

# Estimating divergence times from DNA sequences - supplemental material

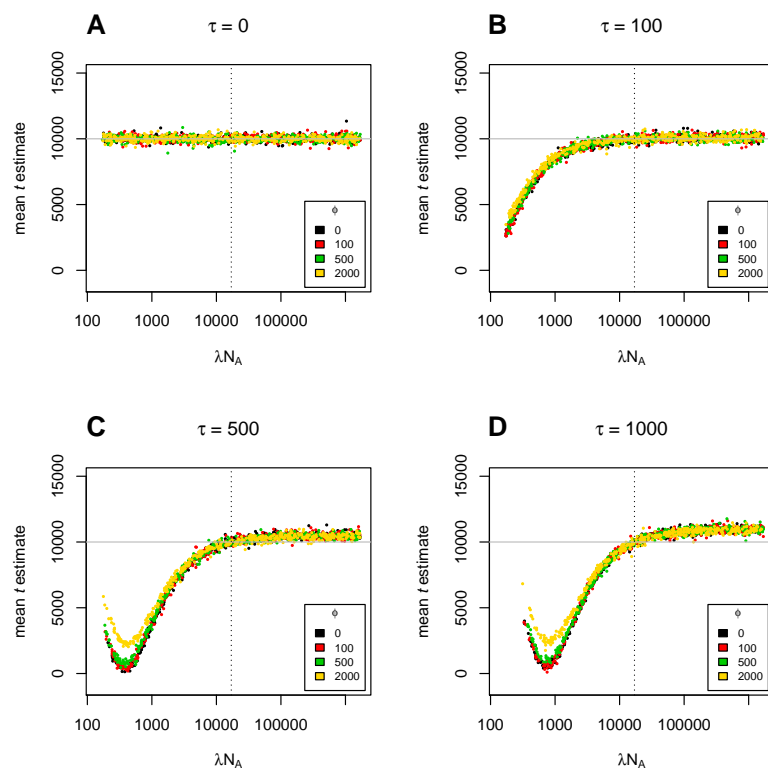
**Per Sjödin<sup>\*,1</sup>, James McKenna<sup>\*</sup> and Mattias Jakobsson<sup>\*,1,†</sup>**

<sup>\*</sup>Human Evolution, Department of Organismal Biology, Uppsala University, Uppsala, Sweden, <sup>†</sup>Science for Life Laboratory, Uppsala University, Uppsala, Sweden

