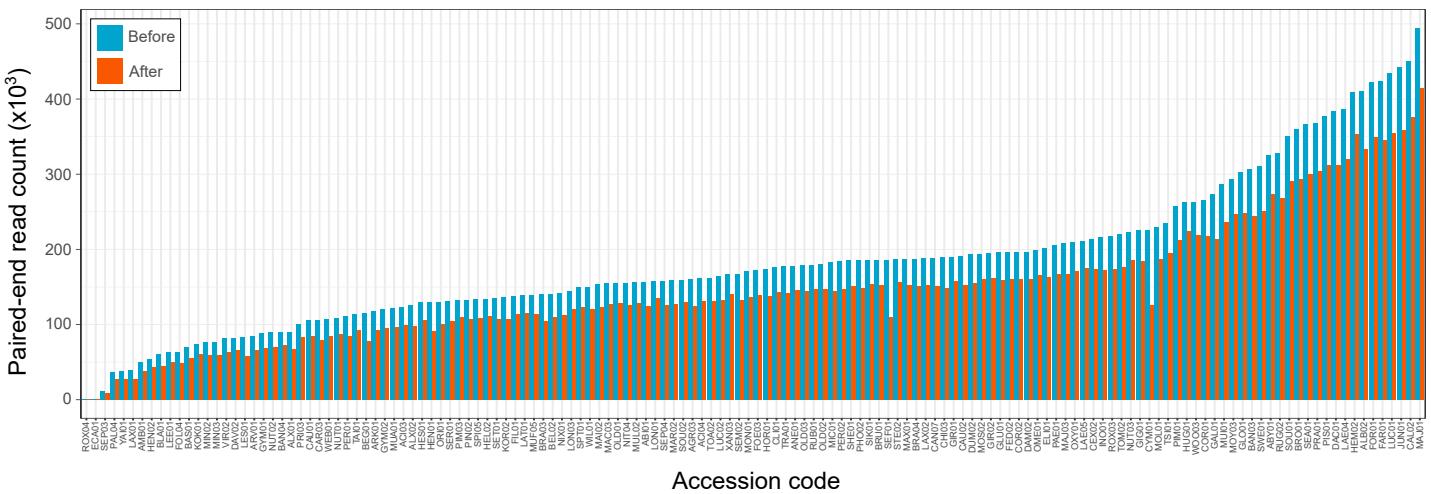
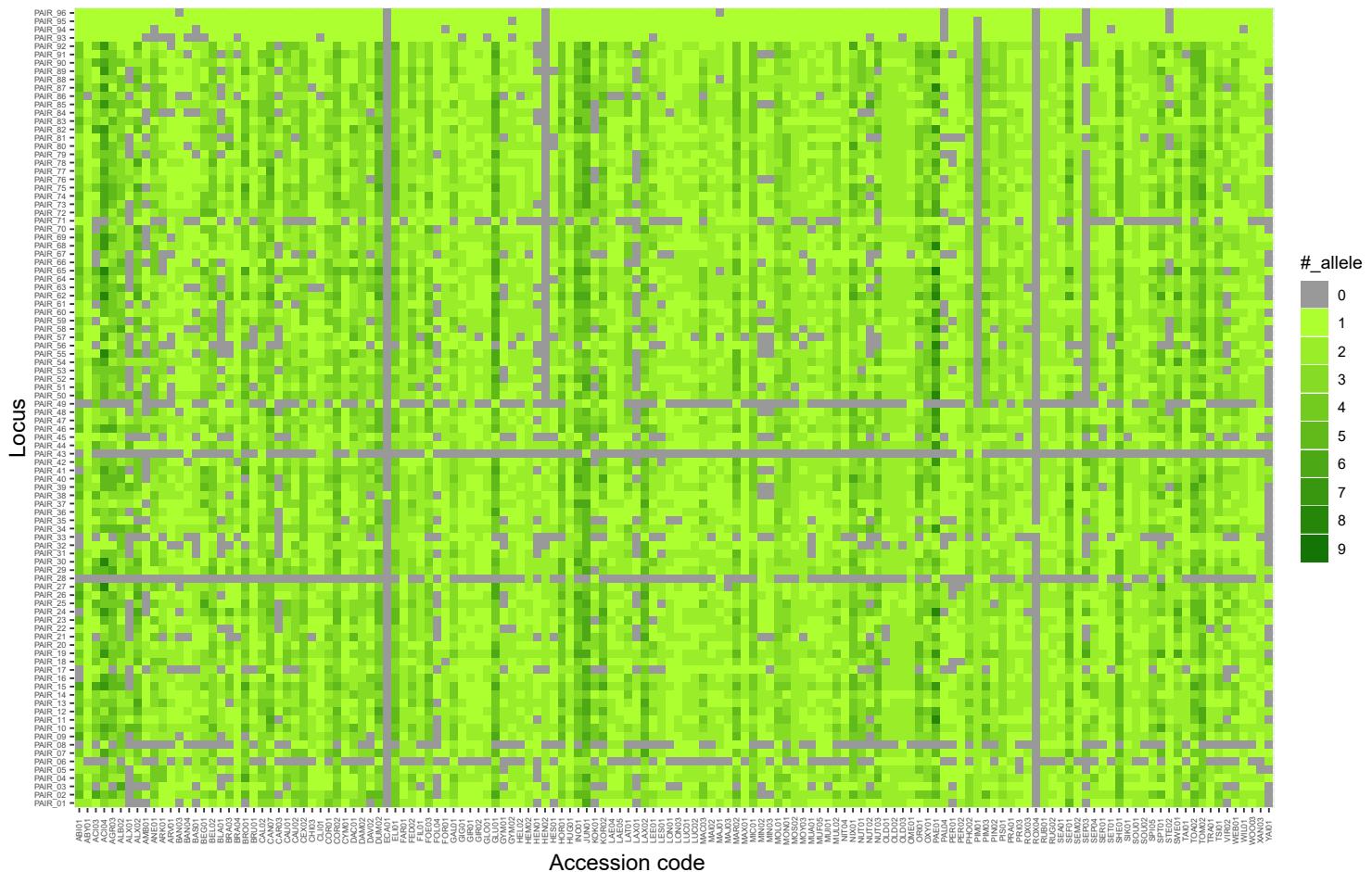


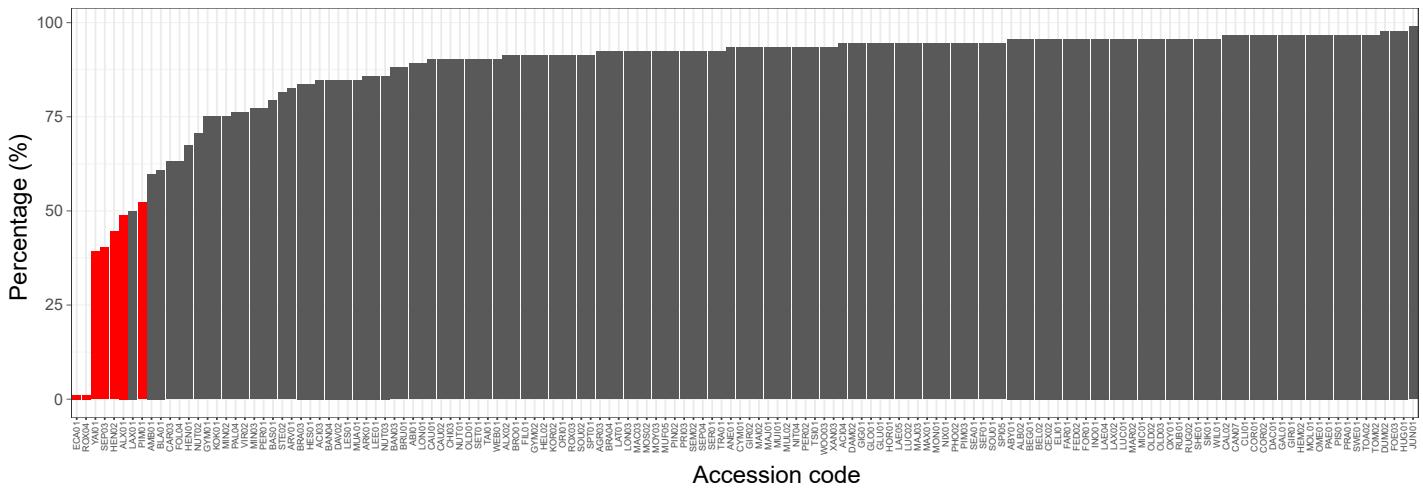
Supplementary figure S1. Geographical origins of the *Rosa* accessions with tissue fragments preserved at IRHS. Localizations were assigned as close as possible according to vouchers. When no precise localization was available, we attribute one region according to literature.

