SI Appendix

Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex

Tao Ma[a,1], Kun Wang[a,1], Quanjun Hu[a,1], Zhenxiang Xi[a,1], Dongshi Wan[b], Qian Wang[a], Jianju Feng[b], Dechun Jiang[b], Hamid Ahani[c], Richard J. Abbott[d], Martin Lascoux[e], Eviatar Nevo[f,2], Jianquan Liu[a,2]

[a]Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, Sichuan, P. R. China;

[b]State Key Laboratory of Grassland Agro-ecosystem, College of Life Sciences, Lanzhou University, Lanzhou 730000, P. R. China

[c]Natural Resources and Watershed management Administration of Khorasan Razavi 9177948974, Iran

[d]School of Biology, University of St Andrews, St Andrews, Fife KY16 9TH, UK

[e]Department of Ecology and Genetics, Evolutionary Biology Center and Science for Life Laboratory, Uppsala University, Uppsala 75236, Sweden

[f]Institute of Evolution, University of Haifa, Mount Carmel, Haifa 3498838, Israel

[1]T.M., K.W., Q.H., and Z.X. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: liujq@nwipb.cas.cn or nevo@research.haifa.ac.il

SI Text

Sample collection, sequencing and mapping. Silica gel dried leaves representing 34 *P. pruinosa* and 122 *P. euphratica* individuals spanning the geographic ranges of the two species were collected from wild populations in western China, Kyrgyzstan, Iran, Israel, Spain and Algeria. Based on previous studies (1, 2) we excluded known hybrids between the two species occurring in western China where both species co-occur. In each population, sampled individuals were at least 500 m apart. For each individual, genomic DNA was extracted with a standard protocol and sequenced on the HiSeq 2000 and HiSeq 2500 Illumina platforms. The generated raw reads were subject to quality control and low-quality reads were removed if they met either of the following criteria: i) $\geqslant$10% unidentified nucleotides (*N*); ii) a phred quality $\leq$7 for $>$65% of read length; iii) reads overlapping more than 10 bp with the adapter sequence, allowing < 2 bp mismatch; iv) duplicate reads. We also trimmed both ends of a read if the consecutive three bases had a phred quality $\leq$13, and discarded reads shorter than 45 bp after trimming. High quality clean reads were mapped to the *P. euphratica* reference genome (3) using BWA-MEM (0.7.10-r789) with default parameters (4), and alignment results and marked duplicate reads were sorted using SAMtools (v0.1.19) (5). The Picard package (http://picard.sourceforge.net/) was subsequently used to assign readgroup information containing library, lane, and sample identity, and the Genome Analysis Toolkit (GATK) (6) was used to perform local realignment of reads to enhance alignments in regions around putative indels.

Filtering alignments and sites. Before SNP calling, we removed low quality alignments which met either of the following criteria to avoid their effects on SNP detection and subsequent analysis: i) alignments with a mapping quality score lower than 30; ii) reads that have multiple best hits; iii) reads with a flag above 255; iv) pairs of reads with mates mapped incorrectly; v) alignments which anchored the short scaffolds of less than 2 kbp. To minimize the influence of sequencing and mapping

bias, we also removed low quality sites according to the following criteria: i) sites with unbalanced allele quality scores as determined using Wilcoxon rank sum test with threshold of P < 1e-5; ii) sites with strand bias (P < 1e-5); iii) bases with a quality below 20; iv) sites with extremely low (<2×) and extremely high coverage (>60×) per individual; v) sites that failed the Hardy-Weinberg Equilibrium test at P < 1e-3 using SAMtools and BCFtools (5). These filters left us with a data set comprising ~358 Mbp, representing 72.4% of the genome.

SNP discovery and genotype calling. We used the SAMtools model (7) implemented in analysis of next generation sequencing data (ANGSD) (8) to estimate genotype likelihoods and to generate Beagle files at the population level. The major/minor state was determined by estimating the minor allele frequency (9) based on genotype likelihoods. A likelihood ratio test statistic for population allele frequency and P < 1e-6 was used as a SNP discovery criterion. SNPs were retained only if they were detected in more than 50% of sampled individuals over all populations. The site frequency spectrum (SFS) was estimated using realSFS (10) implemented in ANGSD. The full data were used to compute posterior probabilities of genotypes at each site for each individual using the sample allele frequency as prior for genotype frequencies under the assumption of Hardy-Weinberg equilibrium. BEAGLE (11, 12) was then used to infer haplotypes of individuals within each population and improve the genotype imputation based on genotype likelihoods previously estimated. In a final filtering strategy, the sites that showed a correlation between the observed and imputed data (r^2) <0.9 were removed, resulting in a total of 5,057,358 SNPs.

Relatedness analysis. We used the KING program (13) to estimate degrees of relatedness between all individuals based on pairwise comparisons of SNP data. Those pairs exhibiting greater than 3rd-degree relationships were removed, leaving a total of 99 individuals (27 and 72 individuals for *P. pruinosa* and *P. euphratica*, respectively) for subsequent population genomics analyses.

Ancestral state reconstruction. To identify the ancestral state of *P. euphratica* and *P. pruinosa*, ten additional individuals representing the following taxa were sequenced to at least 15× depth: *P. ilicifolia*, *P. davidiana*, *P. lasiocarpa*, *P. laurifolia*, *P. nigra*, *P. trichocarpa*, *P. deltoides*, *P. fremontii*, *P. angustifolia* and *P. tremuloides*. We mapped these data to the *P. euphratica* reference genome and genotyped all segregating sites using the same criteria as described above. Sites that were detected in at least three species and exhibited the same homozygous genotype were directly used as the ancestral state. After the removal of tri-allelic sites, we identified the ancestral state at about 64% of the SNPs.

Population structure analysis. We compiled an artificial nucleotide sequence from the nuclear genome data comprising all SNPs and their ancestral states determined above. We then performed a phylogeographic analysis using Neighbor-net (14) implemented in SPLITSTREE (version 4.13.1) (15) with default parameters. After mapping our resequencing data against the *P. trichocarpa* chloroplast genome (Accession Number NC_009143.1), only positions covered by a minimum number of three independent unique reads with base qualities of ≥30 were used to call consensus sequences. The phylogenetic tree was constructed using maximum likelihood with MEGA (16). Principal component analysis (PCA) was performed on all SNPs using the smartpca program in EIGENSOFT (17). A Tracy-Widom test was used to determine the significance level of eigenvectors. The software identified two *P. pruinosa* (Kyrgyzstan) and five *P. euphratica* (Spanish, Algerian, Israeli and two Iranian) individuals as outliers, which were subsequently removed from analysis. In addition, we used the program NGSadmix (18), which is based on genotype likelihoods to directly estimate individual admixture proportions from next generation sequencing data, the results revealed a similar population genetic structure in our samples. To further investigate relationships, patterns of splits and mixtures between these populations, we produced a maximum likelihood drift tree using the software TreeMix (19) and inferred migration events. For further insight into relationships among lineages, we performed identity-by-descent (IBD) blocks analysis using the

algorithm from BEAGLE (11, 12) with the following parameters: window = 100,000; overlap = 10,000; ibdtrim = 100; ibdlod = 10.

Demographic inference. We used individuals with high sequencing coverage (>15×) from each population and applied the Pairwise Sequentially Markovian Coalescent (PSMC) model (20) to reconstruct demographic history. PSMC requires diploid consensus sequences. The consensus was generated from the 'pileup' command of the SAMtools software package. We masked bases with extremely low (less than a third of the average depth) or extremely high coverage (threefold the average depth) and used the parameters -N25 -t15 -r5 -p "4+25*2+4+6" to perform the analysis. Results were scaled using an assumed mutation rate of 3.8×10^{-8} per base pair per generation and a generation time of 15 years (21). The bootstrapping analysis was performed following the software instruction: 1) the program "splitfa" in PSMC package was used to split long sequences into shorter segments (utils/splitfa diploid.psmcfa > split.psmcfa); 2) the PSMC was performed 100 times with "-b" option, which will enable PSMC to randomly sample with replacement from these segments (seq 100 | xargs -i echo psmc -N25 -t15 -r5 -b -p "4+25*2+4+6" -o round-{}.psmc split.fa | sh); 3) combine the result files together and plot with program 'psmc_plot.pl' from PSMC package (cat diploid.psmc round-*.psmc > combined.psmc; utils/psmc_plot.pl combined combined.psmc). We also used coalescent simulations applying the composite likelihood method implemented in fastsimcoal2.1 software (22) to infer demographic parameters based on the site frequency spectrum. The folded spectra between each pair of populations were used to minimize potential biases when determining ancestral allelic states. Alternative models of historical events were fitted to the joint site frequency spectra data. All parameter estimates were global ML estimates from 50 independent fastsimcoal2.1 runs, with 100,000 simulations per likelihood estimation (-n100,000, -N100,000) and 40 cycles of the likelihood maximization algorithm. Confidence intervals of parameter estimates were obtained by parametric bootstrapping, with 100 bootstrap replicates per model. The best model was identified through the maximum value of likelihoods and Akaike's information

criterion (AIC); simulated datasets were compared with the observed site frequency spectra to evaluate the fit of the best demographic model.