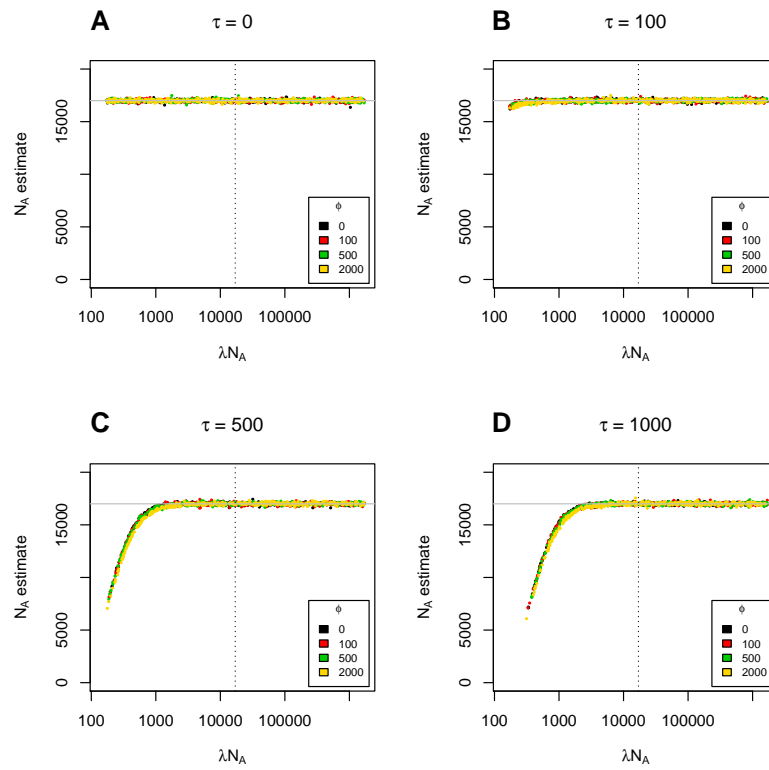
---

## Supplemental Material

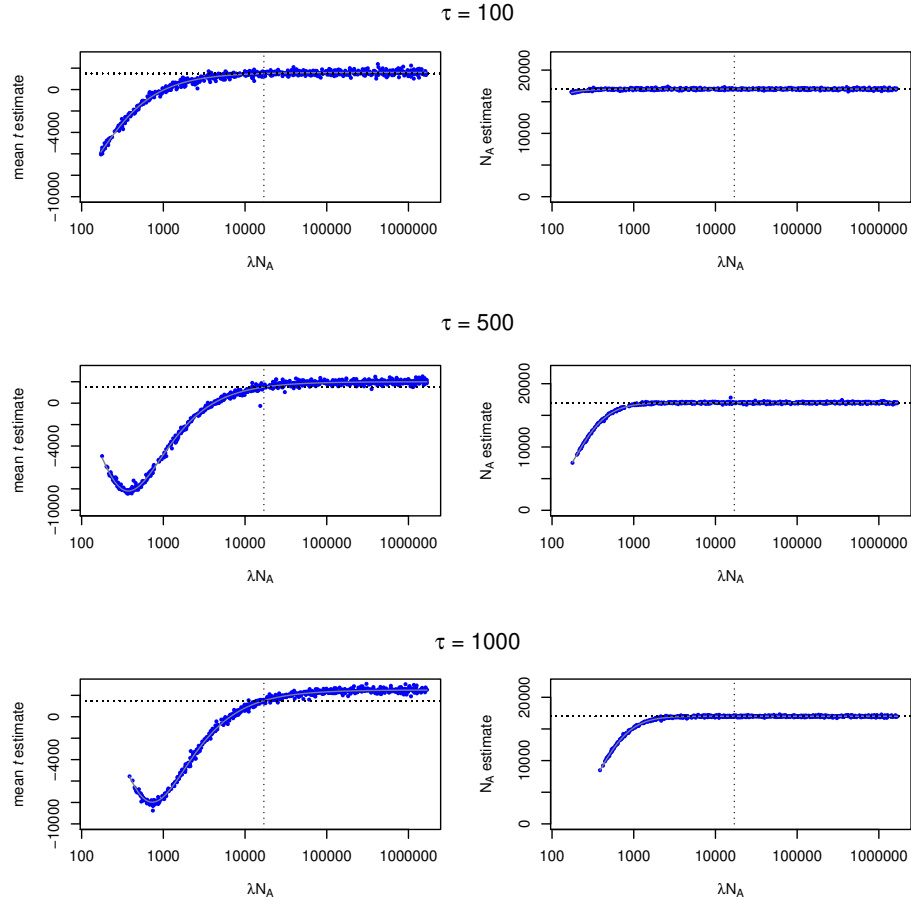
### Figures for simulation studies



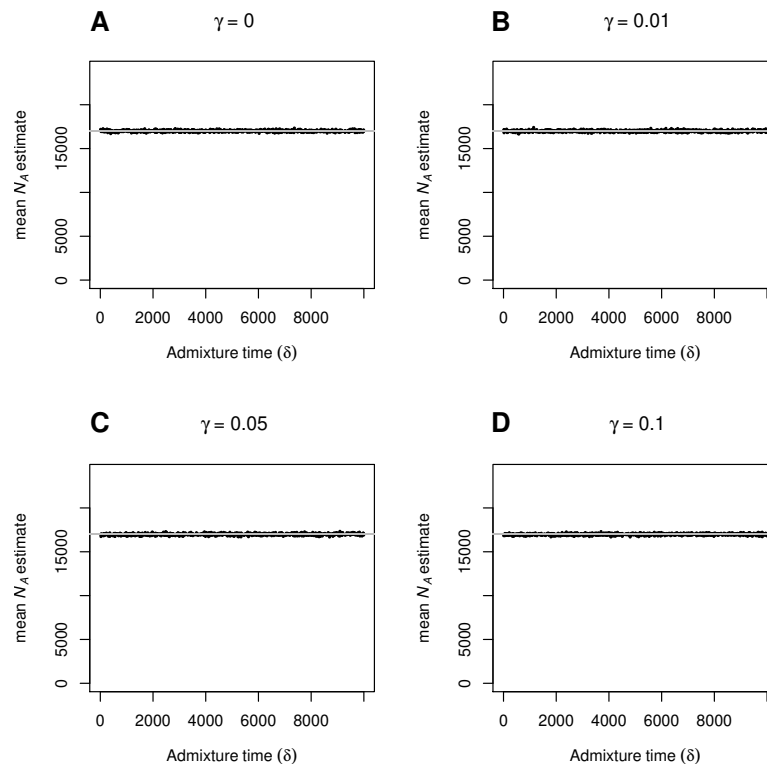
**Figure S1** The effect of ancestral population size change,  $\lambda N_A$ , on TT method estimates of divergence time,  $t$ , with a true split time of 10,000 generations and ancestral population size change lasting (A) 0, (B) 100, (C) 500, and (D) 1000 generations.



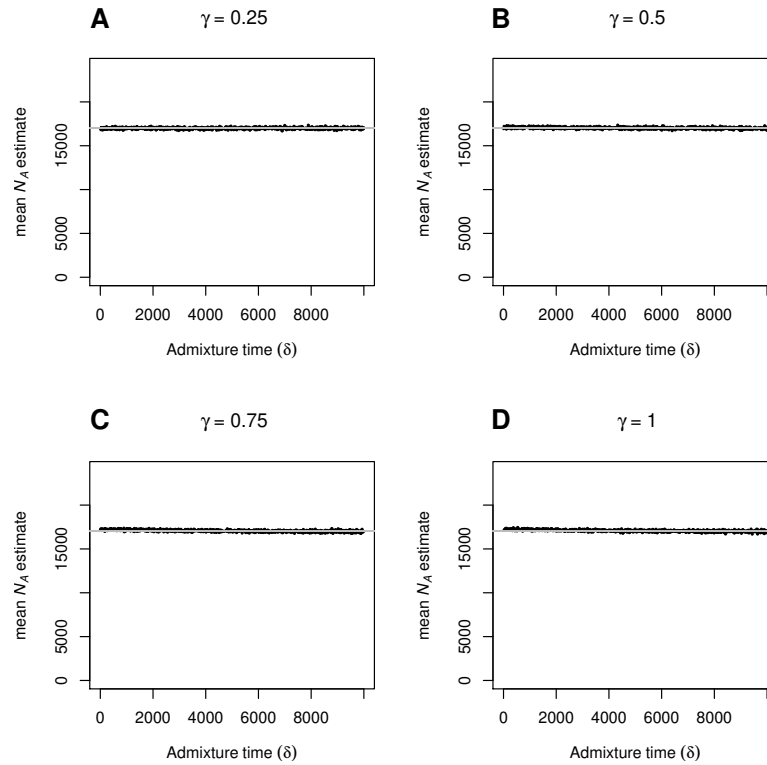
**Figure S2** The effect of ancestral population size change,  $\lambda N_A$ , on TT method estimates of ancestral population size,  $N_A$ , with a true split time of 10,000 generations and ancestral population size change lasting (A) 0, (B) 100, (C) 500, and (D) 1000 generations.