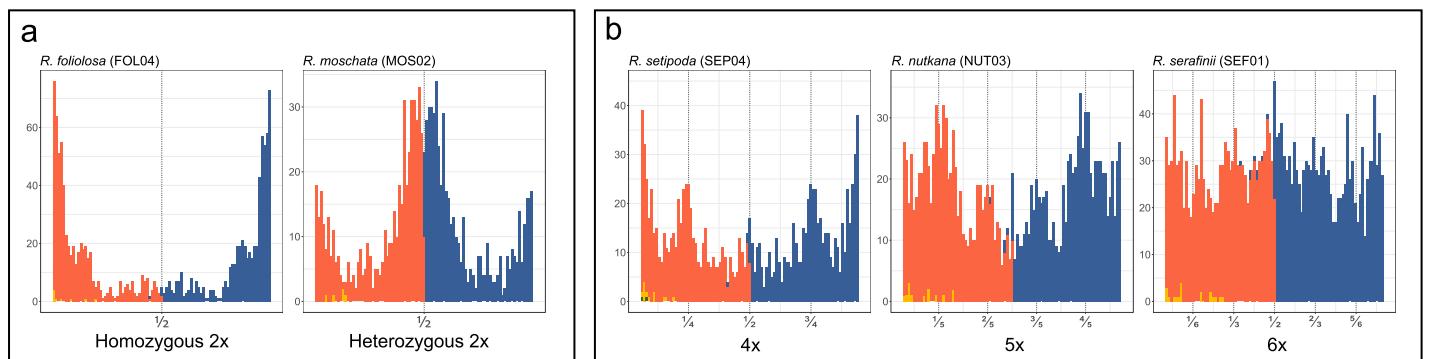


Supplementary figure S2. Paired-end read count before and after read processing. 'Before' refers to raw reads as received from the sequencing platform (end of step 1, Figure 1). 'After' refers to processed reads (step3, Figure 1).