Linkage Disequilibrium and Recombination. We measured and compared patterns of linkage disequilibrium (LD) for each lineage except PeE due to its limited sample size. The correlation coefficient r^2 of pairwise SNPs was calculated using Haploview (23) with the parameters: -dprime -maxDistance 50 -minMAF 0.05 -hwcutoff 0.001 -minGeno 0.6. To minimize any bias due to sample size variation, we randomly reduced the size of each lineage to the same number of individuals (n = 21). Population-scaled recombination rates along each scaffold ($\rho = 4N_e r$ per kb) were estimated using a Bayesian reversible-jump MCMC scheme under the crossing-over model of the program interval in the LDhat (24) package with standard defaults. Rates were estimated for each lineage separately.

Genomic diversity and divergence. Many of the population genetic summary statistics are direct functions of the site frequency spectrum (SFS). We estimated the folded SFS and calculated nucleotide diversity (π), population scaled mutation rate θ_w (Watterson's estimator) and Tajima's D for each lineage, using a probabilistic method (25) implemented in the software ANGSD (8) which takes genotype uncertainty into account. Using the likelihoods of sample allele frequencies for each population and a 2D-SFS as a prior, we quantified population genetic differentiation (F_{ST}) and the number of nucleotide substitutions (D_{xy}) between each pair of populations using the software ngsTools (26, 27). All of the parameters were calculated using a sliding window approach (10 Kbp non-overlapping window) and the windows were discarded if there were less than 3 Kbp left after all of the above filtering steps.

Coalescent simulations. To examine the effects of demographic processes on the heterogeneous pattern of genomic divergence, we conducted coalescent simulations to compare observed patterns of polymorphism and differentiation to those expected under different demographic scenarios. Simulations were performed using the

program *msms* (28) based on the parameters derived from the best-fitting model inferred above. Specifically, we simulated genotypes corresponding to a 10Kbp region with the same sample size as the real data for 500,000 replications, from where we simulated genotype likelihoods using the program msToGlf in ANGSD (8) by assuming a mean sequencing depth of 11× and an error rate of 0.005. We then estimated values of nucleotide diversity (π), Tajima's *D* and F_{ST} for all simulated replicates using ANGSD to test whether the simulated data matched the observed data.

Divergent regions of exceptional differentiation. The outlier windows for each population pair were identified by combining an empirical approach with a permutation approach. We first permuted the SNPs across the genome 5,000,000 times to get a null F_{ST} distribution for each of the windows holding different amounts of variable sites. We then determined the *P*-value by comparing window estimates of F_{ST} against its relative null distribution and corrected for multiple testing using the False Discovery Rate (FDR) adjustment. Only windows falling in the top 1% of the empirical F_{ST} distribution and with FDR lower than 0.01 were considered as putative outlier windows. This approach avoided a statistical power bias caused by a fixed threshold and minimized noise from individual site-based divergence estimates. Finally, adjacent outlier windows were combined to form larger divergent regions and the number and size of these regions were compared among all population pairs.

Genes under positive selection. We applied the Hudson-Kreitman-Aguadé (HKA) test (29) and the population branch statistic (PBS) (30) to verify whether recent selective sweeps had acted specifically in each lineage. Taking into account the very recent split between PeC and PeE, and the relatively few samples for PeE, we discarded PeE and compared the levels of polymorphism and divergence among the other three lineages (PeC, PeW and Pp) to identify genes under positive selection in each lineage. We considered only coding regions and analyzed a total of 28,148 genes. For each gene, we recorded the number of polymorphic sites in one lineage (A) and

the number of fixed differences (the sites with $F_{ST} > 0.95$) between this lineage and both of the other two lineages (B). We then performed the HKA test by comparing the ratio of A/B to the genome-wide average, computed as the sum of A and B values across all genes analyzed and testing the null hypothesis A(gene)/B(gene)=A(genome-wide)/B(genome-wide) using a Pearson's chi-square test on the 2x2 contingency table (31). In addition, we compared the three pairwise F_{ST} values between these populations and used the classical transformation by Cavalli-Sforza, $T = -\log(1 - F_{ST})$ to obtain estimates of lineage divergence time T in units scaled by population size. The length of the branch leading to one population was then obtained as:

$$PBS_{POP1} = \frac{T^{pop1-2} + T^{pop1-3} - T^{pop2-3}}{2}$$

Finally, genes with a significant nominal *P*-value (<0.01) for the HKA test and a ranked PBS above the 95th percentile were considered as positively selected genes. Functional classification of GO categories for these genes was performed using the Blast2GO program (32). Enrichment analysis was performed and Fisher's exact test was used to calculate the statistical significance of enrichment. *P*-values were further adjusted by Benjamini-Hochberg false discovery rate (33) and the adjusted *P*-value cut-off was 0.05.

SI Figures:

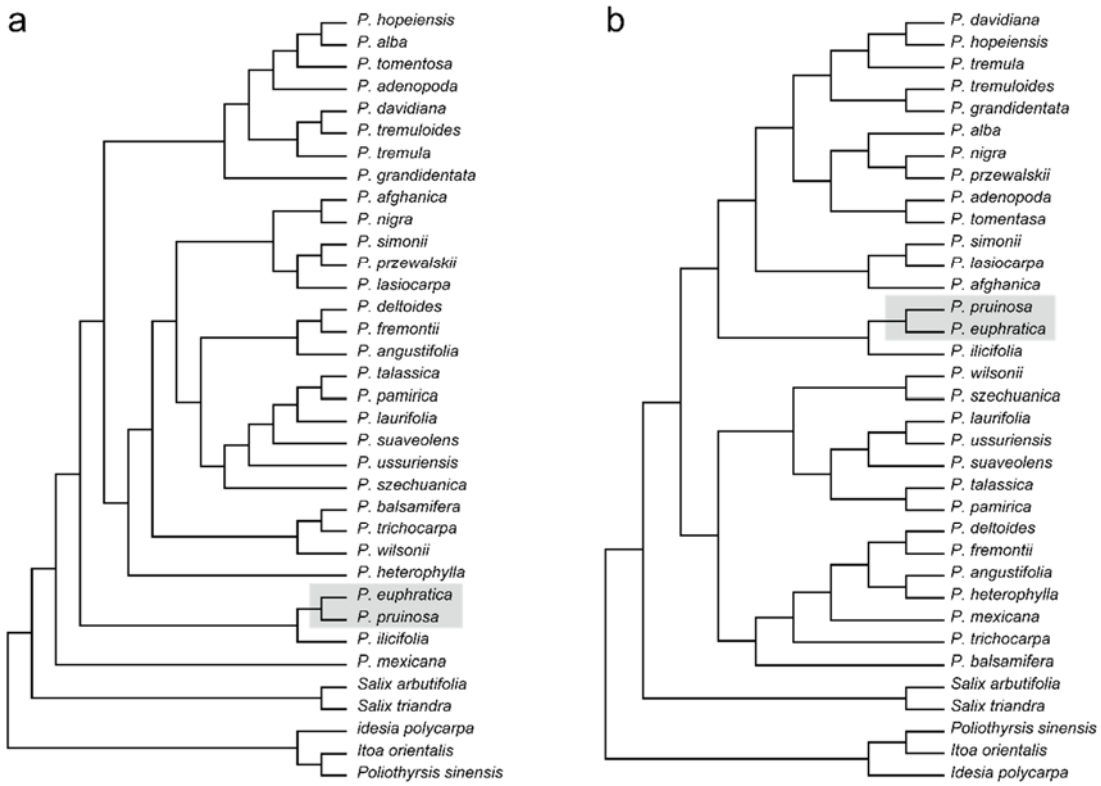


Fig. S1: Phylogeny of *Populus* from the combined (a) 23 single-copy nuclear DNA sequence and (b) 34 plastid fragments. Redrawn from ref. 34.