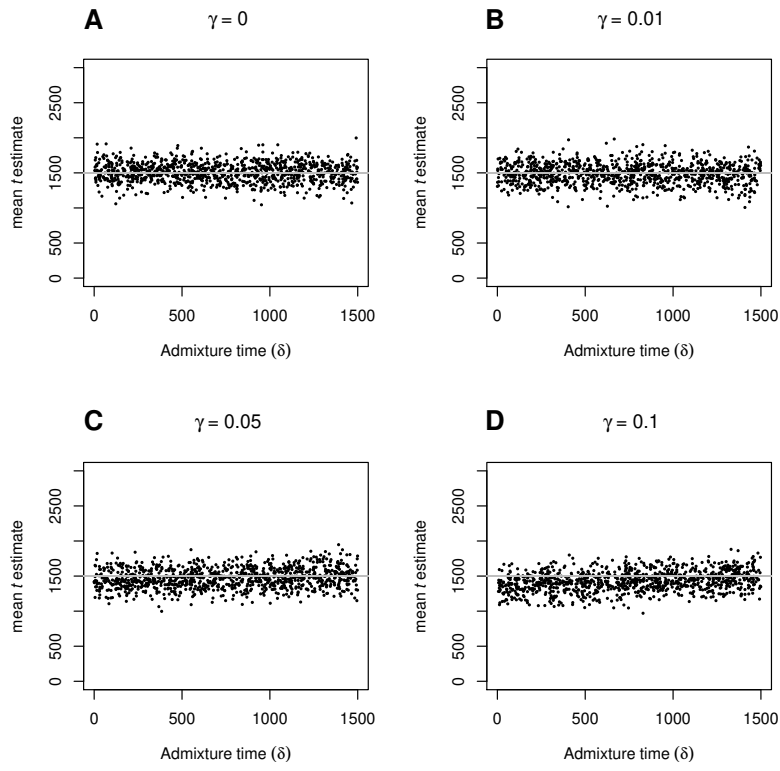
**Figure S3** TT method estimates of divergence time and ancestral population size with changes to ancestral population size of duration ( $\tau$ ) (A, B) 100, (C, D) 500, and (E, F) 1,000 generations. In all cases, the change in ancestral population size occurs immediately prior to the split ( $\phi=0$ ) and the true split time is 1500 generations.



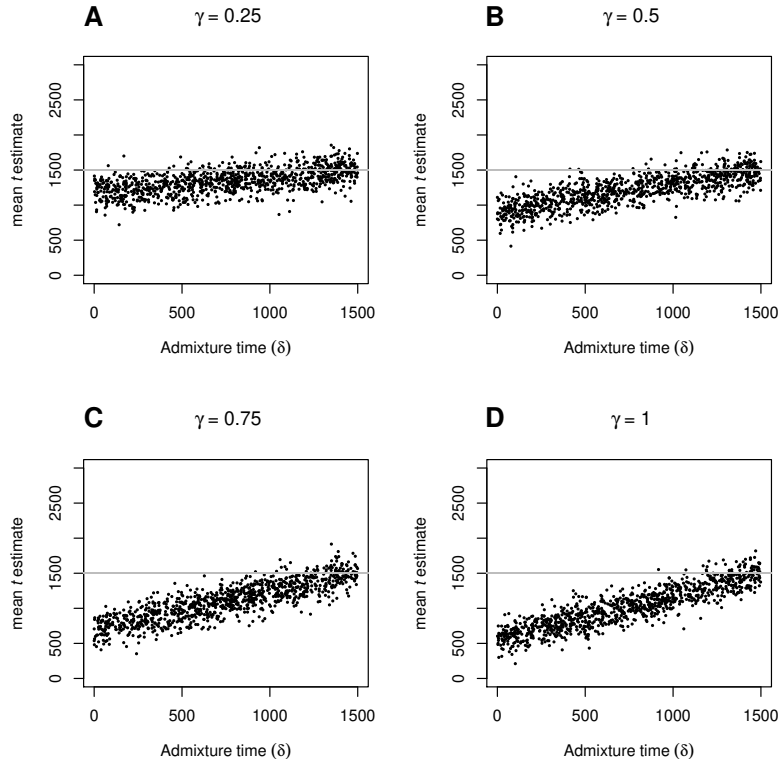
**Figure S4** The effect of varying admixture time ( $\delta$ ) on TT method estimates of ancestral population size ( $N_A$ ), with migration proportions ( $\gamma$ ) of (A) 0, (B) 0.01, (C) 0.05 and (D) 0.1 when true ancestral population size is 17,000 and true split time is 10,000 generations.



**Figure S5** The effect of varying admixture time ( $\delta$ ) on TT method estimates of ancestral population size ( $N_A$ ), with migration proportions ( $\gamma$ ) of (A) 0.25, (B) 0.5, (C) 0.75 and (D) 1 when true ancestral population size is 17,000 and true split time is 10,000 generations.



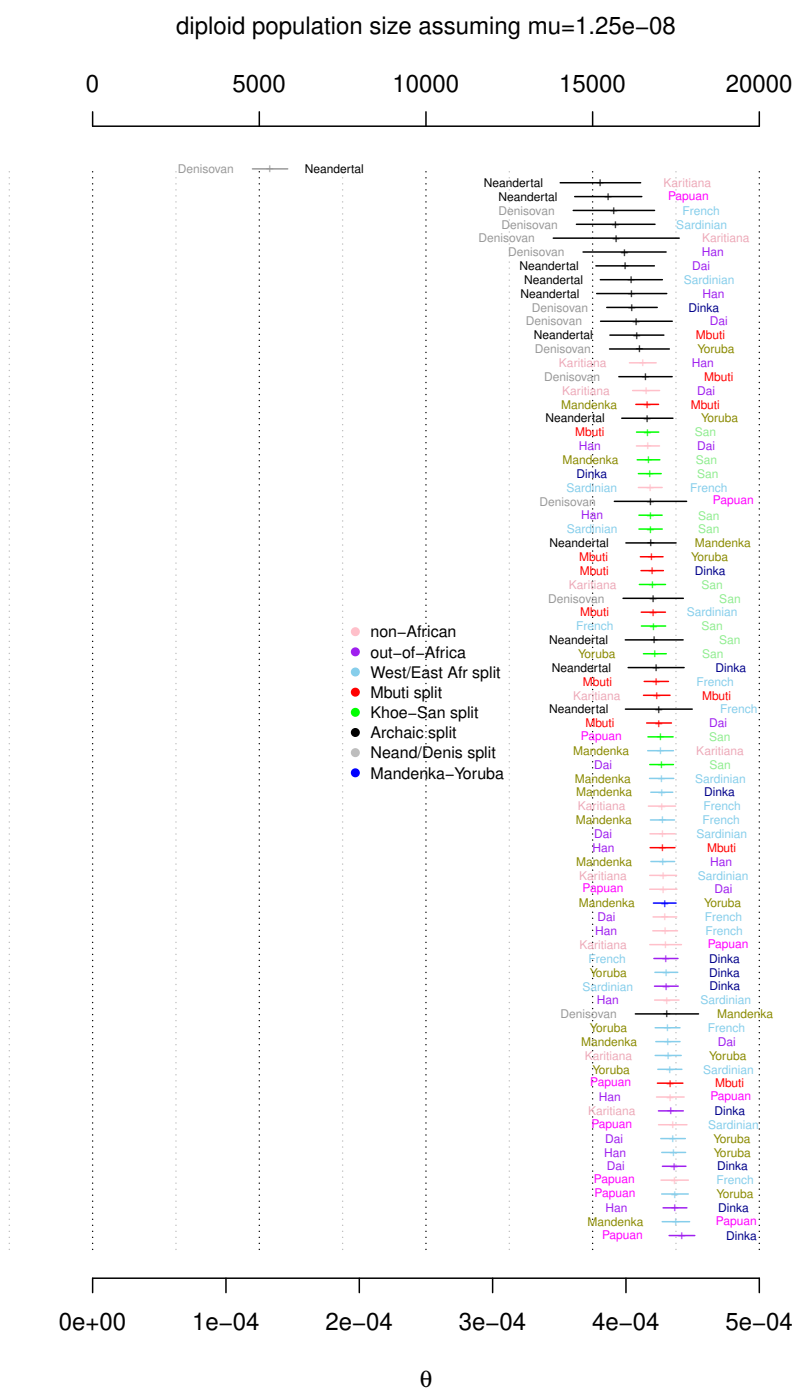
**Figure S6** The effect of varying admixture time ( $\delta$ ) on TT split time estimates ( $\hat{t}$ ), when the proportion of admixture ( $\gamma$ ) is (A) 0, (B) 0.01, (C) 0.05 and (D) 0.1, and true split time is 1500 generations.