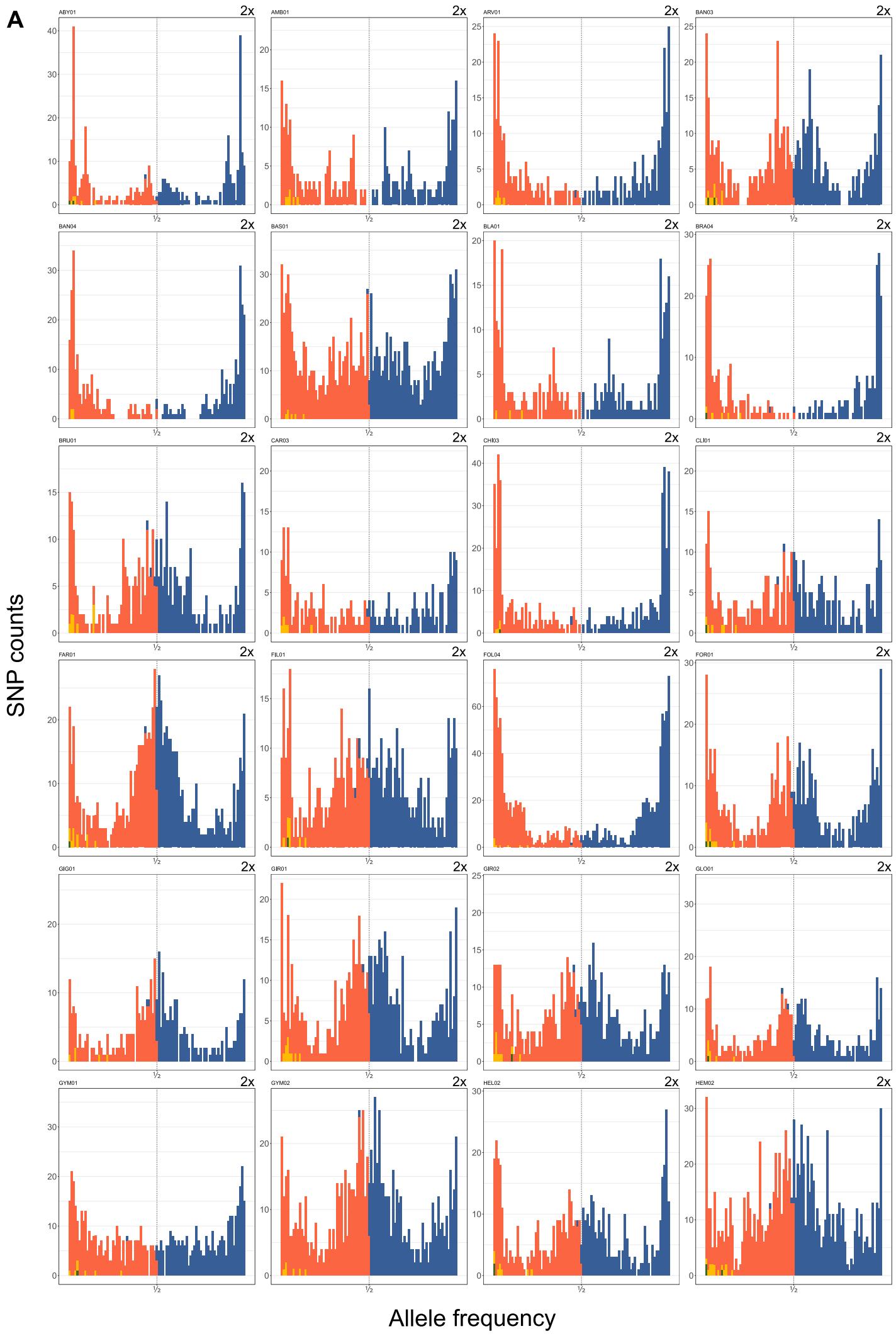


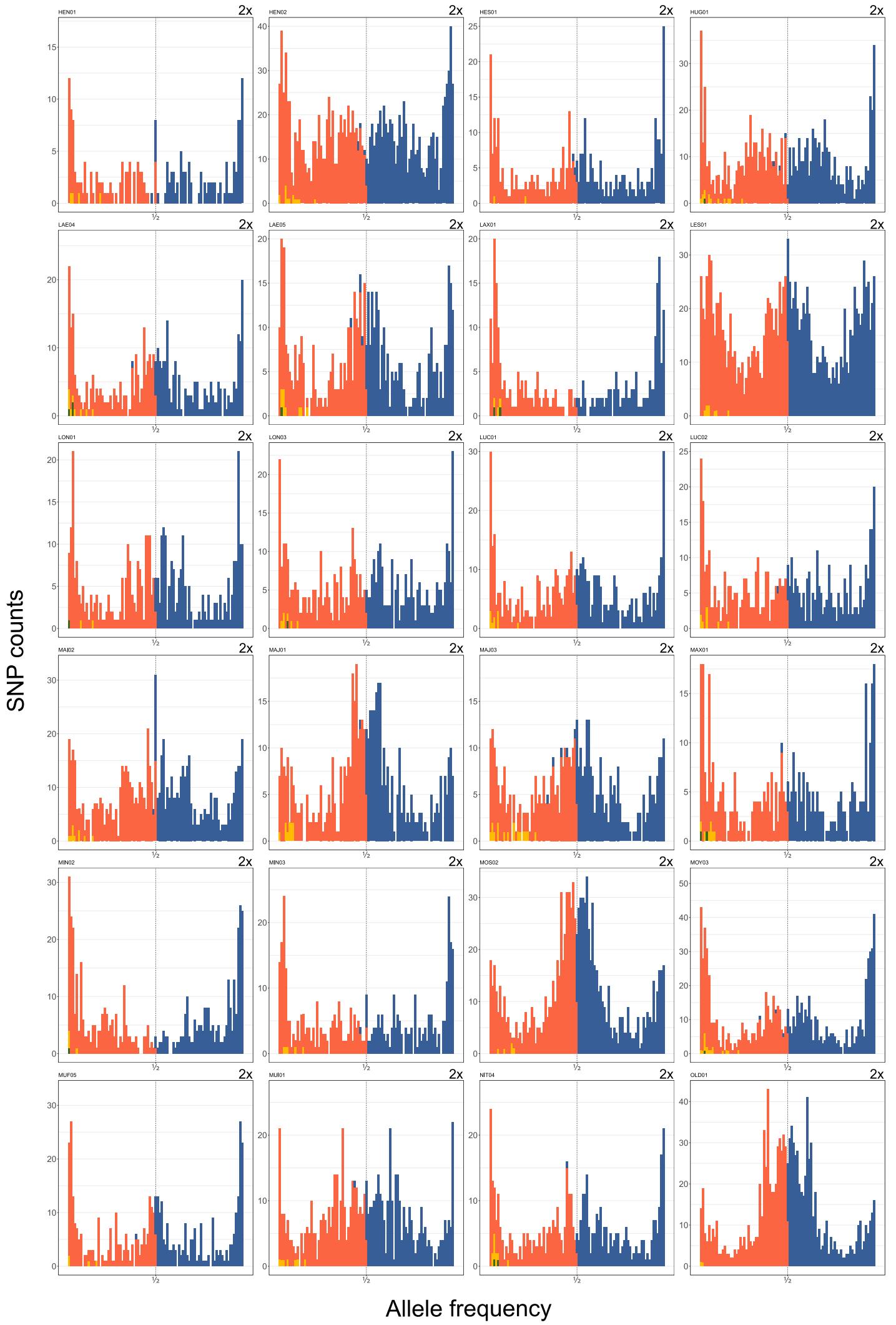
Supplementary figure S3. Heat map showing the number of alleles recovered for each sample at each locus.

