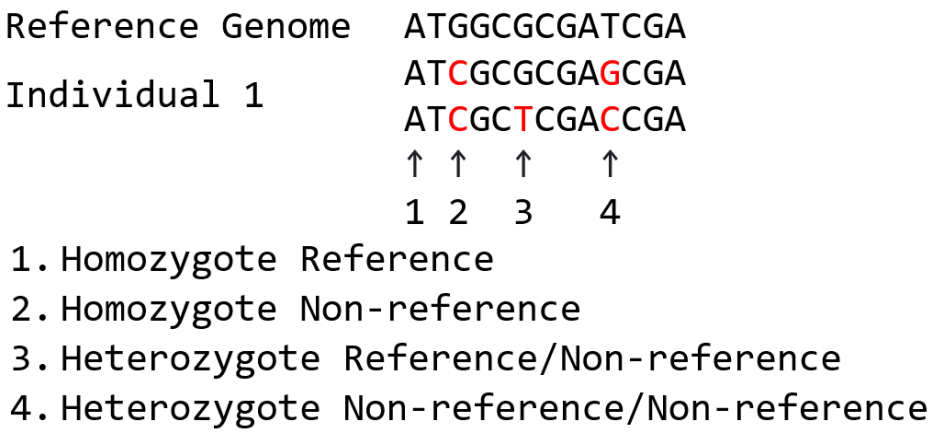
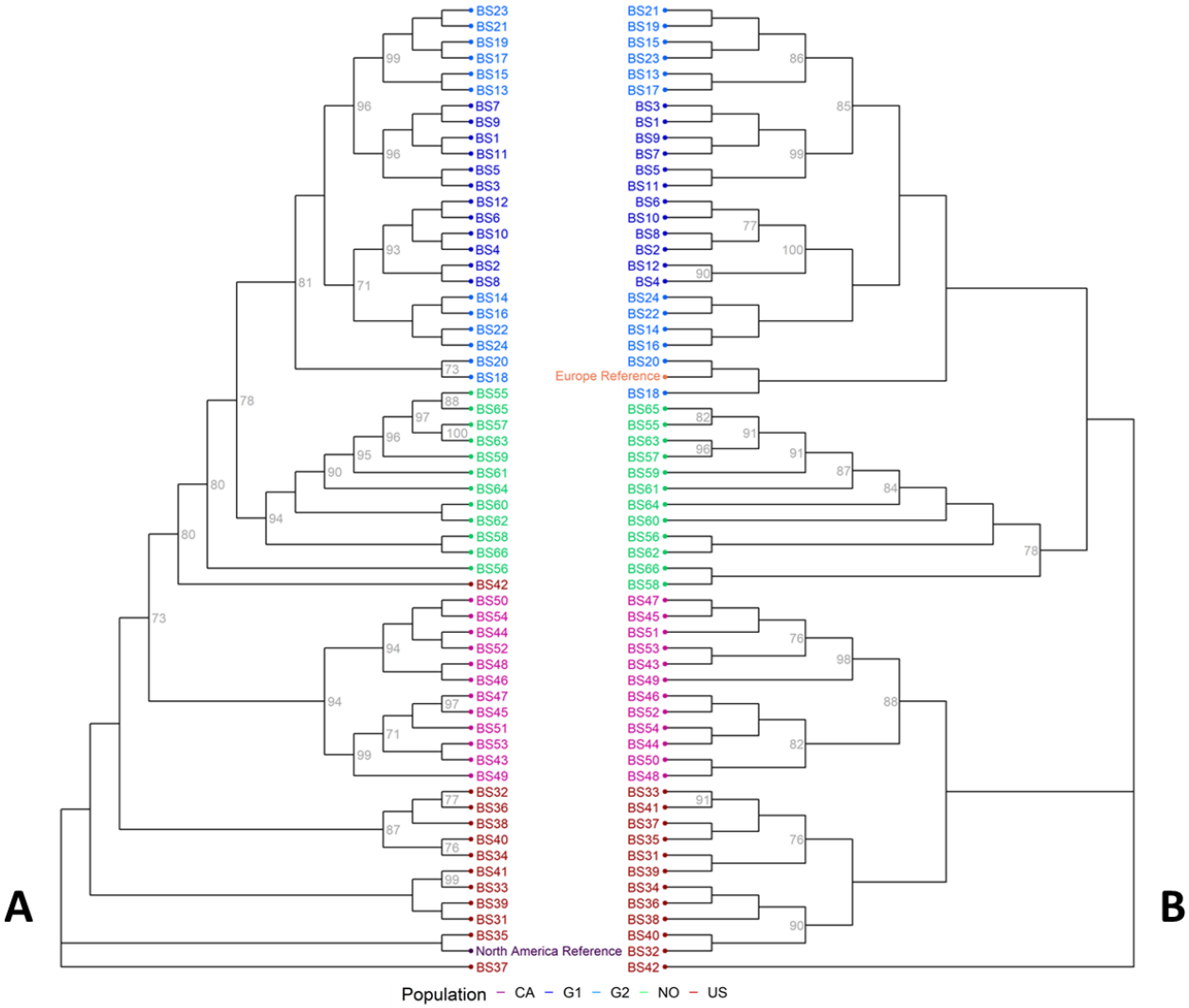
Supplementary Figures