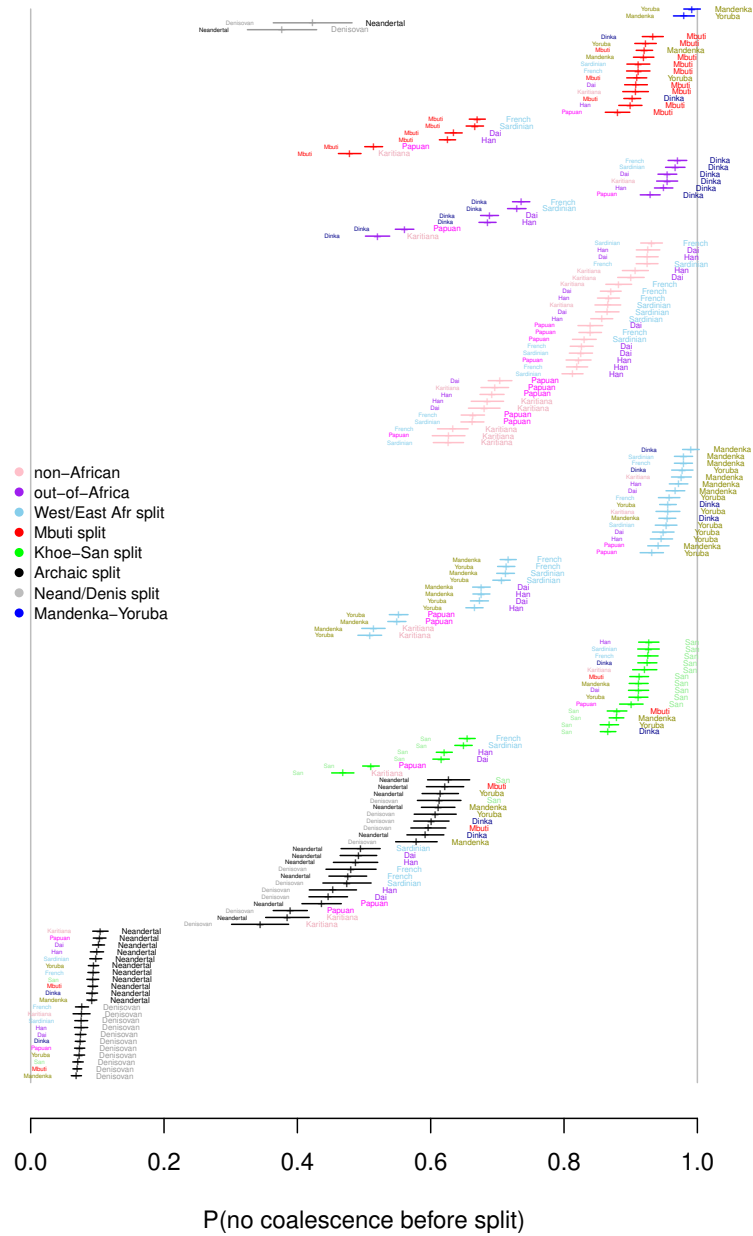
**Figure S7** The effect of varying admixture time ( $\delta$ ) on TT split time estimates ( $\hat{t}$ ), when the proportion of admixture ( $\gamma$ ) is (A) 0.25, (B) 0.5, (C) 0.75 and (D) 1, and true split time is 1500 generations.