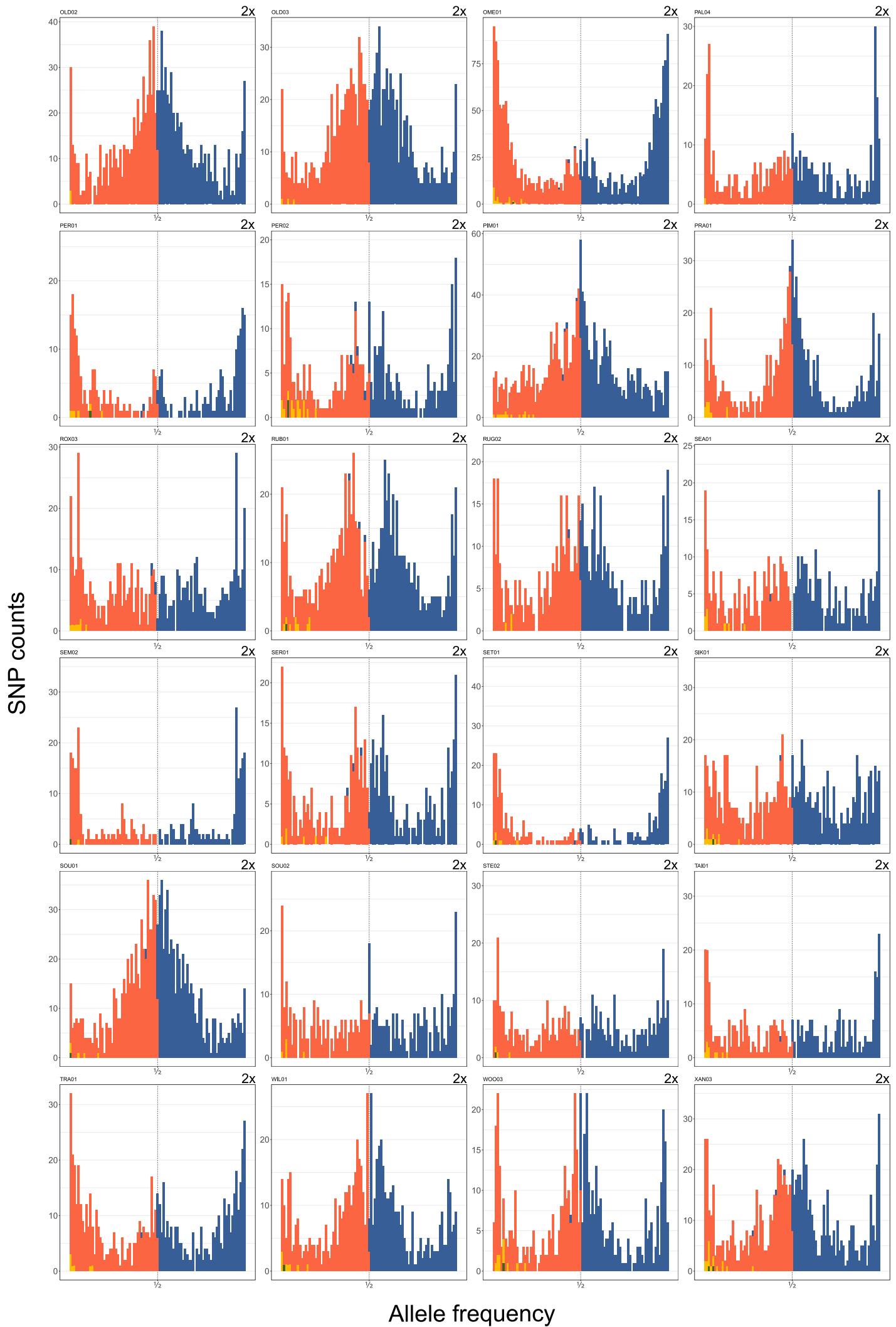


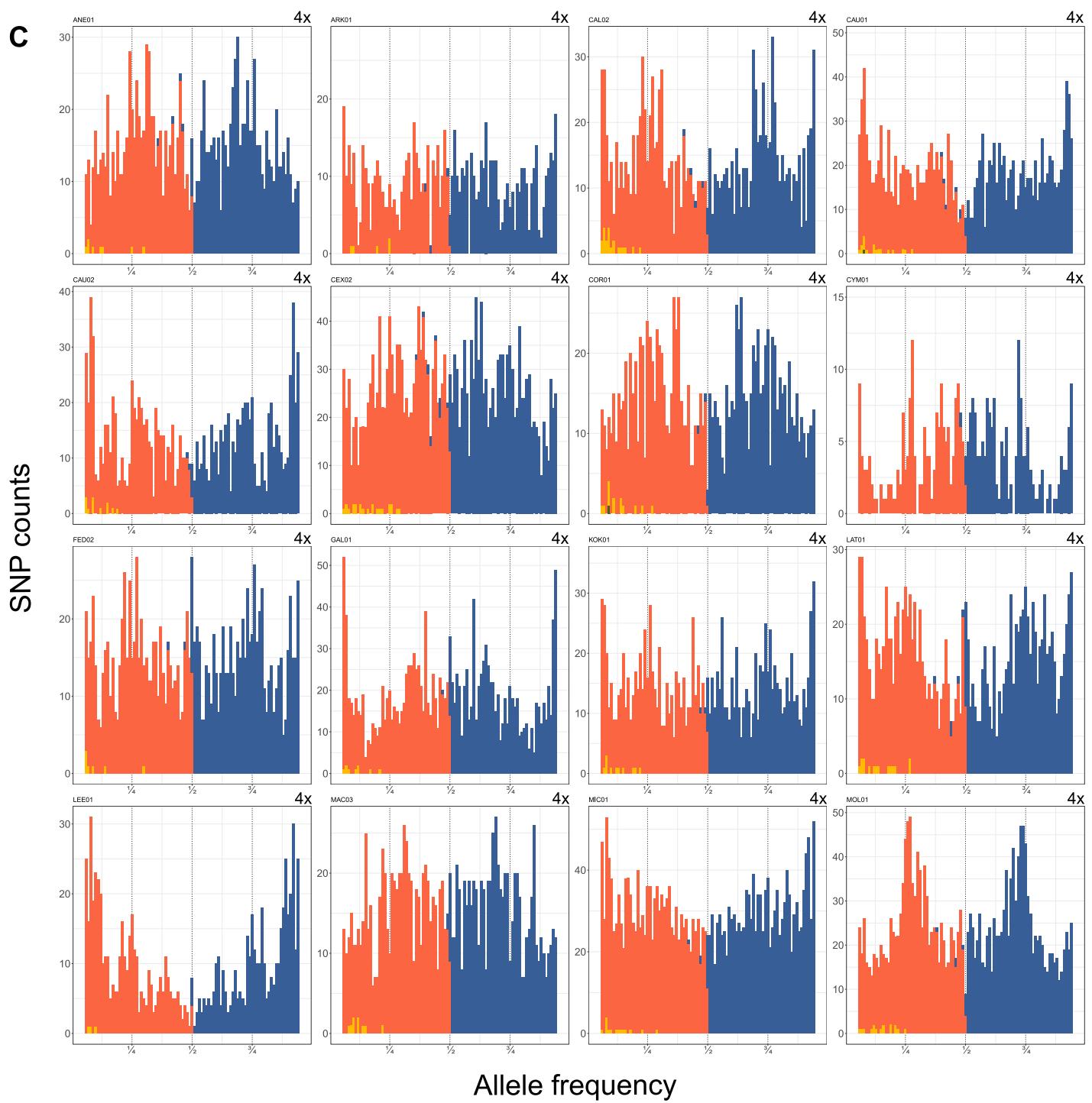
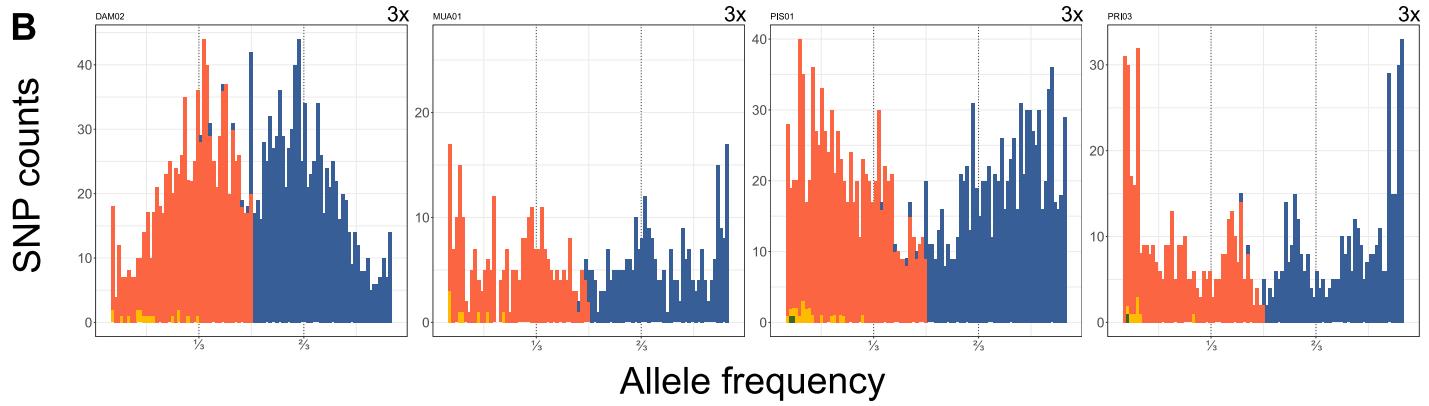