Origin Matters: Using a Local Reference Genome Improves Measures in Population Genomics

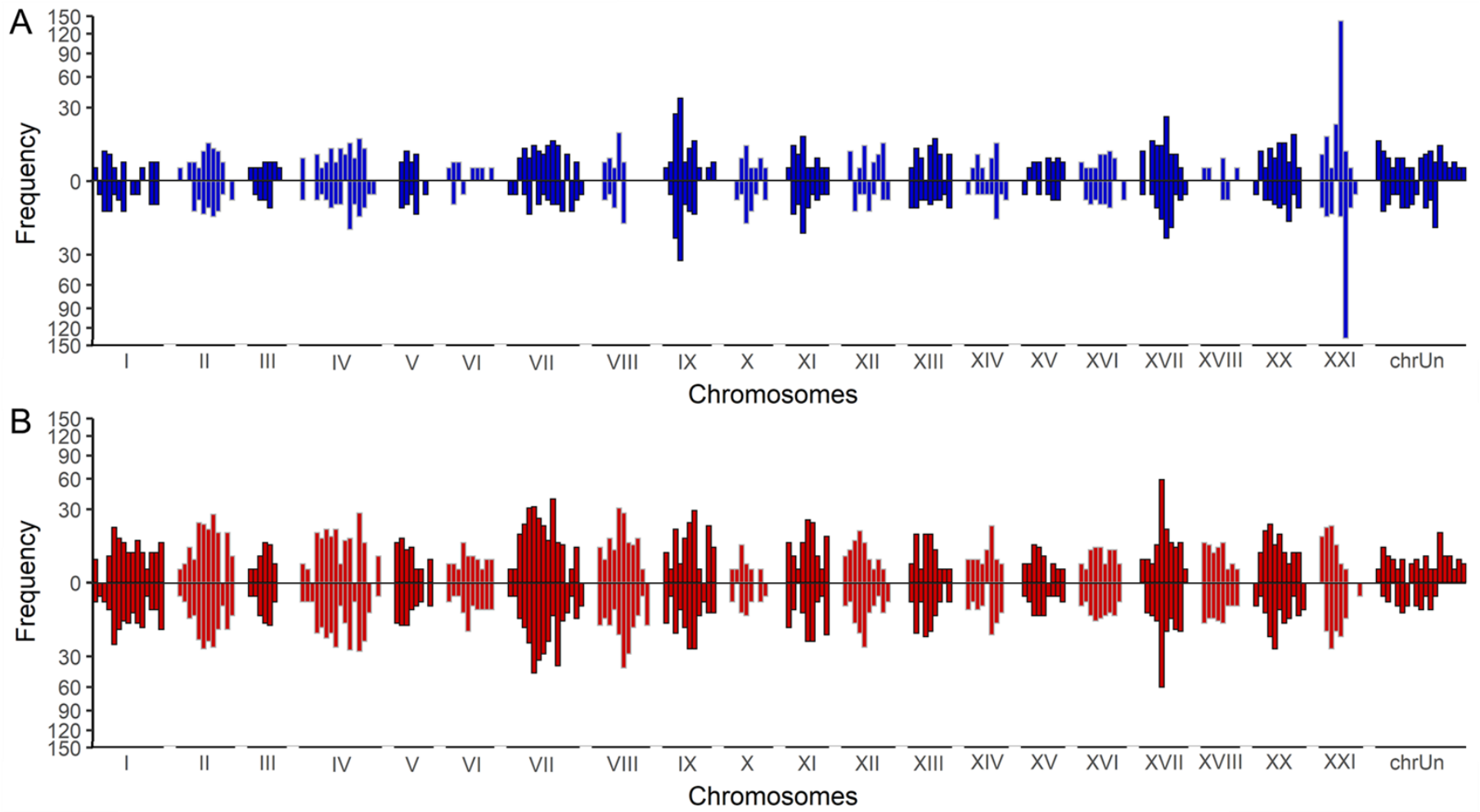
Doko-Miles J. Thorburn[1,2+], Kostas Sagonas[1,3], Mahesh Binzer-Panchal[4], Frederic J.J. Chain[5], Philine G.D. Feulner[6,7], Erich Bornberg-Bauer[8], Thorsten BH Reusch[9], Irene E. Samonte-Padilla[10], Manfred Milinski[10], Tobias L. Lenz[11,12], Christophe Eizaguirre[1]