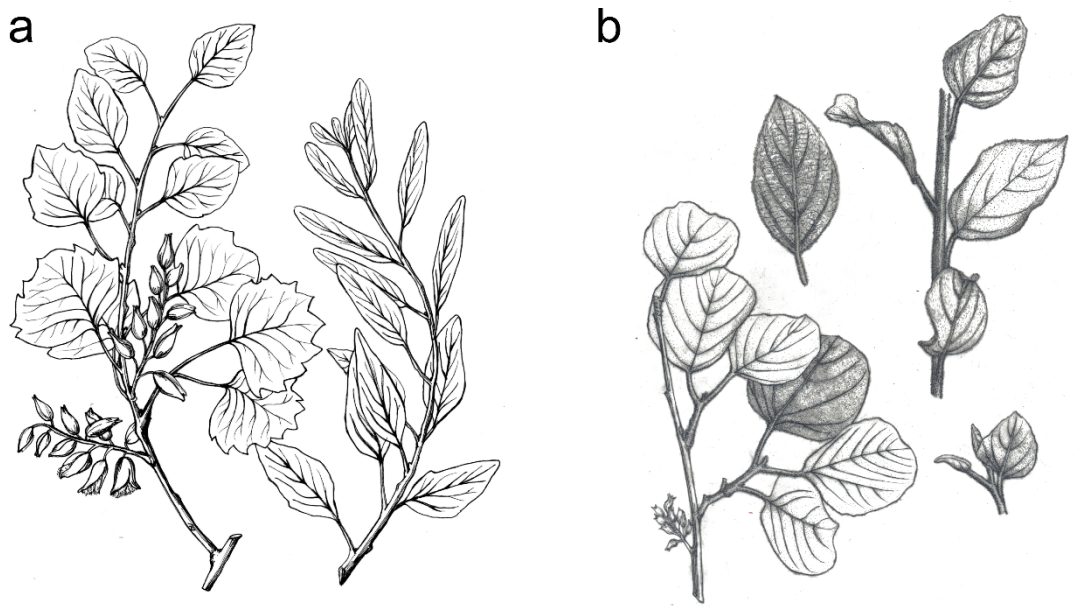


Fig. S2: Diagram of the leaf morphology for *Populus euphratica* (**a**) and *P. pruinosa* (**b**).

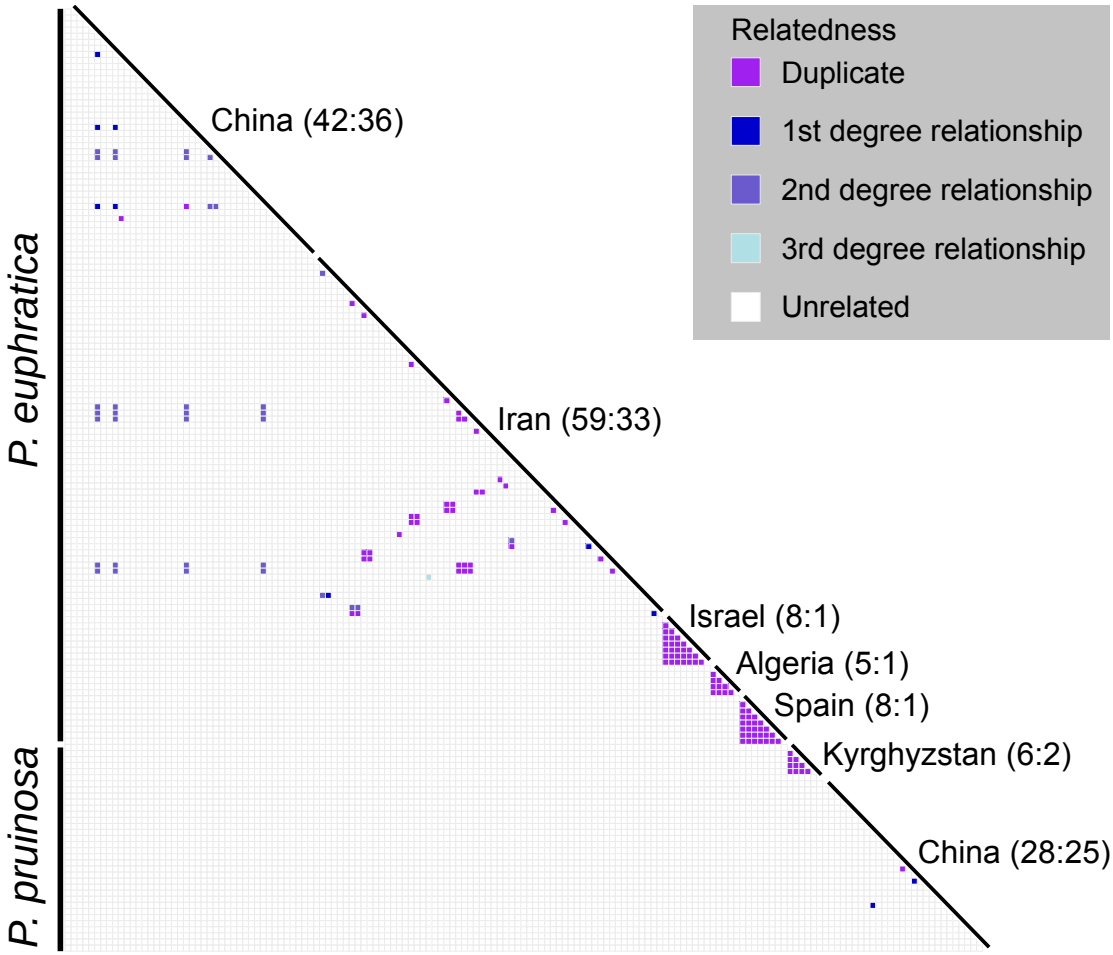


Fig. S3: Estimated genetic relatedness between each pair of 156 individuals using KING software, also showing the number of individuals before and after removing the samples that are more related than 3rd-degree relationships for each sample geographical origin. A total of 99 individuals were retained for subsequent population genetic analyses.

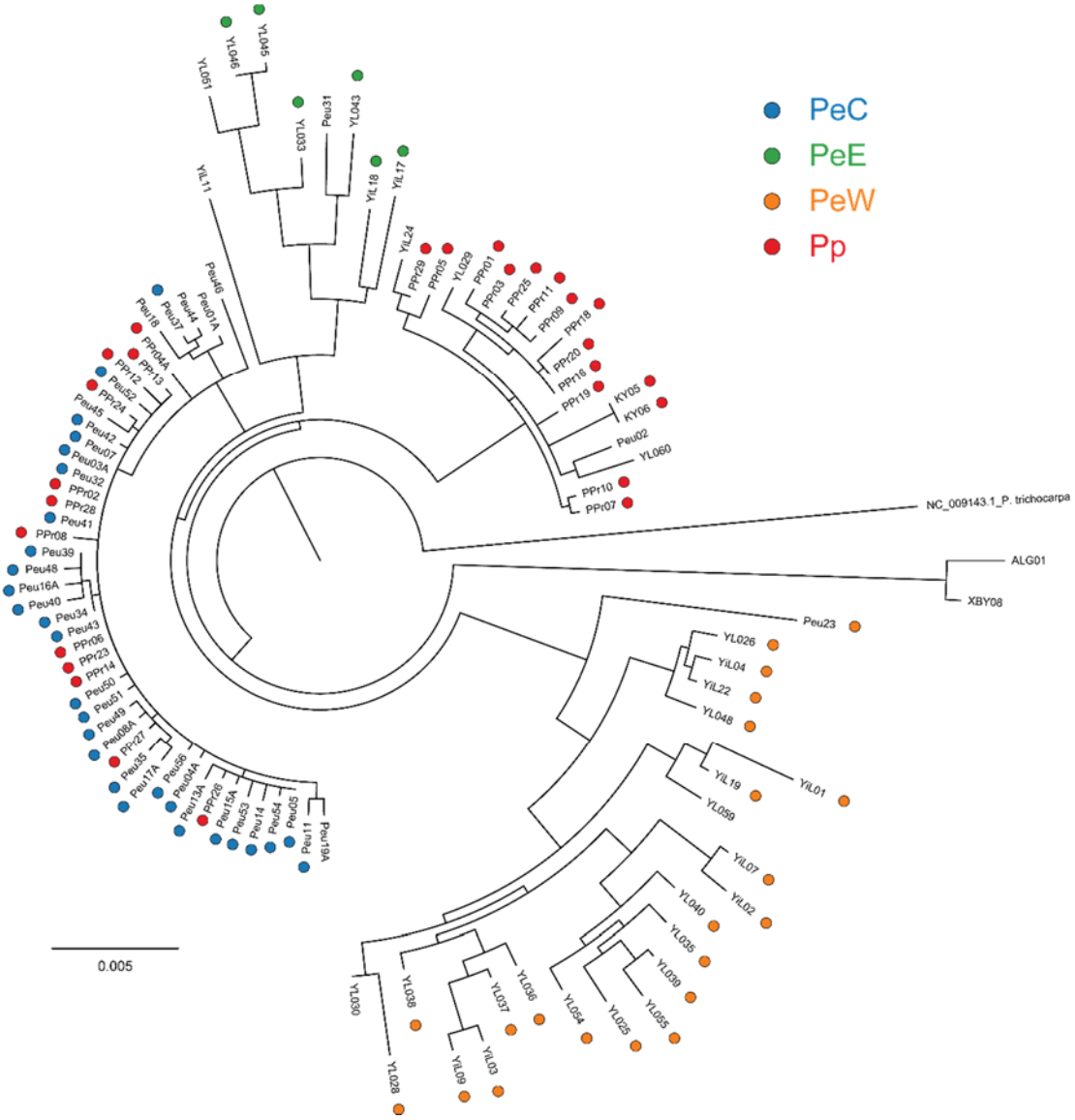


Fig. S4: Maximum-likelihood phylogenetic tree reconstructed from chloroplast genome, using *P. trichocarpa* (NC_009143.1) as outgroup.