## Figures for application to data

### TT-estimates under a constant ancestral population

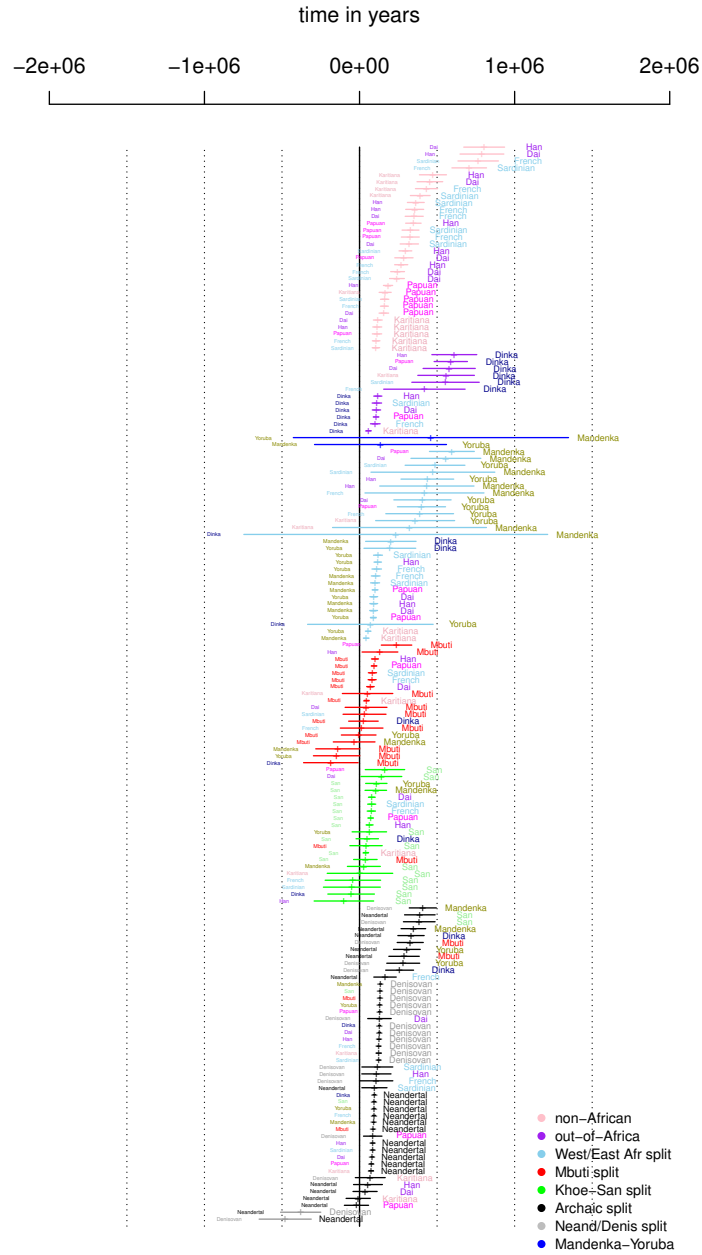


**Figure S8** Estimates of  $\theta = \mu N_A$  and the corresponding diploid ancestral population size ( $N_A/2$ ) assuming that it is constant and a mutation rate of  $1.25 \times 10^{-8}$ .

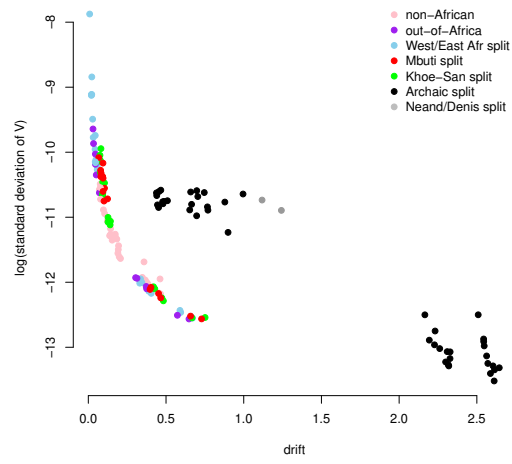


**Figure S9** Estimates of  $\alpha$  assuming a constant ancestral population.



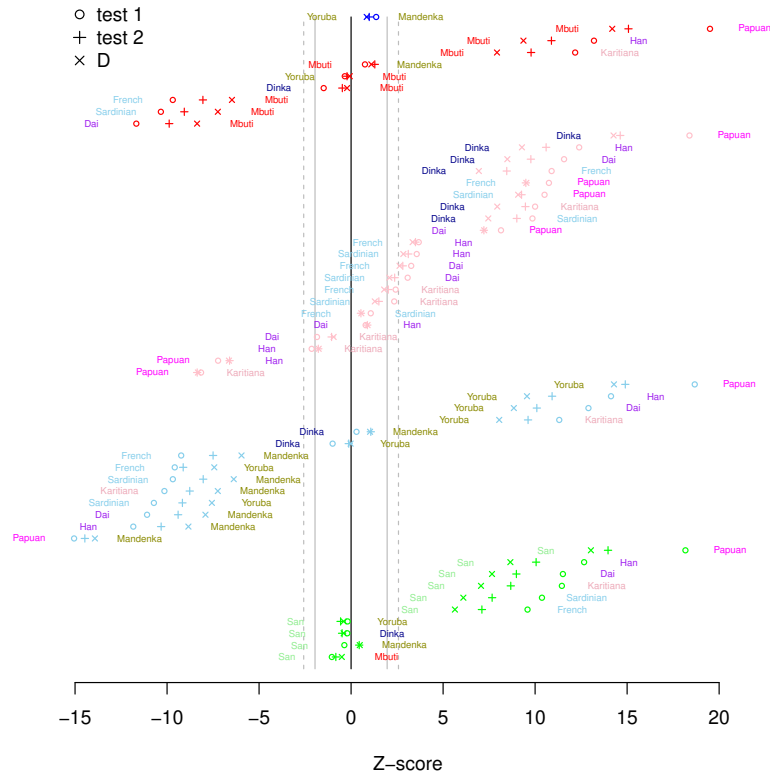


**Figure S10** Estimates of expected coalescence times in years given coalescence before split assuming a constant ancestral population. A mutation rate of  $1.25 \times 10^{-8}$  and a generation time of 30 years is assumed.

