

## SHORT COMMUNICATION

# mtDNA diversity of the Zapotec in Mexico suggests a population decline long before the first contact with Europeans

Jun Gojobori<sup>1</sup>, Fuzuki Mizuno<sup>2</sup>, Li Wang<sup>3</sup>, Keisuke Onishi<sup>2</sup>, Julio Granados<sup>4</sup>, Celta Gomez-Trejo<sup>5</sup>, Víctor Acuña-Alonzo<sup>5</sup> and Shintaroh Ueda<sup>2</sup>

The New World is the last continent colonized by anatomically modern humans, *Homo sapiens*. The first migrants entered the New World from Asia through Beringia. It is suggested that there were three streams of Asian gene flow, one major and two additional minor gene flows. The first major migrants took a Pacific coastal route and began spreading to the American continent before the opening of the ice-free corridor. We investigated the diversity of full-length mitochondrial DNA genomes of the Zapotec population, residing in the Mesoamerican region, and reconstructed their demographic history using Bayesian Skyline Plots. We estimated the initial date of gene flow into the New World by Zapotec ancestors at around 17 000–19 000 years ago, which is highly concordant with previous studies. We also show a population decline after the initial expansion. This decline started 4000 years ago, long before European contact with Native Americans. This indicates that other factors including climate change should be considered to explain the observed demographic pattern.

*Journal of Human Genetics* (2015) 60, 557–559; doi:10.1038/jhg.2015.55; published online 21 May 2015

## INTRODUCTION

It is widely accepted that anatomically modern humans (*Homo sapiens*) from Asia colonized the New World through Beringia during the ice age.<sup>1</sup> Many researchers in archaeology, linguistics and evolutionary genetics have investigated the route, timing and size of the first migration event to America.<sup>2,3</sup> The single migration hypothesis was proposed based mainly on the analysis of uniparental genetic markers of current Native Americans.<sup>1,4</sup> Genome-wide single nucleotide polymorphism data for Native Americans suggest three streams of Asian gene flow, one major plus two additional minor gene flow events.<sup>5</sup> These data also suggest a coastal route for the initial peopling of America, in agreement with the results of mitochondrial DNA (mtDNA) analyses.<sup>4,5</sup>

The Mesoamerican region is adjacent to the Isthmus of Panama, a narrow strip of land that links North and South America. It is an important region for understanding the demography of Native Americans because the ancestors of South Americans presumably passed through this region during their migration from North America. In a previous study, we analyzed mtDNA haplotypes to estimate genetic diversity within the Mesoamerican region.<sup>6</sup> Here, we report a more detailed demographic history of the Zapotec inferred from Bayesian Skyline Plots (BSP) to gain better insight into their

genetic history. Of particular interest is whether a recent population decline corresponded with European contact, as some studies have proposed, or other factors such as climatic events.<sup>7–11</sup>

## MATERIALS AND METHODS

We used 88 complete Zapotec mtDNA genome sequences reported in the study by Mizuno *et al.*<sup>6</sup> The poly(C) tracts beginning at position 16 182 were not considered in this study. Information regarding data collection and data availability can be found in the study by Mizuno *et al.*

To obtain BSP, we used BEAST v1.7 suite.<sup>12</sup> We applied the Tamura and Nei TN93 mutation model using gamma-distributed rates and a proportion of invariant sites. A clock model was used assuming a log-normal distribution. In the Skyline model, we assumed a piecewise-constant model with 10 groups. We set the Markov Chain Monte Carlo chain length to  $1 \times 10^8$  with  $1 \times 10^7$  burn-in steps to collect sufficient samples for parameter estimation. We performed the Markov Chain Monte Carlo simulation twice independently to confirm that the simulation converged to the same state. We also estimated the times to the most recent common ancestor (tMRCA) by this Markov Chain Monte Carlo method for haplogroups A2, B2, C1 and D1 using Beaufi.<sup>12</sup> The time was scaled by the number of mutations for both BSP construction and the estimation of tMRCA for each haplogroup. To estimate divergence time based on the number of mutations, we assumed a molecular clock of  $1.665 \times 10^{-8}$  and applied the correction method of Soares *et al.*<sup>13</sup> To scale the female effective