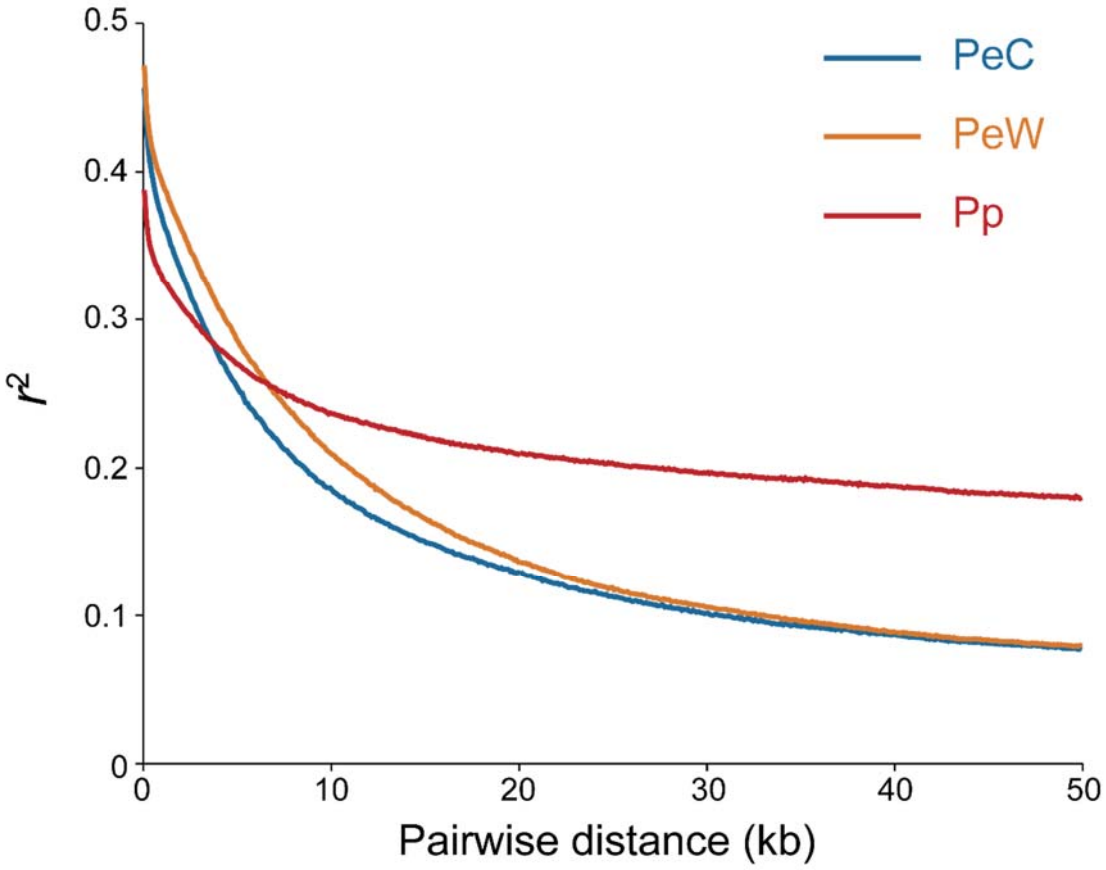


Fig. S5: Decay of linkage disequilibrium (LD), measured by r^2, was calculated for the lineages with at least 21 individuals sequenced. The genome-wide average r^2 dropped to 0.187 ± 0.249 (mean ± standard deviation), 0.210 ± 0.281 and 0.237 ± 0.308 at 10 kb, 0.129 ± 0.192, 0.137 ± 0.216 and 0.209 ± 0.287 at 20 kb for PeC, PeW and Pp, respectively.

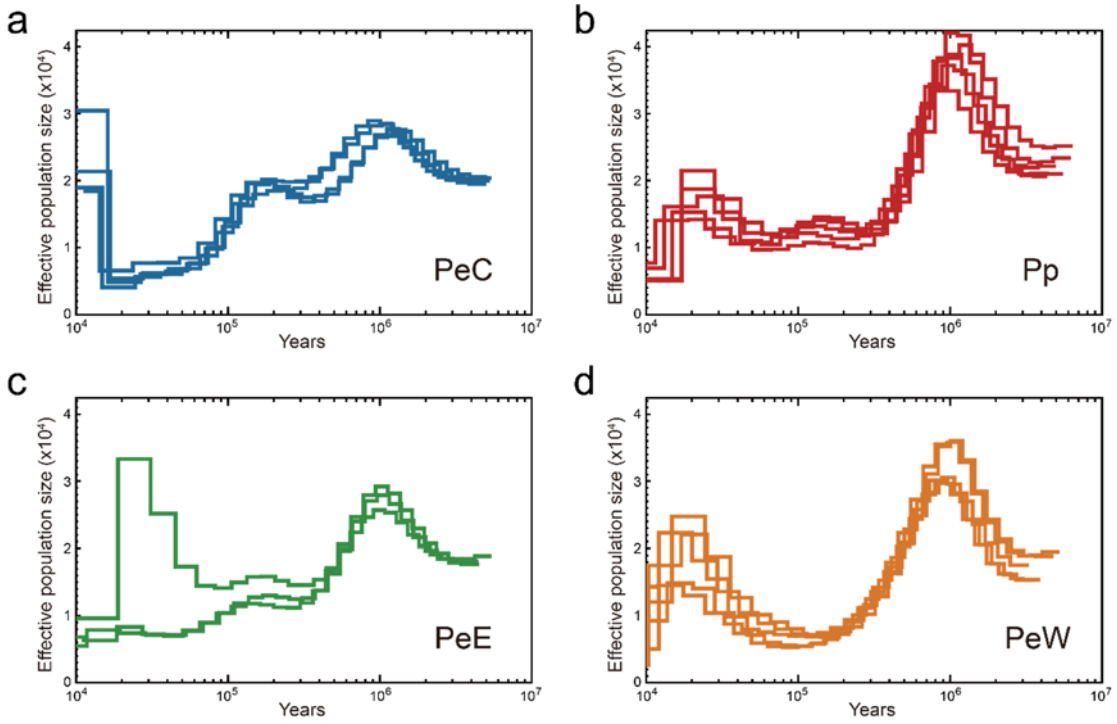


Fig. S6: Changes in effective population size (*Ne*) through time inferred by PSMC for lineages PeC (**a**), Pp (**b**), PeE (**c**) and PeW (**d**). The individuals with high sequencing coverage (>15 ×) were selected for each lineage.