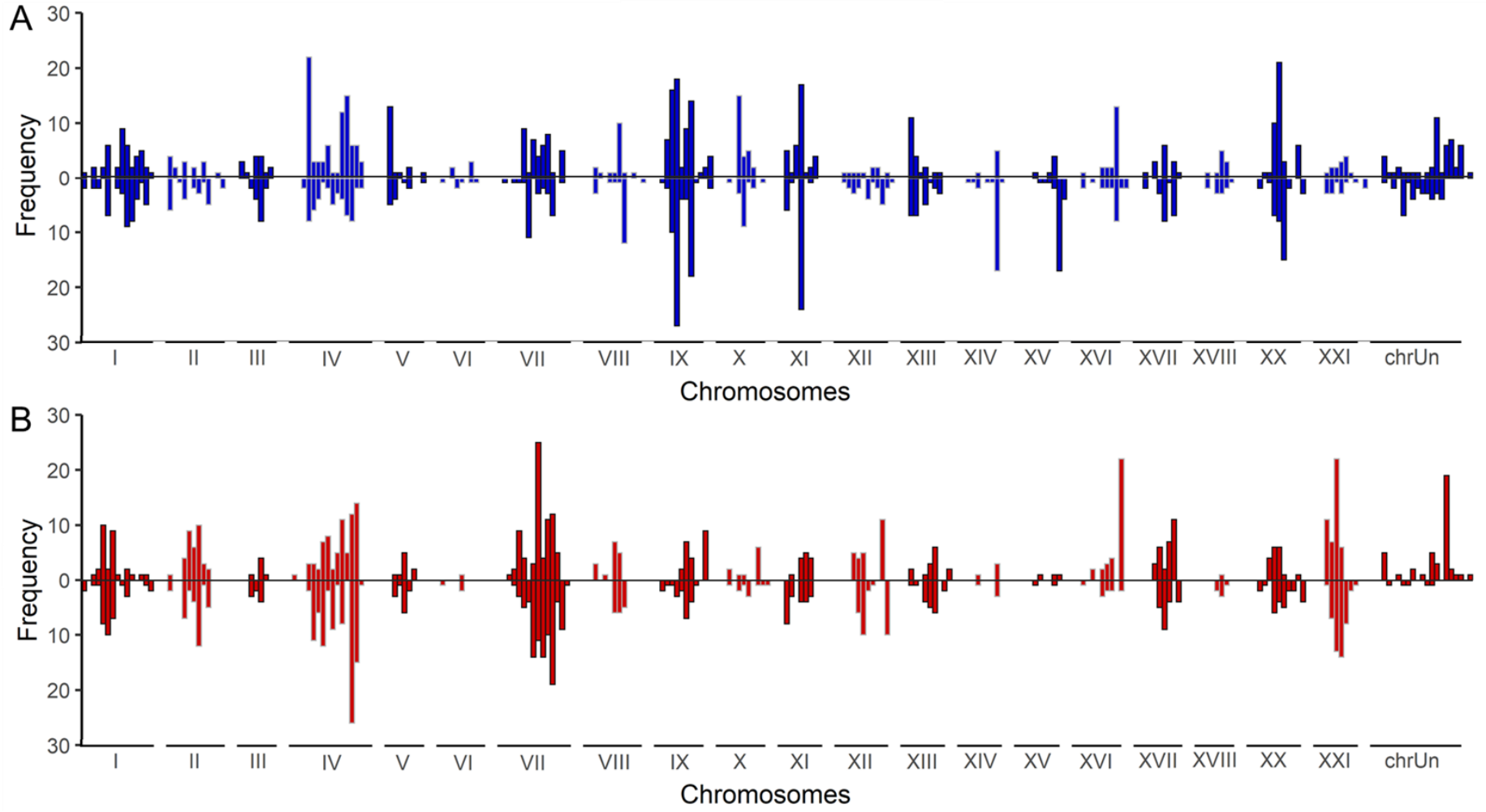
Supplementary figure S1. Theoretical representation of the genotype classes included in the genotype bias analysis for a diploid organism.