<sup>1</sup>School of Advanced Sciences, SOKENDAI (The Graduate University for Advanced Studies), Hayama, Japan; <sup>2</sup>Department of Biological Sciences, Graduate School of Science, the University of Tokyo, Tokyo, Japan; <sup>3</sup>School of Medicine, Hangzhou Normal University, Zhejiang, China; <sup>4</sup>Department of Transplants, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico and <sup>5</sup>Molecular Genetics Laboratory, National Institute of Anthropology and History, Mexico City, Mexico

Correspondence: Dr J Gojobori, School of Advanced Sciences, SOKENDAI (The Graduate University for Advanced Studies), Kamiyamaguchi 1560-35, Hayama, Kanagawa 240-0193, Japan.

E-mail: gojobori\_jun@soken.ac.jp

or Professor L Wang, School of Medicine, Hangzhou Normal University, Zhejiang 311121, China.

E-mail: liwang@hznu.edu.cn

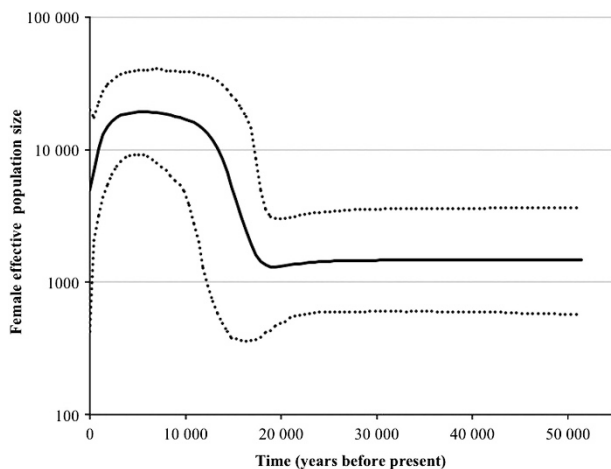
Received 20 November 2014; revised 19 April 2015; accepted 25 April 2015; published online 21 May 2015

population size ( $N_{ef}$ ), we assumed a generation time of 25 years.<sup>14</sup> The simulation results were analyzed using Tracer v1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>). For estimating tMRCAs, we also used another common method,  $\rho$  statistics, to estimate the tMRCAs of the four haplogroups using median joining calculation implemented in Network (<http://www.fluxus-engineering.com>).<sup>15</sup>

## RESULTS

Zapotec mtDNA was classified as haplogroup A2, B2, C1 or D1, and these groups included 59, 16, 10 and 3 sequences, respectively. In total, the data included 183 segregating sites ( $S$ ) and were characterized by a nucleotide diversity per site ( $\pi$ ) of 0.002850. Tajima's  $D$  was  $-1.464$ , and was significantly different from zero ( $P=0.036$ ).<sup>16</sup> This excess of low frequency polymorphisms indicates that the Zapotec experienced a past population expansion. However, these summary statistics enable only a rough estimation of the timing and magnitude of demographic events.

To obtain a more detailed demographic history of the Zapotec, we constructed a BSP using BEAST v1.7 suite.<sup>12</sup> BSP showed that the ancestors of the Zapotec experienced a population expansion around 17 000–19 000 years ago (Figure 1). The  $N_{ef}$  was approximately 1500 and reached approximately 20 000 after the expansion. This observation is highly concordant with previous result;<sup>4</sup> if we assume that this time of population expansion corresponds to the peopling of America, we can infer that the ancestors of the Zapotec migrated



**Figure 1** Bayesian Skyline Plot (BSP) of the Zapotec.  $N_{ef}$  ( $y$  axis, log-scale) was plotted against time ( $x$  axis, years ago). A mutation rate of  $1.664 \times 10^{-8}$  per site per year and a generation time of 25 years were assumed.<sup>14,15</sup> Solid line shows the mean  $N_{ef}$  from the posterior distribution. Dashed lines show the 95% posterior density intervals.

from Beringia to the American continent at least 17 000–19 000 years ago. This time range predates the opening of the ice-free corridor between the Laurentide and Cordilleran ice sheets, estimated to have occurred roughly 12 000 years ago.<sup>4,17</sup> This suggests that the Zapotec ancestors migrated *via* a coastal route rather than the ice-free corridor.<sup>4,5</sup> The BSP estimated from the mtDNA coding regions and the full-length sequences showed the same results (data not shown).