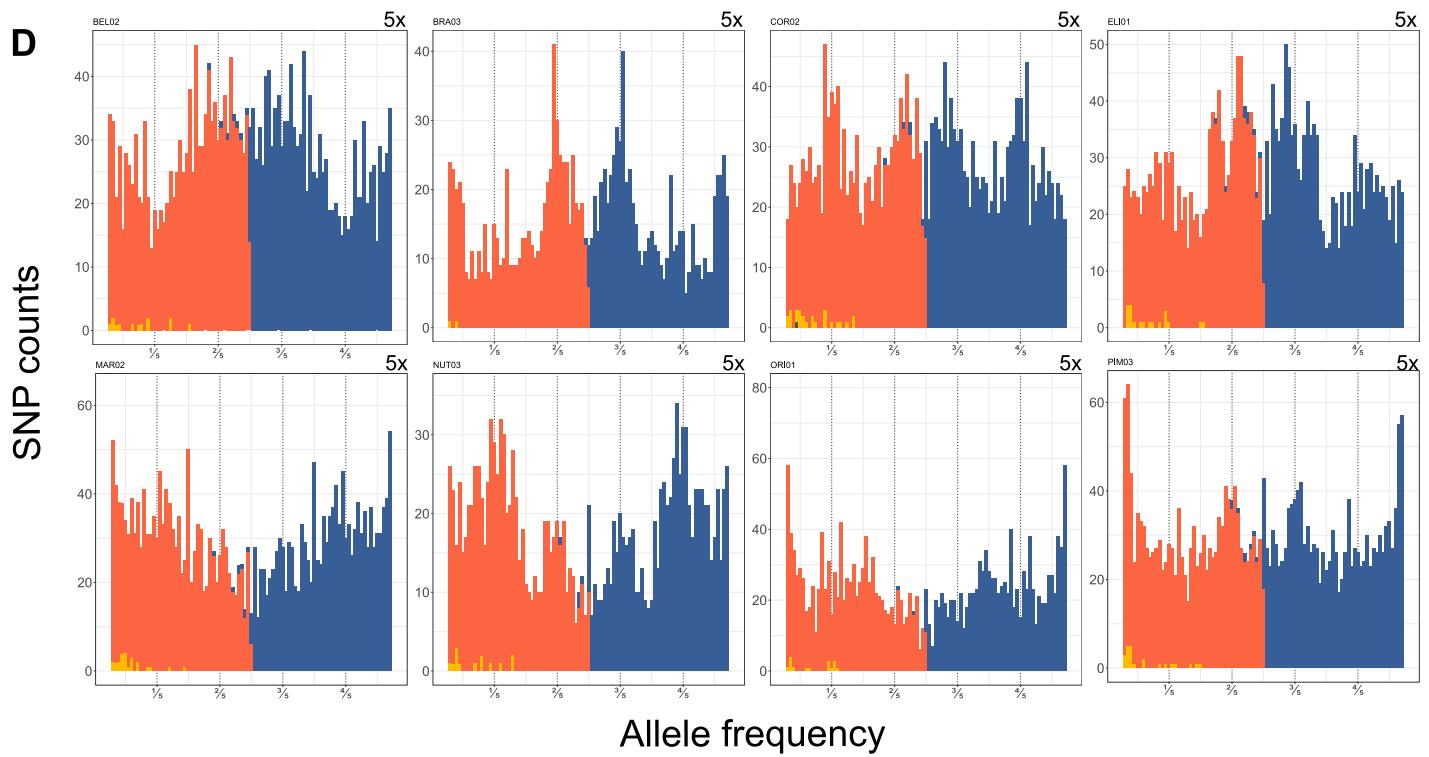
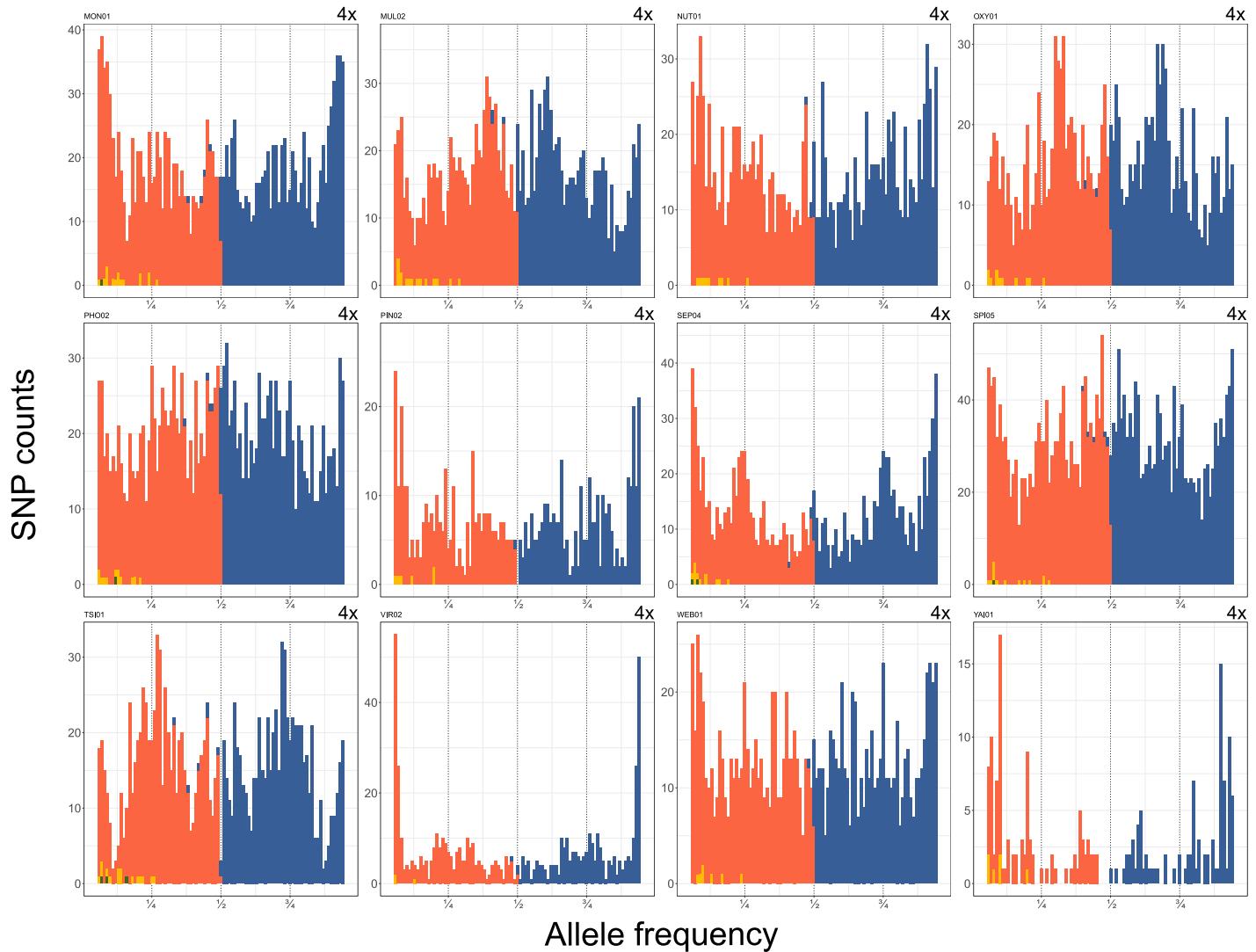
Supplementary figure S5. Typical allele frequency distributions for several ploidy levels. (a) Diploid examples with *left*: U-shaped distribution corresponding to a homozygous diploid specimen. *right*: W-shaped distribution corresponding to a heterozygous diploid specimen. (b) Polyploid examples, from left to right: tetraploid (4x), pentaploid (5x), and hexaploid (6x). Species names are followed by their accession code and are presented above each distribution. Frequencies of the first, second, third and fourth allele at each heterozygous SNP position are represented by blue, orange, yellow and green colors, respectively. The x-axis represents the allele frequency. The y-axis represents the count of heterozygous SNPs. Allele frequency distributions are plotted within the range 5-95%.