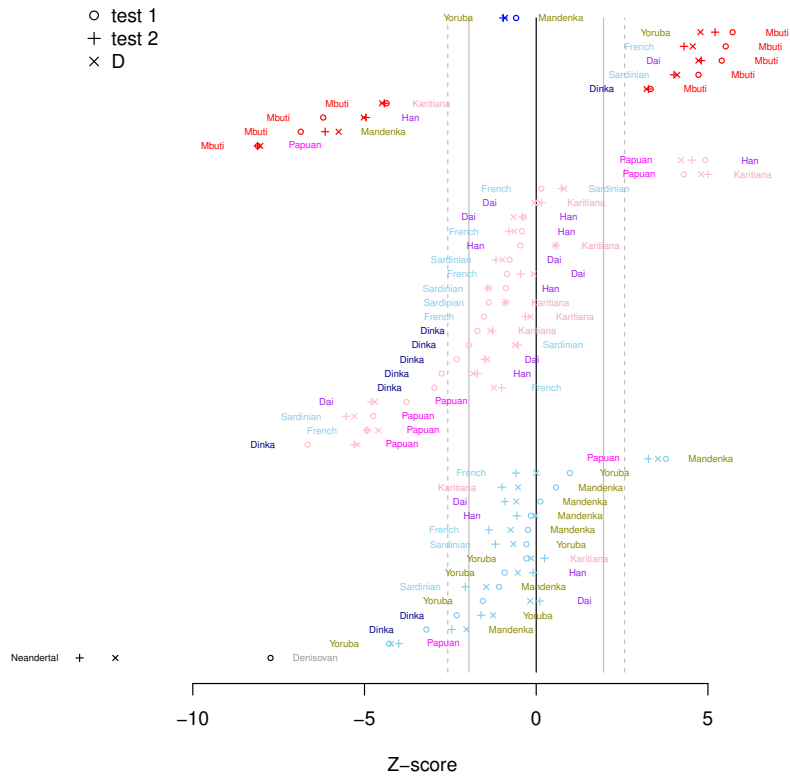


**Figure S11** Drift ( $-\ln(\alpha)$ ) vs estimated standard deviation of  $V = \mu\nu$ .

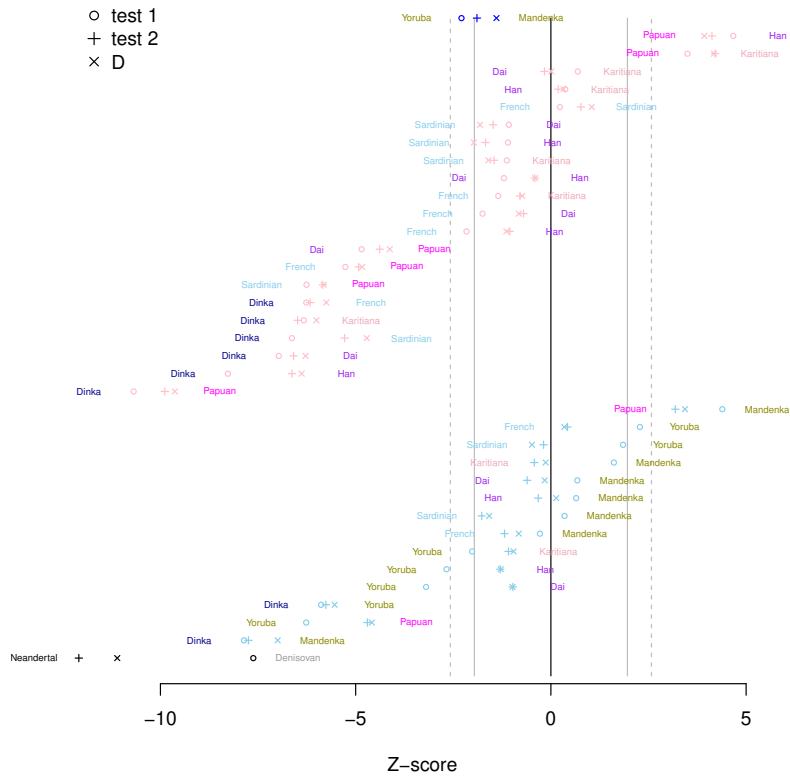
### *TTo-estimates using outgroup ascertainment*



**Figure S12** Tests based on equation 5 (○) and equation 6 (+) when data is conditional on the derived variant being present in either the Neandertal or Denisovan genome. D-tests (scaled by its estimated standard deviation, ×) with D('left population', 'right population', Neandertal+Denisovan, outgroup) are also shown. Solid grey lines shows cut-offs for rejecting null-model at the 0.05 level and the dashed lines at the 0.01 level.



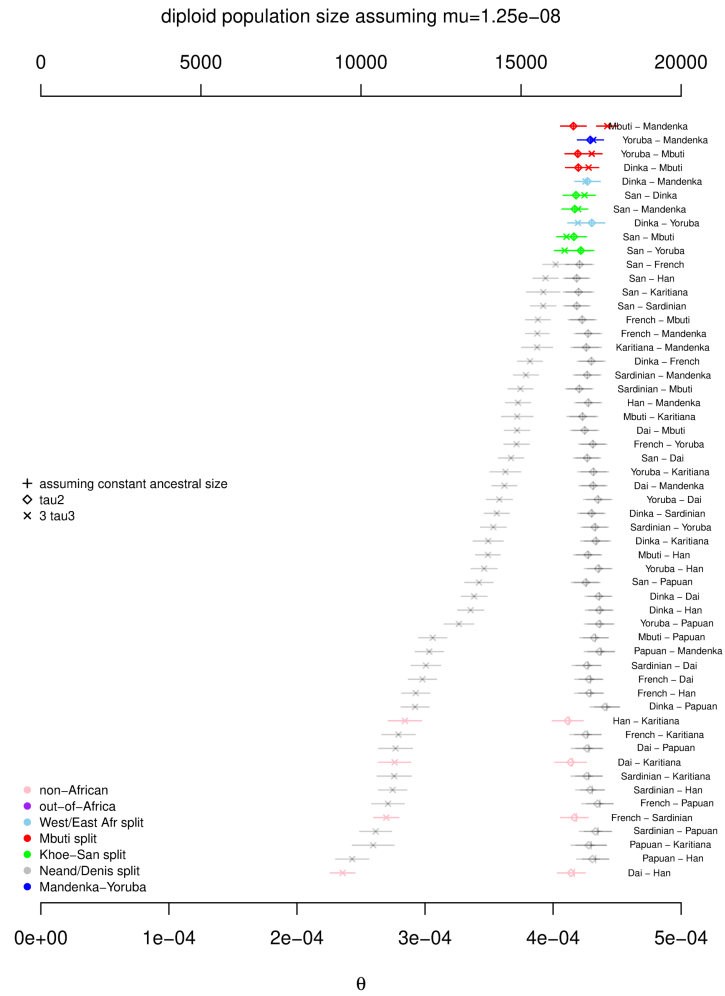
**Figure S13** Tests based on equation 5 (○) and equation 6 (+) when data is conditional on the derived variant being present in Balito Bay A. D-tests (scaled by its estimated standard deviation, ×) with D('left population', 'right population', Balito Bay A, outgroup) are also shown. Solid grey lines shows cut-offs for rejecting null-model at the 0.05 level and the dashed lines at the 0.01 level.