Table 1 shows the estimated tMRCA for the four haplogroups determined using BEAST.<sup>12</sup> Haplogroup A2 showed the oldest tMRCA, while Haplogroup D1 showed the youngest tMRCA. The tMRCAs of Haplogroups B2 and C1 were intermediate. These results are not concordant with a previous study.<sup>4</sup> Because the Zapotec people are considered to be the descendants of the first major migrants to the New World,<sup>5</sup> overlap between the tMRCAs of the four haplogroups was expected.<sup>4</sup> To verify our results, we estimated tMRCA for the four haplogroups using  $\rho$  statistics.<sup>15</sup> Table 1 shows the estimates of  $\rho$  and the associated dates inferred from the statistic, using the molecular clock of Soares *et al.*<sup>13</sup> The estimates based on the  $\rho$  statistics were consistent with the results obtained using BEAST.

We carefully examined very recent demographic events estimated from BSP to gain insights into the population decline of Native Americans. The BSP of the Zapotec shows a population decline after an approximately 10-fold population expansion (Figure 1).  $N_{ef}$  decreased to around 5000 from a peak of around 20 000. If we assume that this population decline was caused by the first contact with Europeans, the time of this decline should correspond to the time of the contact event that occurred 500 years ago.<sup>8</sup> However, our BSP showed that the population decline started approximately 4000 years ago, which is 3500 years earlier than the arrival of the first Europeans to the New World.

## DISCUSSION

The tMRCAs estimated for the four haplogroups showed different result from previous study<sup>4</sup> (Table 1). The sample size of Haplogroup D1 was only three; therefore, it is quite possible that the tMRCA of D1 was underestimated. The sample sizes of Haplogroups B2 and C1 were also small (16 and 10, respectively), which might also result in an underestimate of the mutations and therefore can lead to more recent coalescent time estimates.

Our BSP showed a population decline started approximately 4000 years ago (Figure 1). On the other hand, O'Fallon and Fehren-Schmitz<sup>8</sup> detected a population decline of Native Americans that occurred 500 years ago, which coincides with the first European contact. Unlike our analysis, O'Fallon and Fehren-Schmitz used extended BSP. We performed extended BSP with our Zapotec data, applying the same parameters used in the study by O'Fallon and

**Table 1** Estimated tMRCAs in years for Haplogroups A2, B2, C1 and D1 from the Zapotec using BEAST and  $\rho$  statistics

| Haplogroups | BEAST                     | $\rho$ statistics |                           |
|-------------|---------------------------|-------------------|---------------------------|
|             | tMRCA in years (95% c.i.) | $\rho$ (s.d.)     | tMRCA in years (95% c.i.) |
| A2          | 23126.9 (16888.6–30542.3) | 8.31 (1.54)       | 24432.6 (15727.1–33472.3) |
| B2          | 17769.5 (12896.4–23625.8) | 7.75 (1.67)       | 21168.5 (11908.1–30833.5) |
| C1          | 18048.6 (12457.8–24820.1) | 6.90 (1.52)       | 18726.7 (10393.8–27402.3) |
| D1          | 15137.7 (8636.2–20663.4)  | 5.00 (1.37)       | 13370.5 (6030.8–21006.2)  |

Abbreviations: c.i., confidential interval; s.d., standard deviation; tMRCA, times to the most recent common ancestor.

Fehren-Schmitz, and detected a population decline starting 5000–6000 years ago. When we applied the molecular clock of Soares *et al.*<sup>13</sup> in this extended BSP analysis, the time of population decline was estimated at around 4000 years ago, which matches our BSP results (data not shown). Therefore, the difference between our results and those of O’Fallon and Fehren-Schmitz was not due to methodological differences. We assume that the difference may reflect the different samples used in the studies.