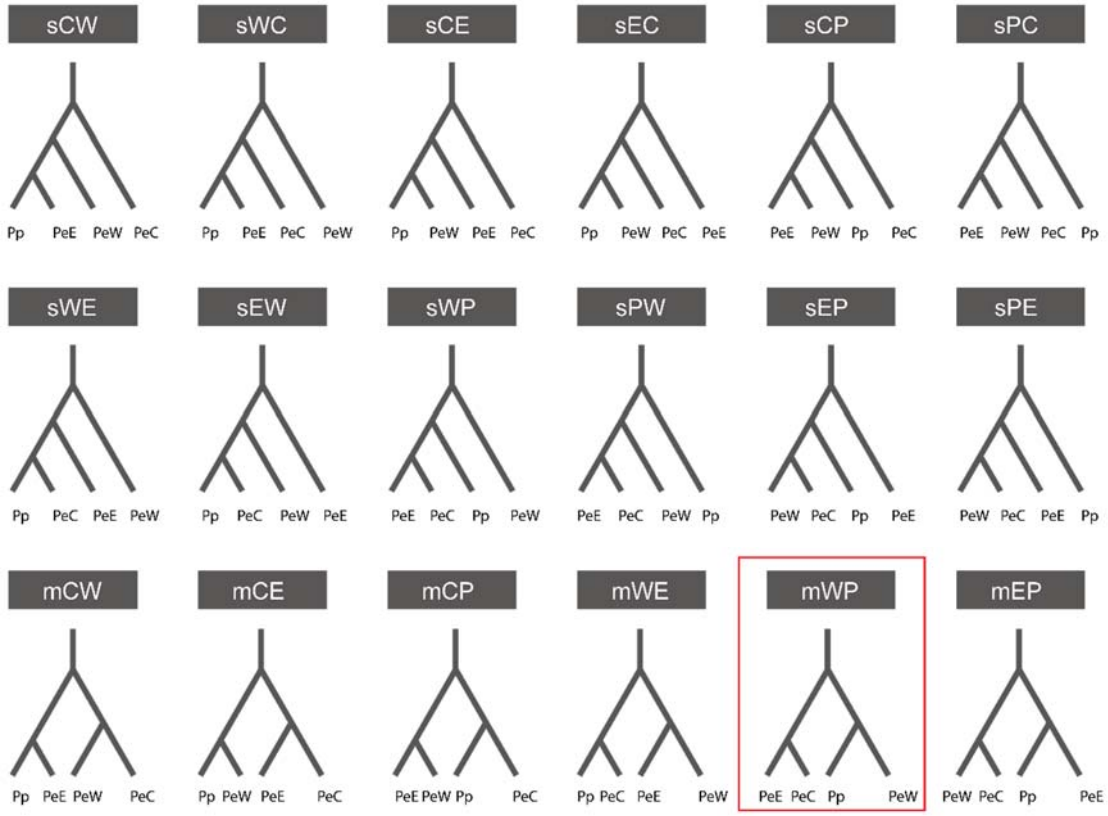


Fig. S7: Schematic diagram of all possible topological structures of these four lineages used in fastsimcoal2 to infer demographic parameters. For each topological structure, the parameters of gene flow, divergence time and effective population sizes were flexible, then we performed parameter estimation for 50 independent runs and chose the model with the highest likelihood. Note that the topological structure of model 'mWP' was best-supported according the value of the likelihoods and Akaike's information criterion.

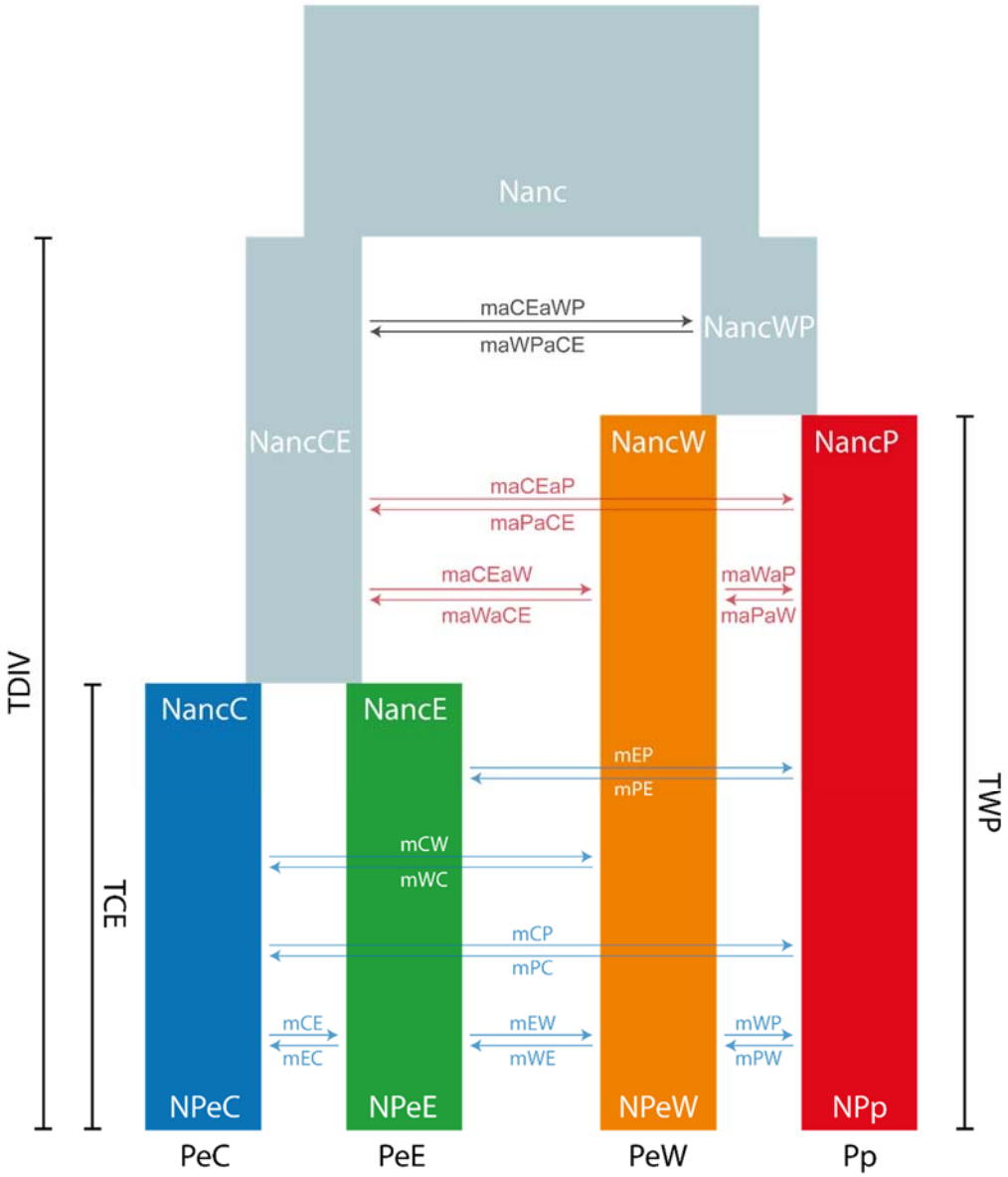


Fig. S8: Schematic diagram of the best-fitting model inferred by fastsimcoal2 with the corresponding parameter tags. Note that this model is a modified version of model 'mWP' (Fig. S6) with exponential change in population size following their divergence. The point estimates and 95% confidence intervals of demographic parameters are shown in Table S4.

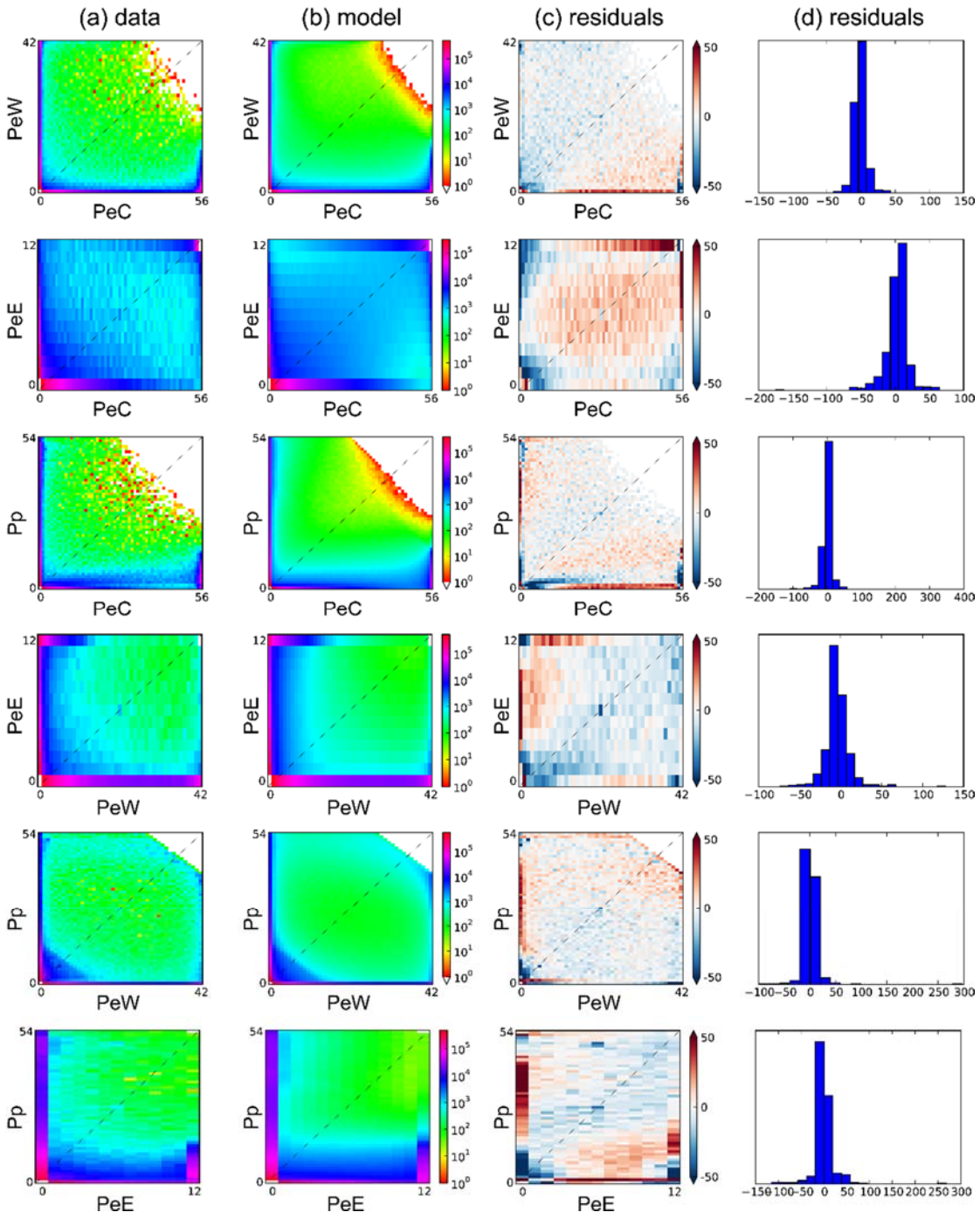


Fig. S9: Assessment of model fit for the best-fitting demographic model for all pairwise comparisons. (**a**) Observed joint SFS for each pair of lineages studied. (**b**) Expected joint SFS under the global ML parameters inferred by fastsimcoal2.1 for this model. (**c**) Model residuals across joint SFS. (**d**) Histograms of model residuals.

