

Supplementary Note – Method for calculating dates of events in *DAZ*'s evolutionary history

We estimated the dates of the chromosome 3 transposition event and the initial *DAZ* duplication event using human-rhesus substitution rates for calibration.

1. Calculation of Y-chromosome and autosome mutation rates (assume human-rhesus divergence time of 25 million years ago).

k = observed substitutions per site

k^* = corrected distance (using REV model as implemented in PAML [1,2])

μ = mutation rate

| | k | k^* | μ | reference |
|--------------|--------|--------|-----------------------|-----------|
| Y chromosome | 0.0939 | 0.118 | 2.36×10^{-9} | [3] |
| Autosomes | 0.0616 | 0.0727 | 1.45×10^{-9} | [4] |
| average | 0.0778 | 0.0953 | 1.91×10^{-9} | |

2. Dating chromosome 3 transposition event.

The time since transposition should be proportional to the divergence between the sequence on chromosome 3 and the homologous sequence on the rhesus MSY. We use human chromosome 3 sequence in this analysis because it is higher quality than the draft rhesus genome sequence. We exclude all coding sequences from the analysis. For the mutation rate, we use the Y chromosome-autosome average rate calculated in part 1.

From the alignment of human chromosome 3 sequence and rhesus MSY, we arrive at $k = 37342/335742 = 0.111$ subst/site; $k^* = 0.147$ subst/site.

Therefore, time = $k^* / (\mu \times 2) = 0.147 / (1.91 \times 10^{-9} \times 2) = 38.5$ million years.

3. Dating initial *DAZ* duplication event.

The time since the duplication event should be proportional to the divergence between the duplicated regions containing *DAZ1* and *DAZ2* on the rhesus MSY. We exclude all coding sequences from the analysis. The alignment is based on *DAZ1* and *DAZ2* intronic sequences as well as a duplicated 17.7-kb region downstream of the genes. We exclude the 30-kb upstream regions that are undergoing gene conversion. For the mutation rate, we use the Y chromosome rate calculated in part 1.

From the alignment of rhesus *DAZ1* and *DAZ2* duplicated regions, we arrive at $k = 2798/23707 = 0.118$ subst/site; $k^* = 0.157$ subst/site.

Therefore, time = $k^* / (\mu \times 2) = 0.157 / (2.36 \times 10^{-9} \times 2) = 33.3$ million years.

LITERATURE CITED

1. **Yang Z.** 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences* **13**:555-6.
2. **Yang Z.** 1994. Estimating the pattern of nucleotide substitution. *J Mol Evol* **39**:105-11.
3. **Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Graves T and others.** 2012. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature* **482**:82-6.
4. **Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM and others.** 2007. Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**:222-34.