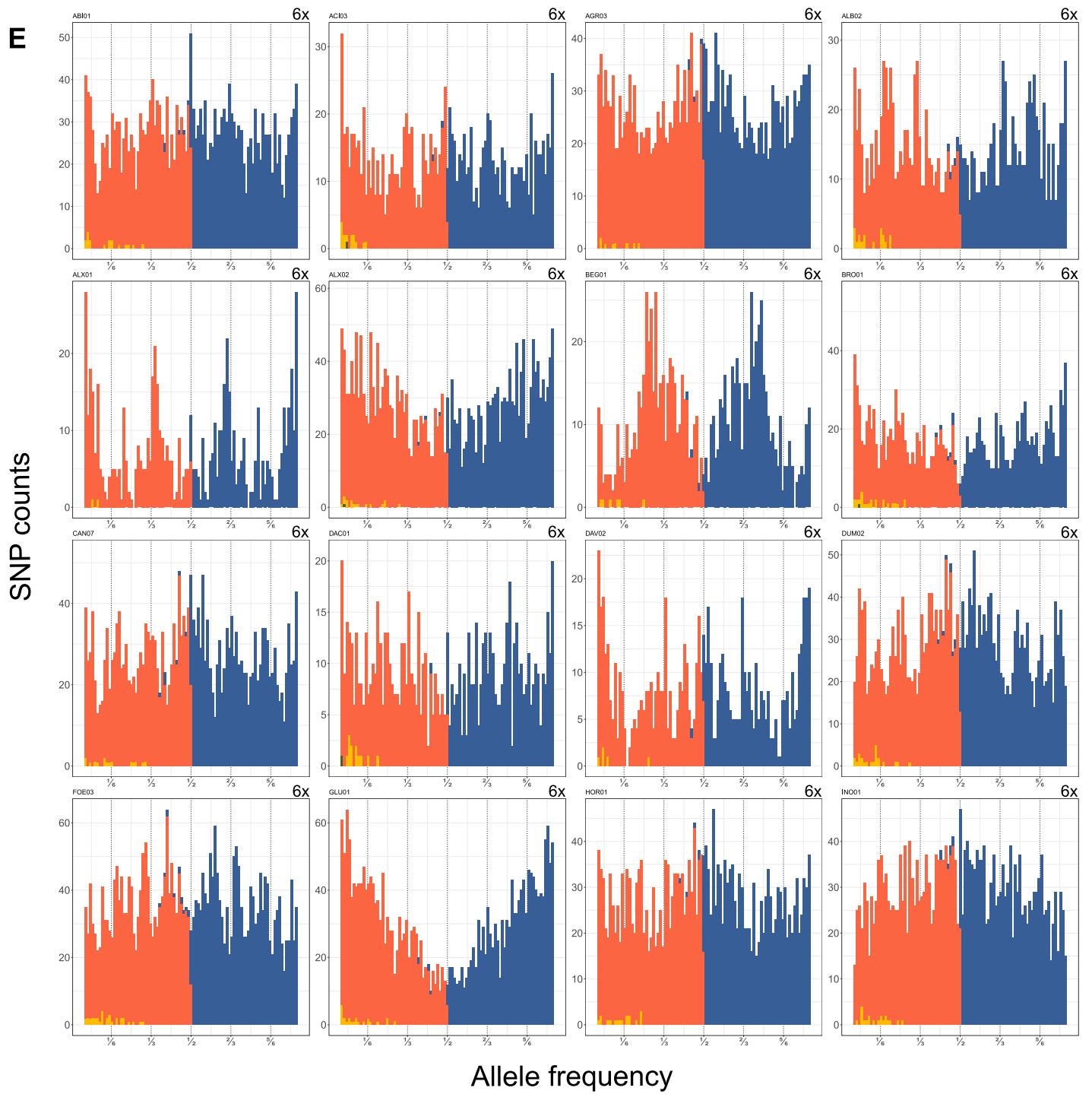
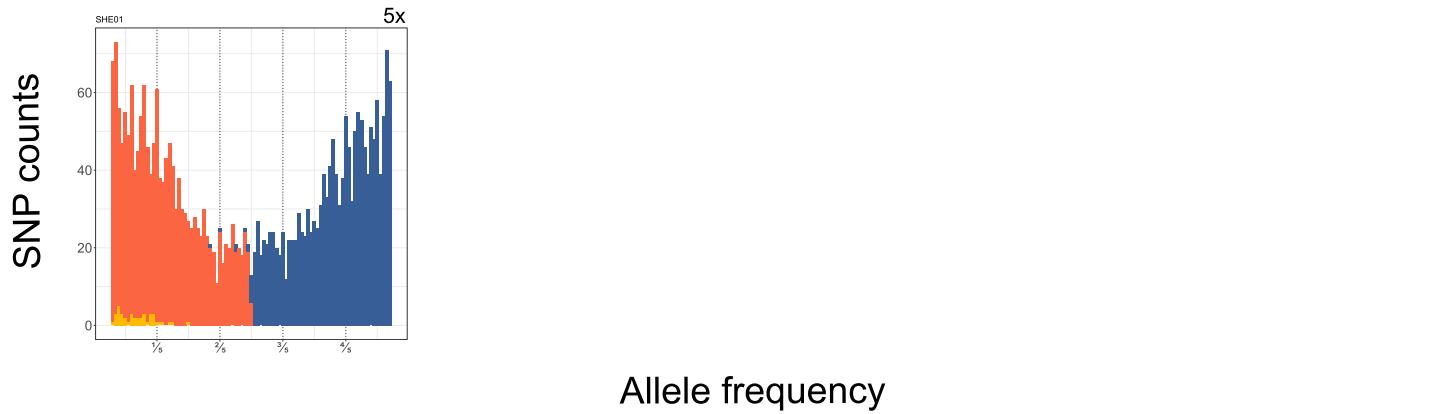
**A**

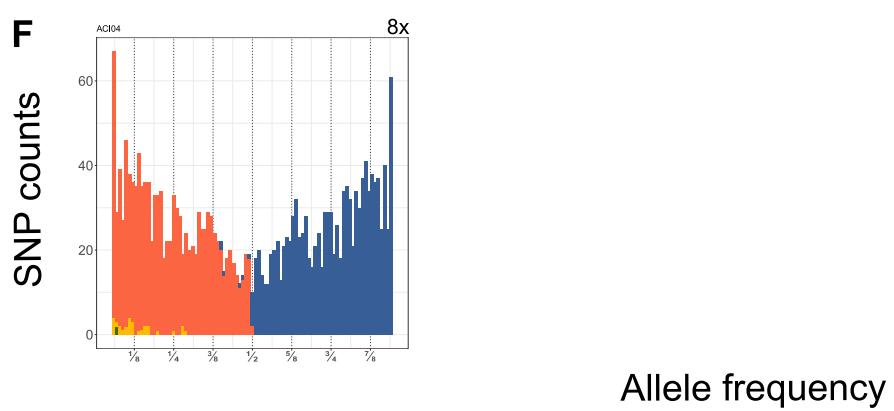
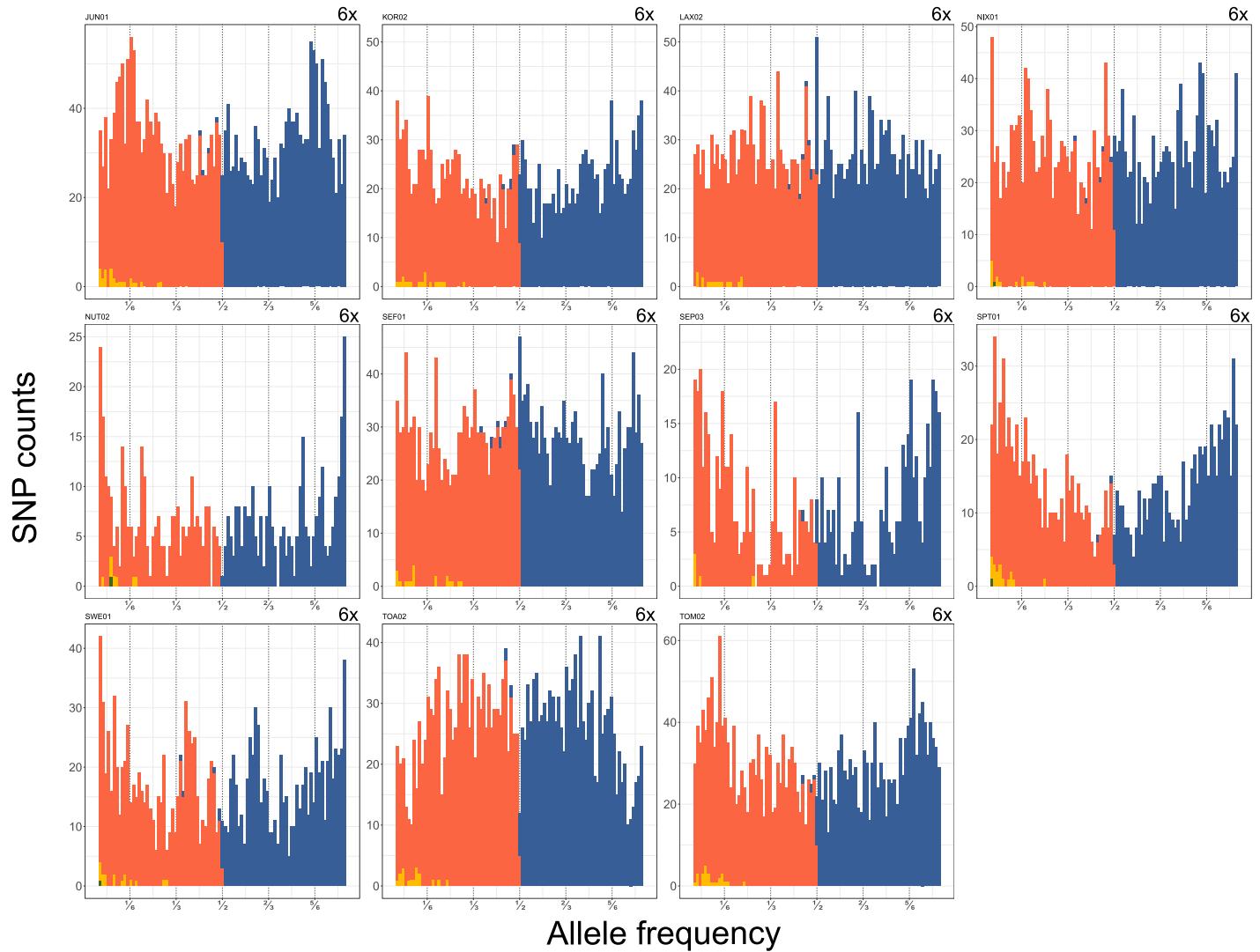