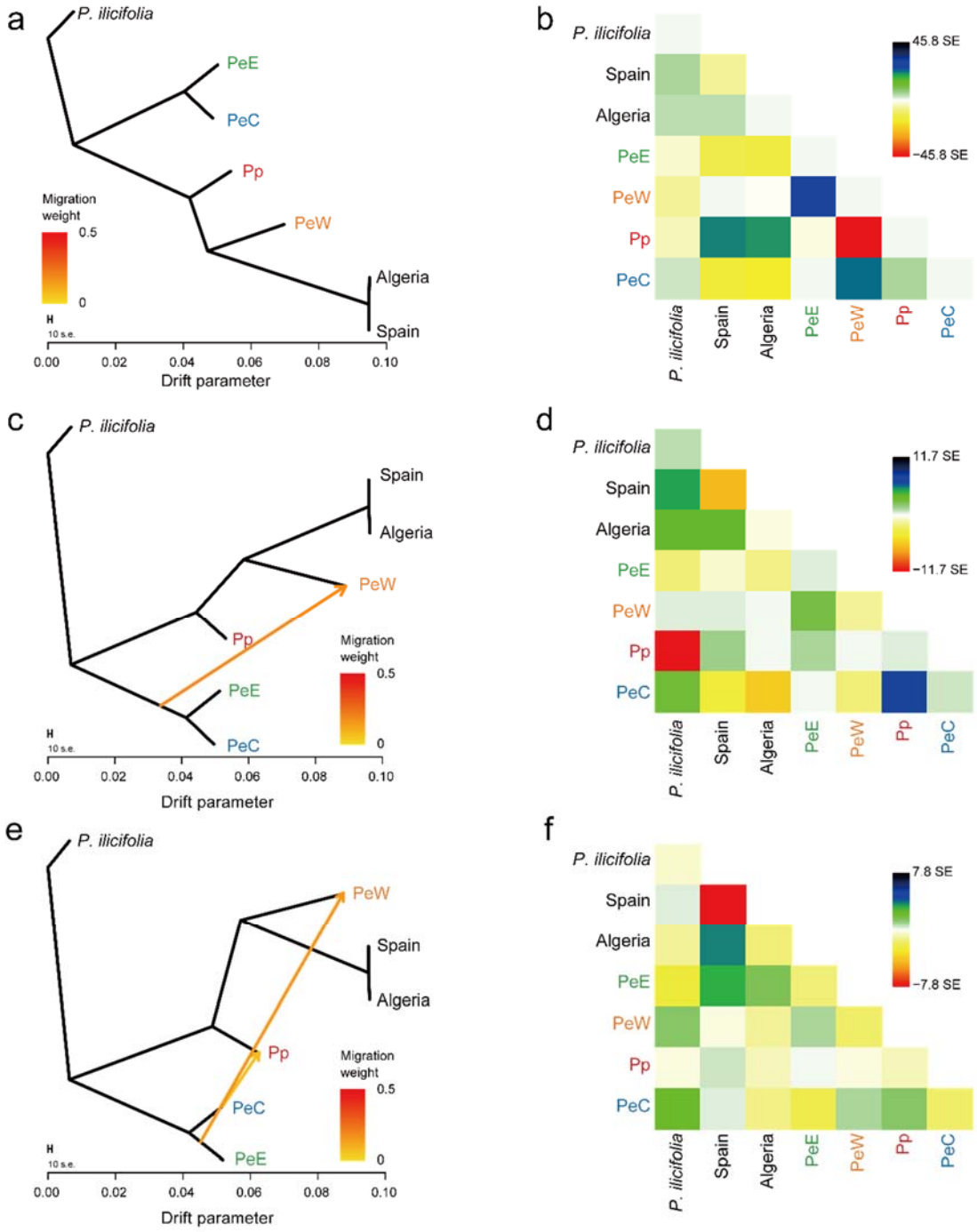
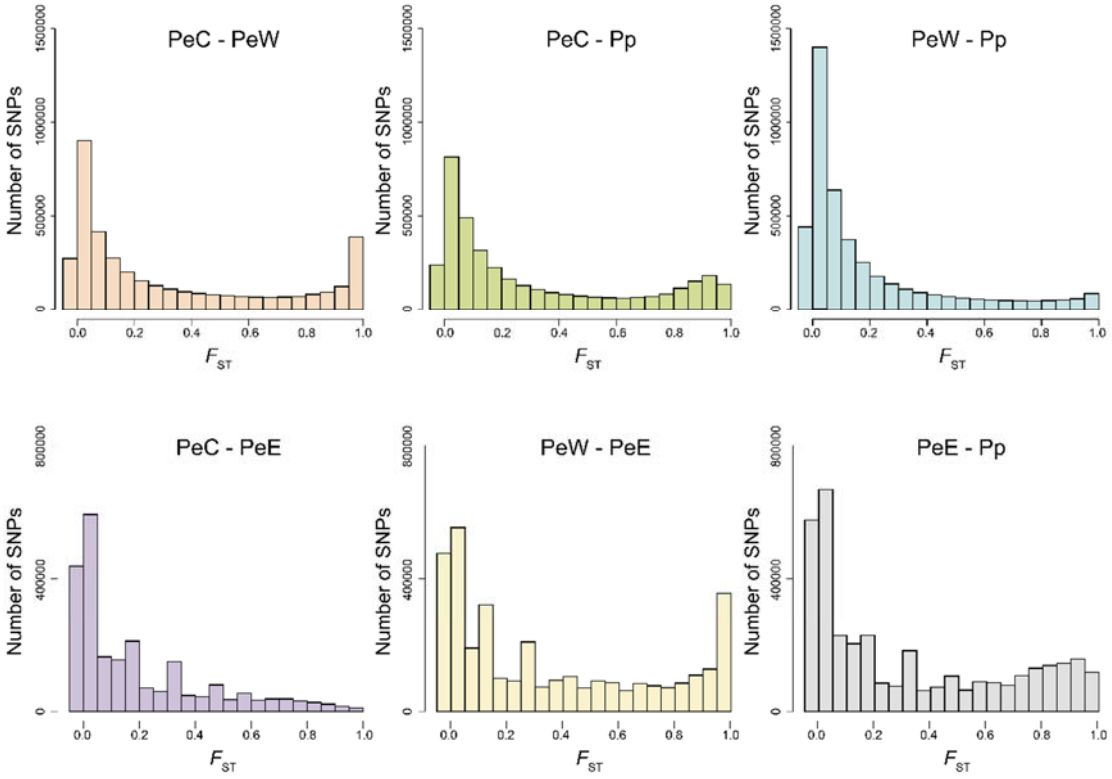
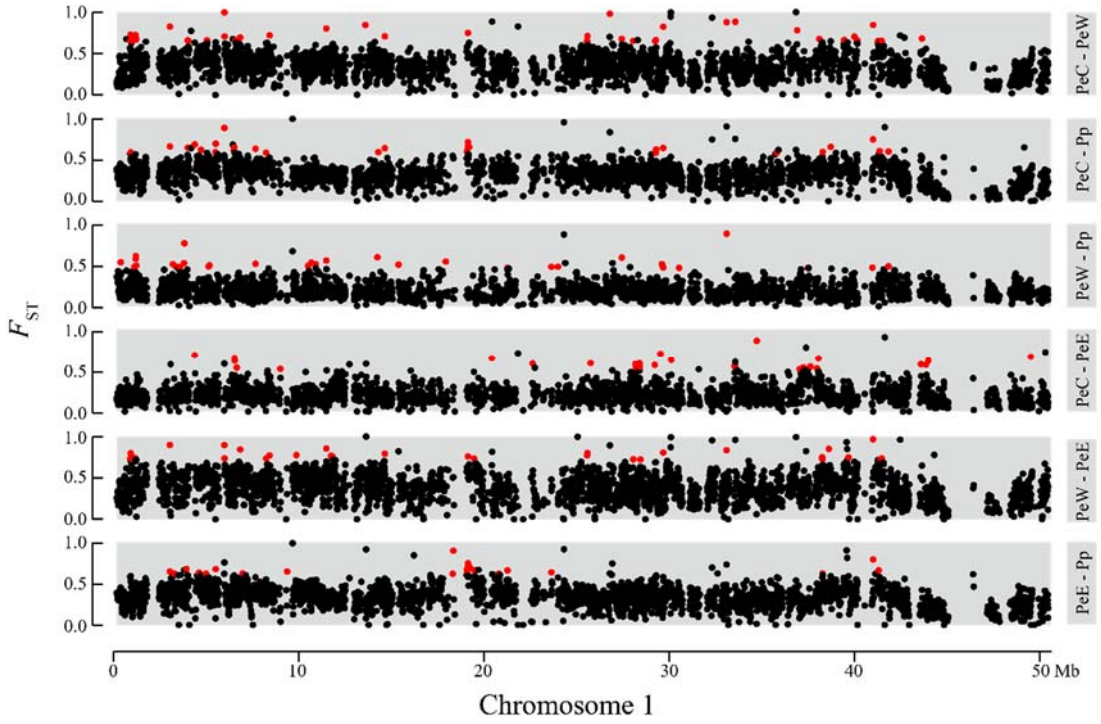