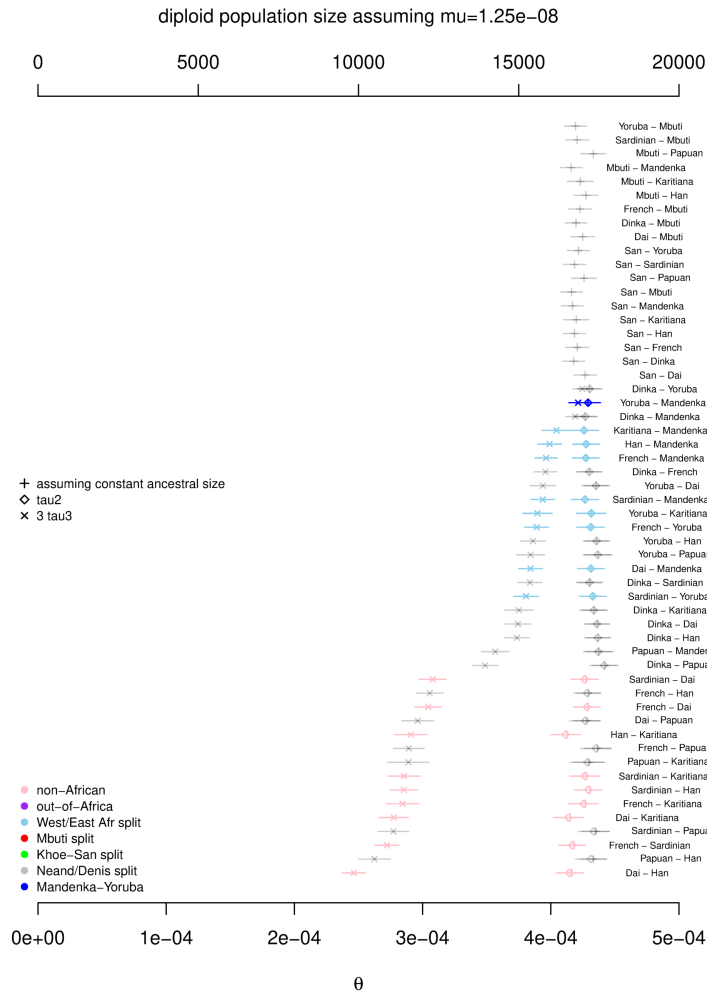
**Figure S14** Tests based on equation 5 (○) and equation 6 (+) when data is conditional on the derived variant being present in Mbuti. D-tests (scaled by its estimated standard deviation, ×) with D('left population', 'right population', Mbuti, outgroup) are also shown. Solid grey lines shows cut-offs for rejecting null-model at the 0.05 level and the dashed lines at the 0.01 level.



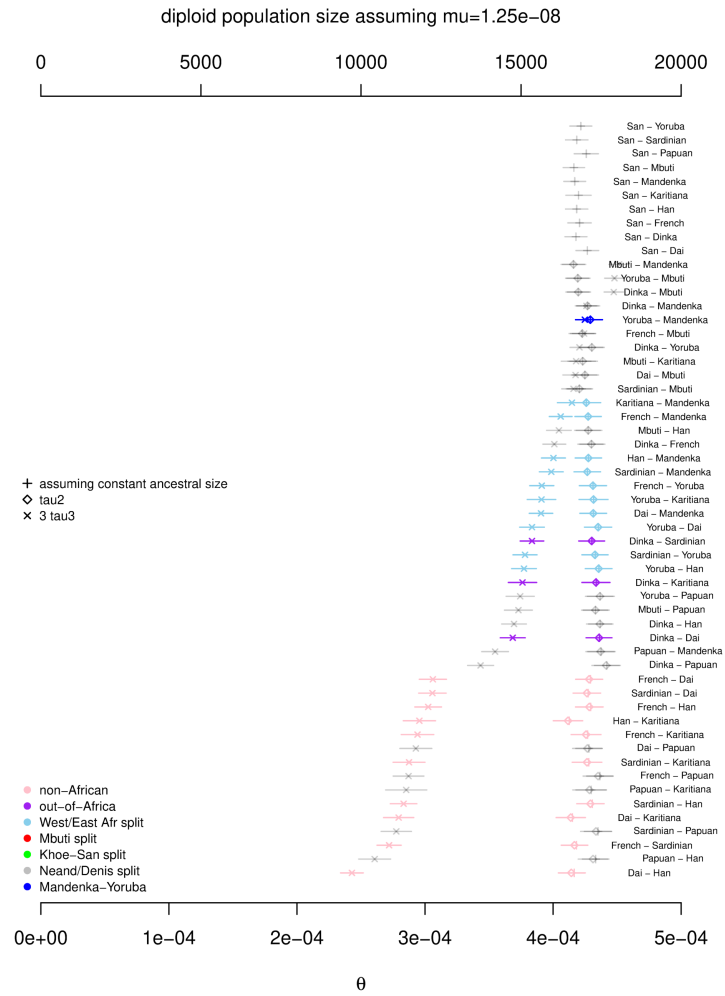
**Figure S15**  $\alpha$ -estimates conditional on the derived variant being present in the outgroup. ‘Cross’ show estimates when the derived variant is present in Mbuti, ‘circle’ show estimates when the derived variant is present in Neandertal or Denisovan, ‘plus’ show estimates when the derived variant is present in Balito Bay A. Black was used for estimates that did not fail the outgroup tests while transparent grey was used for comparisons that failed at least one of the two outgroup tests.



**Figure S16** Different estimates of the ancestral population size. Here Neandertal+Denisovan was used as outgroup to estimate  $\tau_2$  and  $\tau_3$ . Transparent grey was used for comparisons that failed at least one of the two outgroup tests.