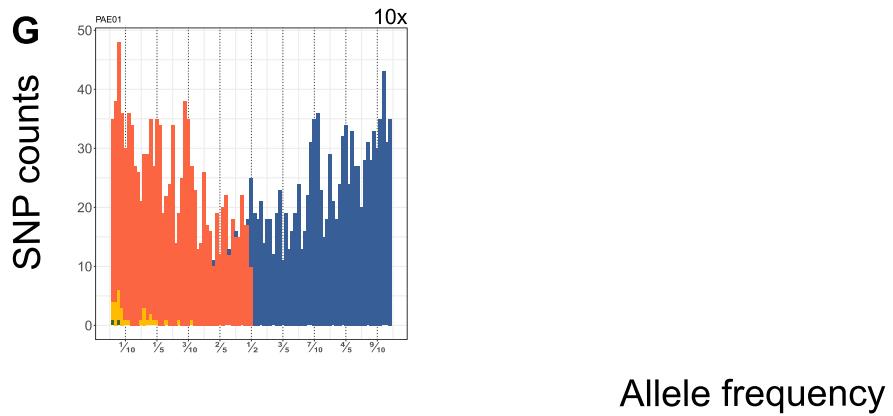






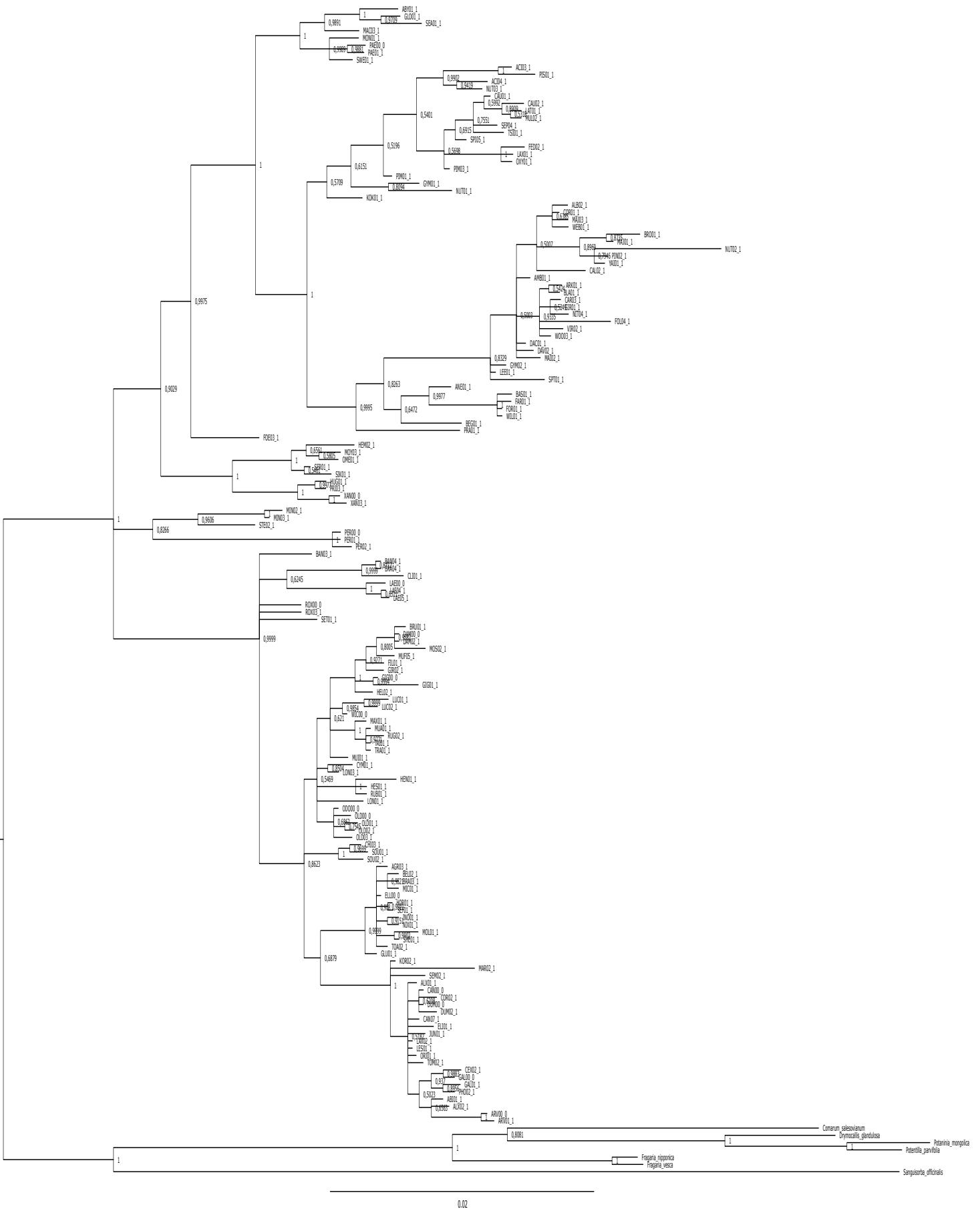




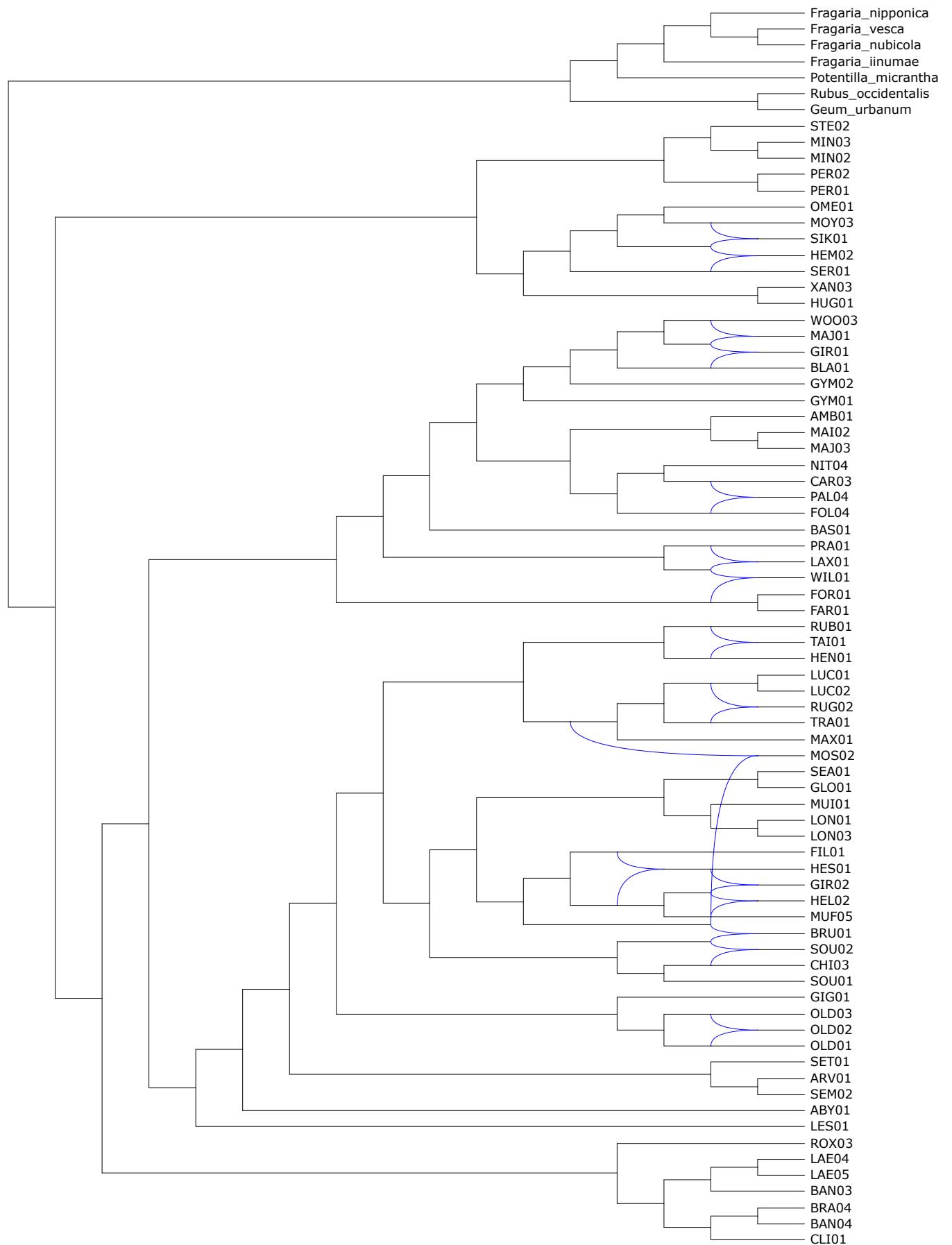


Supplementary figure S6. Estimation of the ploidy level for each accession of the study.

Accessions are grouped according to their putative ploidy level as estimated by the distribution of allele frequencies with A) diploid, B) triploid, C) tetraploid, D) pentaploid, E) hexaploid, F) octoploid and G) decaploid. Each accession's sequencing reads are mapped to a reference made of concatenated nuclear single-copy orthologous tags (nrSCO<sub>Tag</sub>s) identified in *Rosa 'Old Blush'* (Hibrand Saint Oyant et al. 2018). Heterozygous SNP positions within the read set are used to extract allele frequencies. Maximum four alleles can be retrieved at each heterozygous SNP position corresponding to the four nucleotides (A, C, T, and G). Allele frequencies at all heterozygous SNP positions are plotted as a distribution. First, second, third and fourth allele at each heterozygous position is colored in blue, orange, yellow and green, respectively. The x-axis correspond to allele frequency, i.e. the fraction of the reads that contain each allele. Distributions are shown within the range 5%-95% because homozygous SNP positions are not informative. The distribution of allele frequencies is used as a proxy to estimate sub-genome dosage and therefore ploidy (e.g. a distribution with spikes around 25%, 50%, and 75% is likely to reflect a tetraploid accession). For each subfigure, the fractions where spikes were observed are labeled on the x-axis. Two accessions (ECA01 (*R. ecae*), ROX04 (*R. roxburghii*) were not presented here due to a lack of read coverage.



Supplementary figure S7. Maximum clade credibility tree obtain after bayesian search on the concatenation of the four plastid loci. Node numbers correspond to posterior probabilities. Branches supporting bipartitions present in less than 50% of the sampled tree in the posterior distribution were collapsed. Leaf names correspond to accession codes plus the allele number.