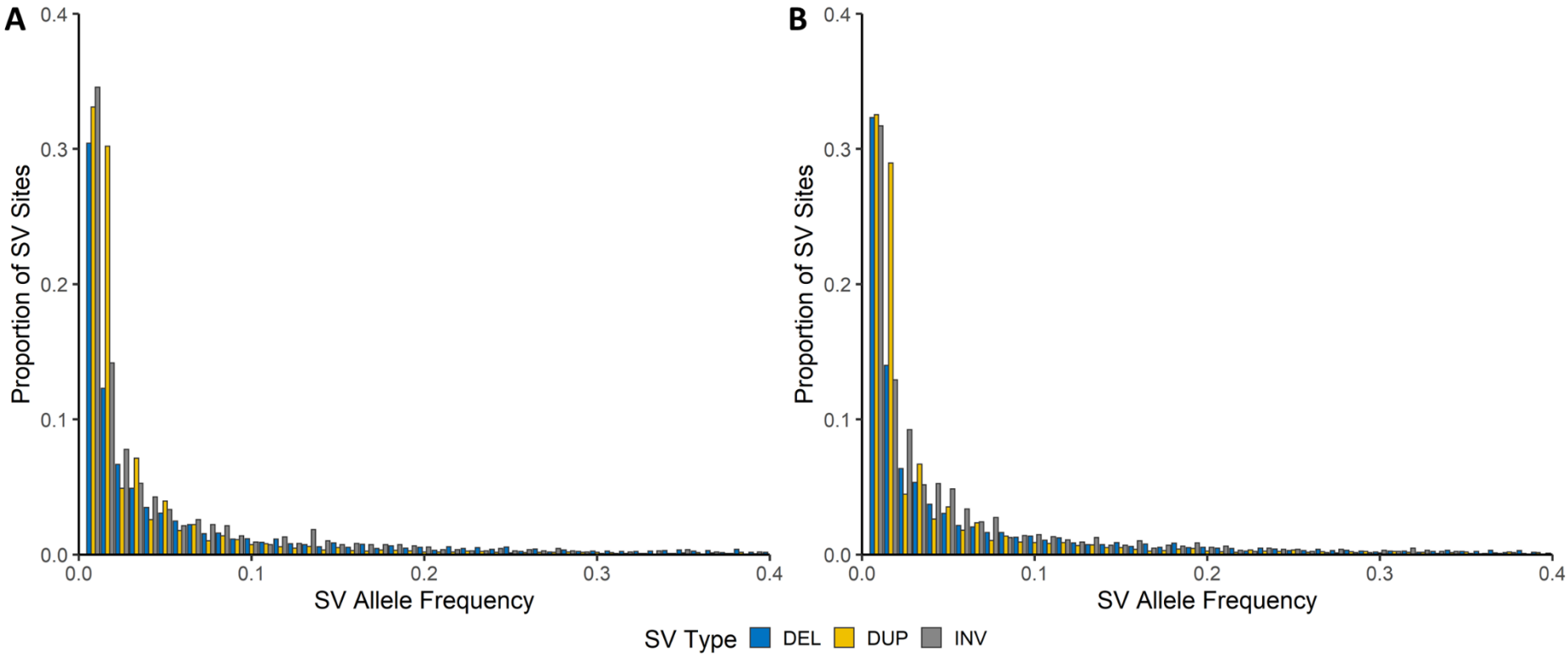
Supplementary Figure S2. Maximum likelihood tree based on polymorphism data called from [A] the North American reference genome and [B] the European reference genome. Both reference genomes were added to the phylogeny by randomly sampling 1% (~100k) of segregating sites across the genome. Nodes with higher than 70% bootstrap support are shown in grey at the node. Individuals are coloured by population-pair



Supplementary Figure S3. Comparing distributions of outlier Tajima's *D* windows across the genome for (top) European and (bottom) North American populations mapped to the (*A*) European or (*B*) North American reference genome. Axes are square root transformed.



Supplementary Figure S4. Comparing distributions of outlier F_{ST} windows across the genome for (top) European and (bottom) North American populations mapped to the (*A*) European or (*B*) North American reference genome. Axes are square root transformed.



Supplementary Figure S5. The allele frequency spectrum of bi-allelic SVs across all 60 individuals, exhibiting that the majority of deletions (blue), duplications (yellow) and inversions (grey) occur at low frequencies when using both the (*A*) European or the (*B*) North American reference genomes.