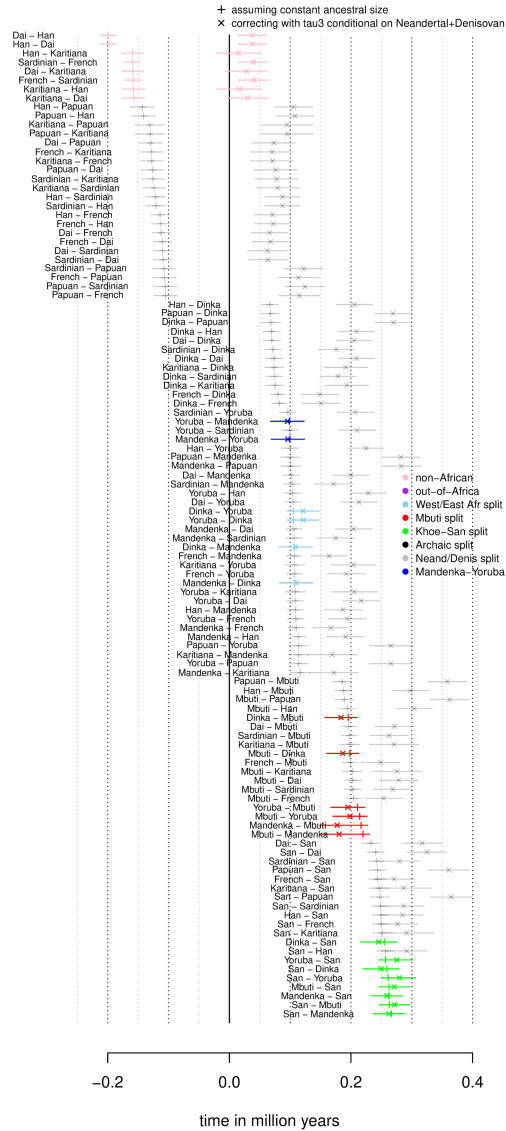


**Figure S17** Different estimates of the ancestral population size. Here Mbuti was used as outgroup to estimate  $\tau_2$  and  $\tau_3$ . Transparent grey was used for comparisons that failed at least one of the two outgroup tests.

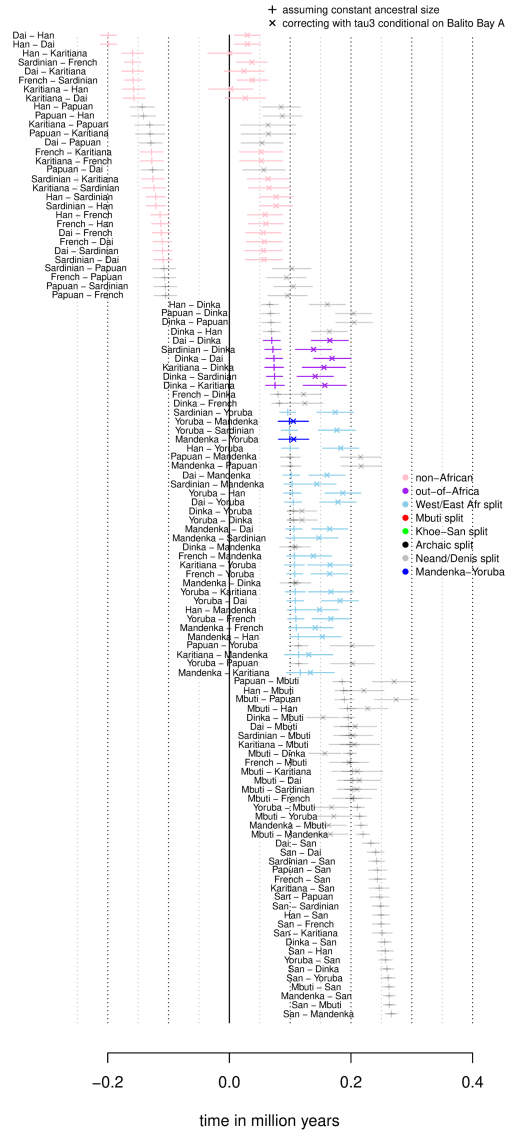


**Figure S18** Different estimates of the ancestral population size. Here Balito Bay A was used as outgroup to estimate  $\tau_2$  and  $\tau_3$ . Translucent grey was used for comparisons that failed at least one of the two outgroup tests.

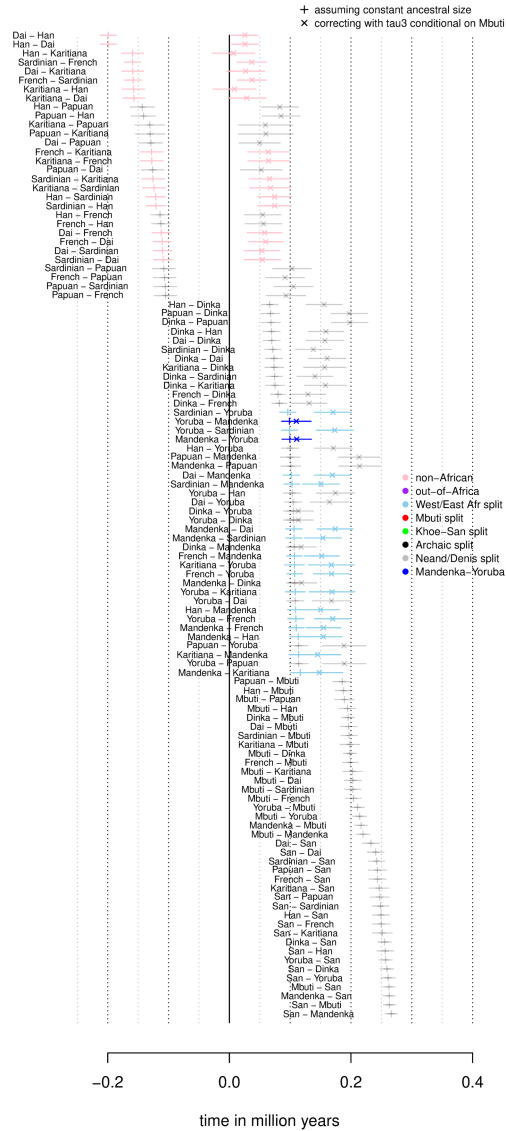




**Figure S19** Different estimates of split times assuming a mutation rate of  $1.25 \times 10^{-8}$  and a generation time of 30 years. Two estimates are shown: estimates assuming a constant ancestral (+) and estimates relying on an external estimate of  $\alpha$  (x) obtained by outgroup ascertainment in Neandertal+Denisovan. Transparent grey was used for comparisons that failed at least one of the two outgroup tests.



**Figure S20** Different estimates of split times assuming a mutation rate of  $1.25 \times 10^{-8}$  and a generation time of 30 years. Two estimates are shown: estimates assuming a constant ancestral (+) and estimates relying on an external estimate of  $\alpha$  (x) obtained by outgroup ascertainment in Balito Bay A (an ancient Khoe-San genome). Transparent grey was used for comparisons that failed at least one of the two outgroup tests.



**Figure S21** Different estimates of split times assuming a mutation rate of  $1.25 \times 10^{-8}$  and a generation time of 30 years. Two estimates are shown: estimates assuming a constant ancestral (+) and estimates relying on an external estimate of  $\alpha$  (x) obtained by outgroup ascertainment in Mbuti. Transparent grey was used for comparisons that failed at least one of the two outgroup tests.