In the Mesoamerican area, agriculture began more than 5000 years ago<sup>18</sup> and the growth of civilizations occurred afterward.<sup>19</sup> Steady population expansion is generally expected after the introduction of agriculture, but such steady growth was not observed in Europe.<sup>20</sup> Our study is the first to uncover unexpected demographic patterns in a population of Mesoamericans after the introduction of agriculture. During the Holocene, millennial or multi-century periodical cycles of climate change have been suggested.<sup>9,10</sup> The timing of these cycles is related to Bond event no. 3 of the Holocene Bond cycles.<sup>10</sup> An abrupt climate change occurred in Mexico around 4000 years ago.<sup>11</sup> We concluded that the  $N_{ef}$  of the ancestors of the Zapotec population decreased around 4000 years ago, long before European contact, possibly owing to severe climate change.<sup>9–11</sup>

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

We thank the volunteers in Zapotec who agreed to provide their DNA. We also thank Prof. Saburo Sugiyama and Dr John A. Eimes for the valuable comments on the manuscript. We thank three anonymous reviewers for providing comments that improved the manuscript. This study was supported by State Key Development Program for Basic Research of China, 973 Program (No. 2014CB541701), Zhejiang Provincial Natural Science Foundation of China Grant (No. LZ13H02001) and “Qian Jiang Distinguished Professor” program (to L.W.).

- 1 Fagan, B. M. *The Great Journey: The Peopling of Ancient America*. London (University Press of Florida, Gainesville, FL, USA, 2004).
- 2 Perez, S. I., Bernal, V., Gonzalez, P. N., Sardi, M. & Politis, G. G. Discrepancy between cranial and DNA data of early Americans: implications for American peopling. *PLoS ONE* **4**, e5746 (2009).
- 3 Goebel, T., Waters, M. R. & O’Rourke, D. H. The late Pleistocene dispersal of modern humans in the Americas. *Science* **319**, 1497–1502 (2008).
- 4 Fagundes, N. J. R., Kantiz, R., Eckert, R., Valls, A. C. S., Bogo, M. R., Salzano, F. M. *et al.* Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am. J. Hum. Genet.* **82**, 583–592 (2008).
- 5 Reich, D., Patterson, N., Campell, D., Tandon, A., Mazieres, S., Ray, N. *et al.* Reconstructing Native American population history. *Nature* **488**, 370–374 (2012).
- 6 Mizuno, F., Gojobori, J., Wang, L., Onishi, K., Sugiyama, S., Granados, J. *et al.* Complete mitogenome analysis of indigenous populations in Mexico: its relevance for the origin of Mesoamericans. *J. Hum. Genet.* **59**, 359–367 (2014).
- 7 Cook, N. D. *Born to Die: Disease and New World Conquest, 1492–1650* (Cambridge University Press, Cambridge, UK, 1998).
- 8 O’Fallon, B. D. & Fehren-Schmitz, L. Native Americans experienced a strong population bottleneck coincident with European contact. *Proc. Natl Acad. Sci. USA* **108**, 20444–20448 (2011).
- 9 Denton, G. H. & Karlén, W. Holocene climatic variations—Their pattern and possible cause. *Quaternary Research* **3**, 155–205 (1973).
- 10 Bond, G. A pervasive millennial-scale cycle in North Atlantic Holocene and glacial climates. *Science* **278**, 1257–1266 (1997).
- 11 Wanner, H., Beer, J., Bütikofer, J. & Crowley, T. J. Mid- to late Holocene climate change: an overview. *Quaternary Science Reviews* **27**, 1791–1828 (2008).
- 12 Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).
- 13 Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A. *et al.* Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am. J. Hum. Genet.* **84**, 740–759 (2009).
- 14 Fenner, J. N. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.* **128**, 415–423 (2005).
- 15 Bandelt, H. J., Forster, P. & Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).
- 16 Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
- 17 Arnold, T. G. Radiocarbon dates from the ice-free corridor. *Radiocarbon* **44**, 437–454 (2006).
- 18 Brown, C. H. (2010) *Pre-Columbian Foodways* (Springer, New York, NY, USA, 2010).
- 19 Hendon, J. A. & Joyce, R. *A Mesoamerican Archaeology: Theory and Practice*, Blackwell, Malden, MA, USA, 2004).
- 20 Shennan, S., Downey, S. S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T. *et al.* Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nat. Commun.* **4**, 2486 (2013).