Supplementary figure S8. Network showing the reticulate phylogenetic relationships among diploid accessions. Nuclear allele trees were searched for a coalescent super allele tree that was further converted into a MUL-tree to obtain a hybrid network. Leaf names correspond to accession codes.

Supplementary table S1. Genbank references of extra *Rosa* accessions and outgroup species used for plastid phylogenies. Species for which a *de novo* plastid genome assembly was required have their names in bold.

	Species	Accession code	GenBank reference
Ingroup	<b><i>Rosa arvensis</i></b>	ARV00	SRX3286288
	<b><i>Rosa canina</i></b>	CAN00	ERX1733250
	<b><i>Rosa × damascena</i></b>	DAM00	SRX3286290, SRX3286291
	<b><i>Rosa dumalis</i></b>	DUM00	ERX1733252
	<b><i>Rosa elliptica</i></b>	ELL00	ERX1733251
	<b><i>Rosa gallica</i></b>	GAL00	SRX4006794
	<b><i>Rosa gigantea</i></b>	GIG00	SRX3286283, SRX3286284
	<b><i>Rosa laevigata</i></b>	LAE00	SRX4006792
	<b><i>Rosa × odorata</i></b>	ODO00	SRX3286293
	<b><i>Rosa 'Old Blush'</i></b>	OLD00	PRJNA445774
	<b><i>Rosa persica</i></b>	PER00	SRX4006789
	<i>Rosa praelucens</i>	PAE00	MG450565
	<i>Rosa roxburghii</i>	ROX00	PRJNA356521
	<b><i>Rosa wichurana</i></b>	WIC00	SRX3286280, SRX3286281
	<b><i>Rosa xanthina</i></b>	XAN00	SRX4006788
Outgroup	<i>Fragaria nipponica</i>	na	KY769125
	<i>Fragaria vesca</i>	na	JF345175
	<i>Potentilla parvifolia</i>	na	KY420033
	<i>Potaninia mongolica</i>	na	KY419959
	<i>Drymocallis glandulosa</i>	na	KY420015
	<i>Comarum salesovianum</i>	na	KY420034
	<i>Sanguisorba officinalis</i>	na	KY419975

Supplementary table S2. Genbank references of outgroup species used to root nuclear phylogenies.

	Species	BioProject / SRA code reference
Outgroup	<i>Fragaria iinumae</i> Makino	PRJDB1478
	<i>Fragaria nipponica</i> Makino	PRJDB1479
	<i>Fragaria nubicola</i> Lindl. Ex Lacaita	PRJDB1480
	<i>Fragaria vesca</i> L.	PRJNA383733
	<i>Geum urbanum</i> L.	PRJEB23412
	<i>Potentilla micrantha</i> Ramond ex DC.	PRJEB18433
	<i>Rubus occidentalis</i> L.	PRJNA430858