

Supplemental Note S1: Inference of gene conversion from conserved palindromes in human, chimpanzee, and rhesus macaque

Using the logic of Hallast et al. 2013, we inferred gene conversion events from fixed nucleotide replacements in X palindrome arms that occurred in either human or chimpanzee, using rhesus macaque as an outgroup. For simplicity, the scenarios below describe a fixed replacement in the human lineage. We propose that fixed nucleotide replacements result from a substitution in humans after divergence from chimpanzee, followed by gene conversion that homogenizes the substitution between arms (Scenario 1). In theory, other scenarios could lead to the same result. In one alternative scenario, the ancestral palindrome was heterozygous at the site in question, with gene conversion occurring in one direction in rhesus macaque and chimpanzee, and the opposite direction in human (Scenario 2). We consider this scenario highly unlikely because it requires the initial site to remain heterozygous for 1.1 million generations before undergoing gene conversion in human and chimpanzee (see Figure 3A). Given our inferred intrachromosomal gene conversion rate of 4.5×10^{-5} events per nucleotide per generation, the probability of any given site not undergoing gene conversion over 1.1 million generations is $(1 - 4.5 \times 10^{-5})^{1100000}$, which is effectively zero ($< 2.22 \times 10^{-308}$).

We also considered a scenario in which the initial substitution occurred in the human-chimpanzee common ancestor, then underwent gene conversion in opposite directions in human and chimpanzee (Scenario 3). Given that X palindromes have on average only 1 difference between arms for every 2200 nucleotides, this scenario could explain at most observed nucleotide replacements in 1 out of 2200 positions in X palindrome arms (0.045%), if all heterozygous sites resolved in opposite directions in each lineage. We observed nucleotide replacements in 2567 out of 409,579 positions in X palindrome arms (0.65%), suggesting that Scenario 3 can account for no more than 7% (0.045% / 0.65%) of our observations.

Finally, we used evolutionary simulations with event tracing to estimate what fraction of fixed nucleotide replacements in human, chimpanzee, and rhesus macaque would arise through each scenario under reasonable evolutionary parameters (see Figure 3, Methods). We found that the vast majority of fixed nucleotide replacements (93.3%) arose through Scenario 1, while around 2.5% arose through Scenario 3. As predicted, we never observed fixed replacements arising from Scenario 2. The remaining fixed nucleotide replacements (4.2%) resulted from other scenarios that involved multiple substitution events. Importantly, our conclusions in Figure 3 are agnostic to the method by which each fixed nucleotide replacement arose, and depend only on the ability of a given set of evolutionary parameters to reproduce the replacement patterns seen in Figure 2B.

