, J.M.; Calacal, G.C.; Perdigon, H.B.; Tabbada, K.A.; Villamor, L.P.; Halos, S.C.; Gunnarsdóttir, E.; Myles, S.; Hughes, D.A.; et al. The Y-chromosome landscape of the Philippines: extensive heterogeneity and varying genetic affinities of Negrito and non-Negrito groups. *Eur. J. Hum. Genet*. **2010**, *19*, 224–30.
19. Tofanelli, S.; Bertoncini, S.; Castri, L.; Luiselli, D.; Calafell, F.; Donati, G.; Paoli, G. On the origins and admixture of Malagasy: new evidence from high-resolution analyses of paternal and maternal lineages. *Mol. Biol. Evol*. **2009**, *26*, 2109–24.
20. Bekaert, B.; Zainuddin, Z.; Hadi, S.; Goodwin, W. A comparison of mtDNA and Y chromosome diversity in Malay populations. *Int. Congr. Ser.* **2006**, *1288*, 252–5.
21. Cai, X.; Qin, Z.; Wen, B.; Xu, S.; Wang, Y.; Lu, Y.; Wei, L.; Wang, C.; Li, S.; Huang, X.; et al. Human migration through bottlenecks from Southeast Asia into East Asia during Last Glacial Maximum revealed by Y chromosomes. *PLoS One*. **2011,** *6*, e24282
22. Chen, Y.; Chu, J.Y.; Yu, J.K.; Yu, L.; Sun, H.; Lin, K.Q.; Tao, Y.F.; Shi, L.; Huang, X.Q.; Shi, T.L.; et al. Polymorphism of DYS287 on Y chromosome in 28 ethnic populations of China. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao*, **2006**, *28*, 196–201. (In Chinese)
23. Razafindrazaka H. Le peuplement humain de Madagascar : Anthropologie génétique de trois groupes traditionnels. Toulouse: Université Toulouse III - Paul Sabatier, Toulouse, France; **2010**. (In French)
24. Van Oven, M.; Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **2009**, *30*, E386–E394. https://doi.org/10.1002/humu.20921.
25. Weissensteiner, H.; Pacher, D.; Kloss-Brandstätter, A.; Forer, L.; Specht, G.; Bandelt, H.; Kronenberg, F.; Salas, A.; Schönherr, S. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucl. Acids Res*. **2016,** *44*, W58–W63

26. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. **1989**, *123*, 585–95.
27. Fu, Y.X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*. **1997,** *147*, 915–25.
28. Excoffier, L.; Lischer, H.E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour*. **2010**, *10*, 564–7.
29. Hammer, O.; Harper, D.A.T.; Ryan, P.D. Past: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol. Electron*. **2001,** *4,* 9. Available online: http://palaeo-electronica.org/2001_1/past/issue1_01.htm (accessed on 17 July 2022).
30. Pickrell, J.K.; Pritchard, J.K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet*. **2012**, *8*, e1002967.
31. Bandelt, H.J.; Forster, P.; Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol*. **1999**, *16*, 37–48.
32. Zhivotovsky, L.A.; Underhill, P.A.; Cinnioğlu, C.; Kayser. M.; Morar, B.; Kivisild, T.; Scozzari, R.; Cruciani, F.; Destro-Bisol, G.; Spedini, G.; et al. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am. J. Hum. Genet.* **2004**, *74*, 50–61