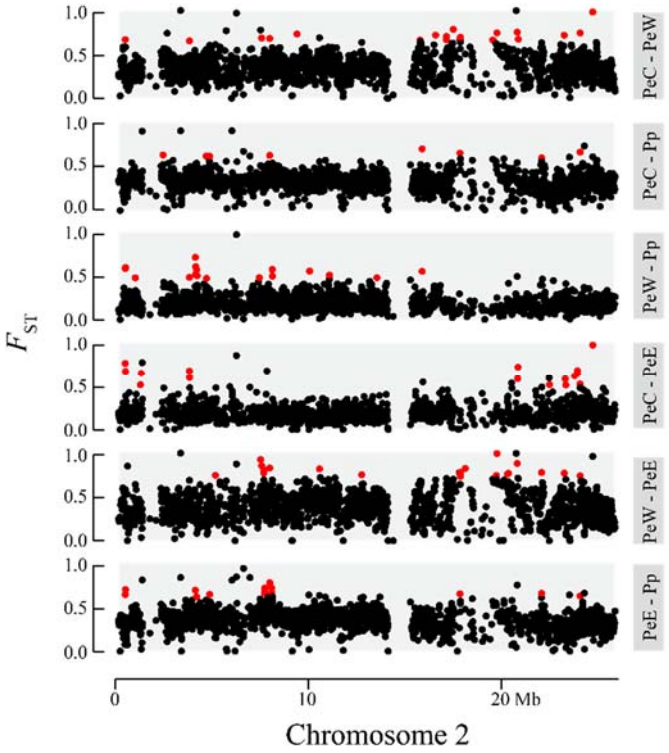
Fig. S10: The maximum likelihood tree for populations of *P. euphratica* and *P. pruinosa* inferred by TreeMix rooted with *P. ilicifolia*, allowing no (**a**), one (**c**) or two (**e**) migration events, respectively. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix. Migration edges are depicted as arrows colored by migration weight. The residual fit from these graphs are shown in (**b**), (**d**) and (**f**), respectively.

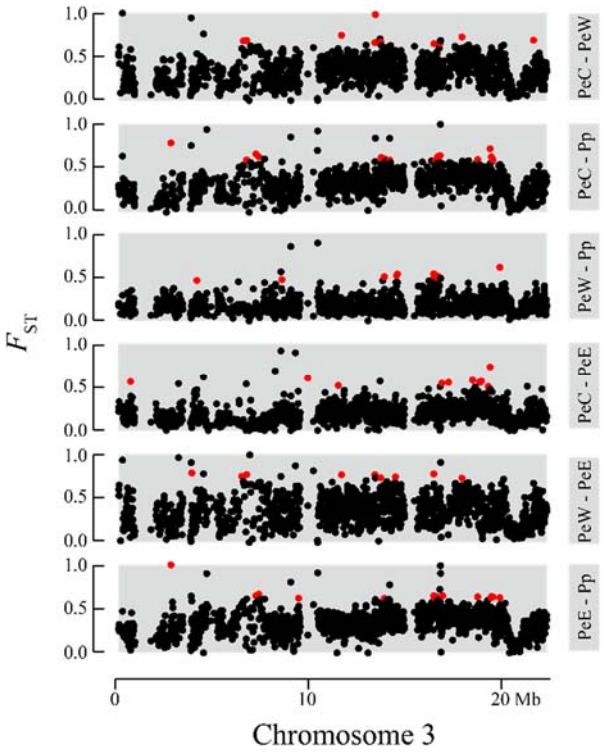


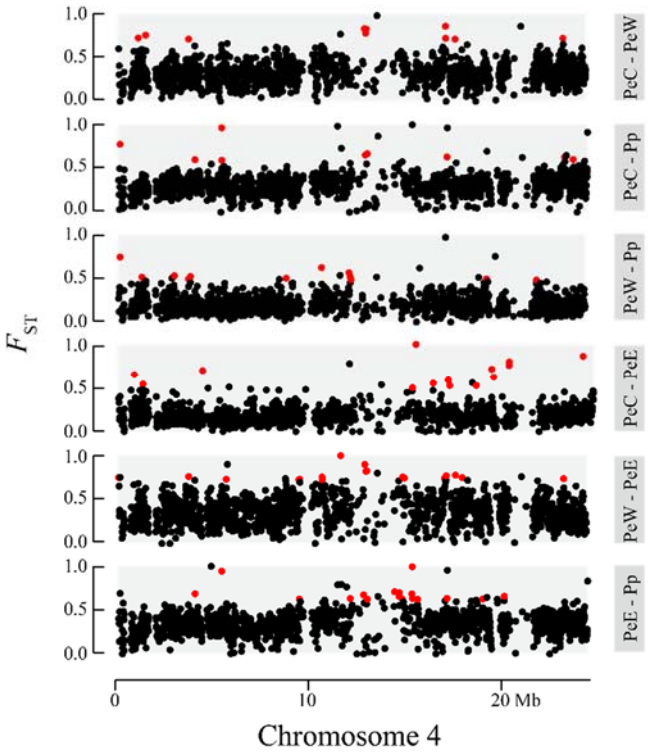
	PeC - PeW	PeC - Pp	PeW - Pp	PeC - PeE	PeW - PeE	PeE - Pp
PeC - PeW	-	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16
PeC - Pp	0.0656	-	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16
PeW - Pp	0.1786	0.1606	-	< 2.2e-16	< 2.2e-16	< 2.2e-16
PeC - PeE	0.2448	0.2268	0.1977	-	< 2.2e-16	< 2.2e-16
PeW - PeE	0.1334	0.1369	0.2559	0.2084	-	< 2.2e-16
PeE - Pp	0.1558	0.1369	0.2181	0.1782	0.0739	-

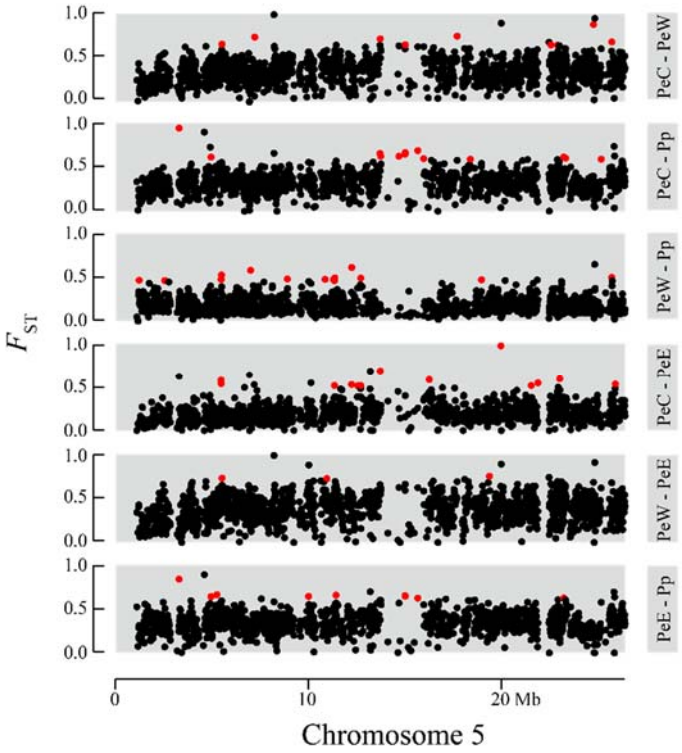
Fig. S11: F_{ST} distributions for each pair of lineages. The differences of the distribution shape between lineage pairs were determined by Kolmogorov-Smirnov test. The values of the Kolmogorov-Smirnov statistic and *p*-values were showed below and above the diagonal, respectively.

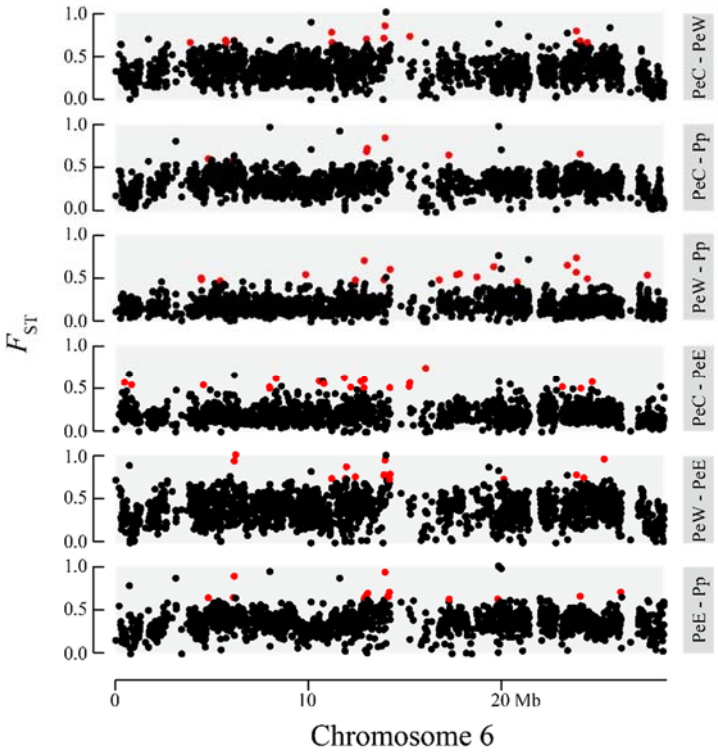

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20
30
40
50 Mb
Chromosome 1

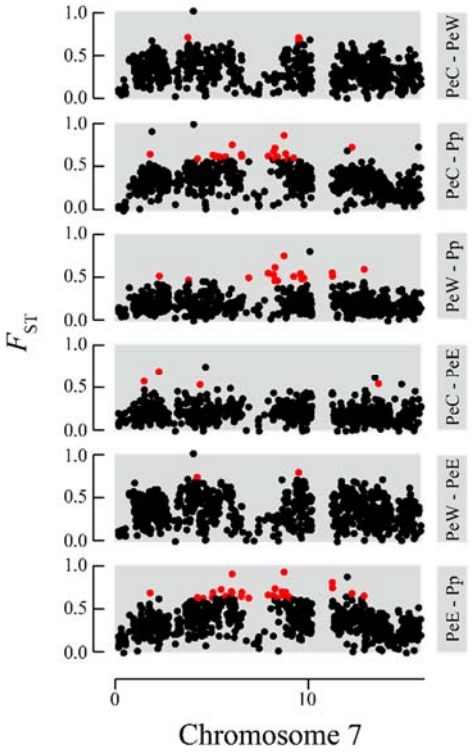

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20 Mb
Chromosome 2

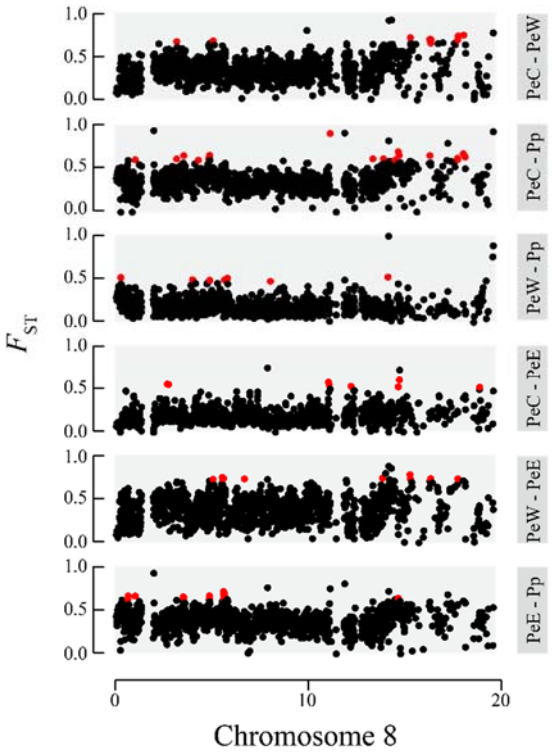

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20 Mb
Chromosome 3

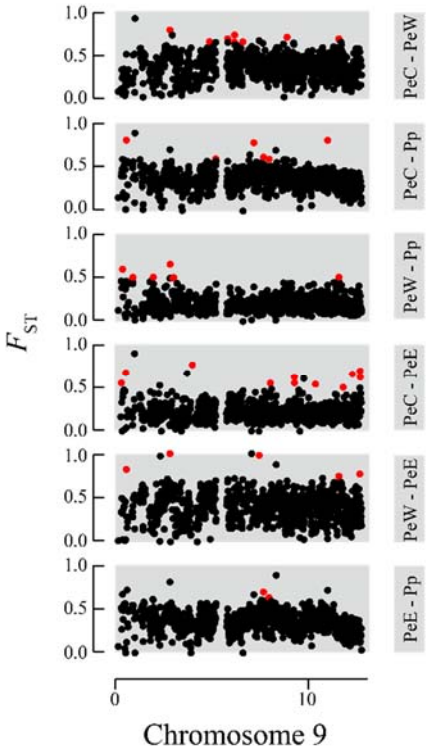

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20 Mb
Chromosome 4

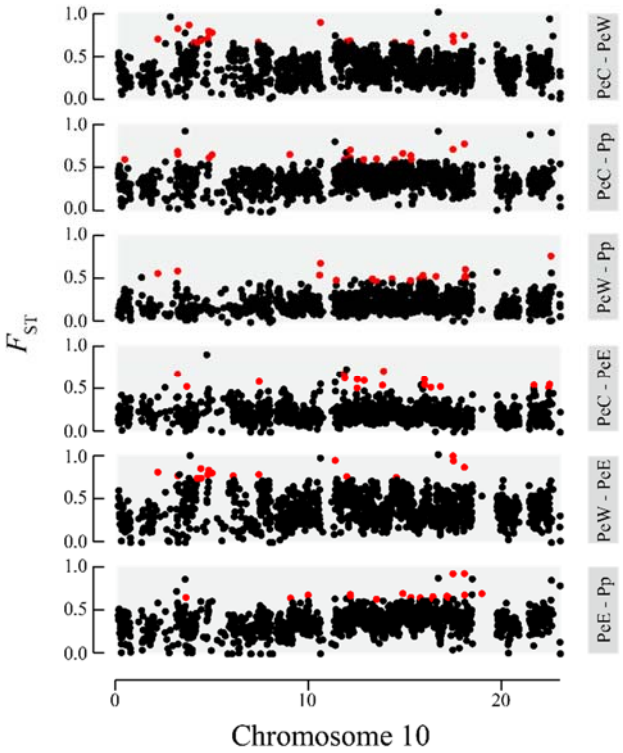

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20 Mb
Chromosome 5

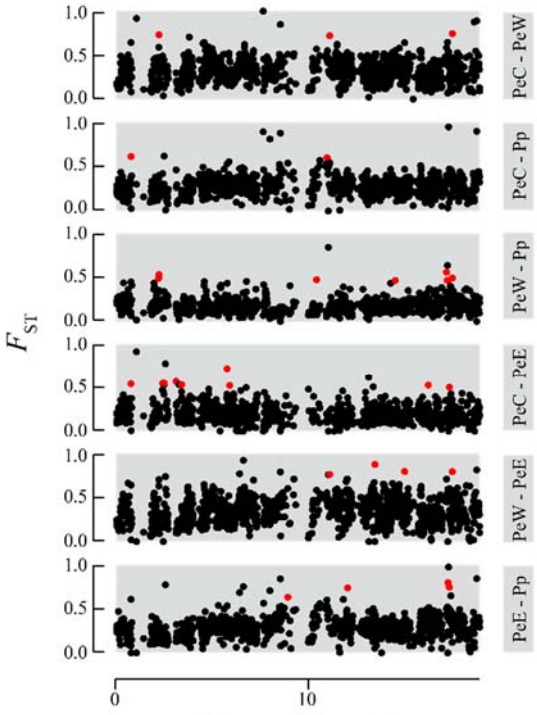

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20 Mb
Chromosome 6

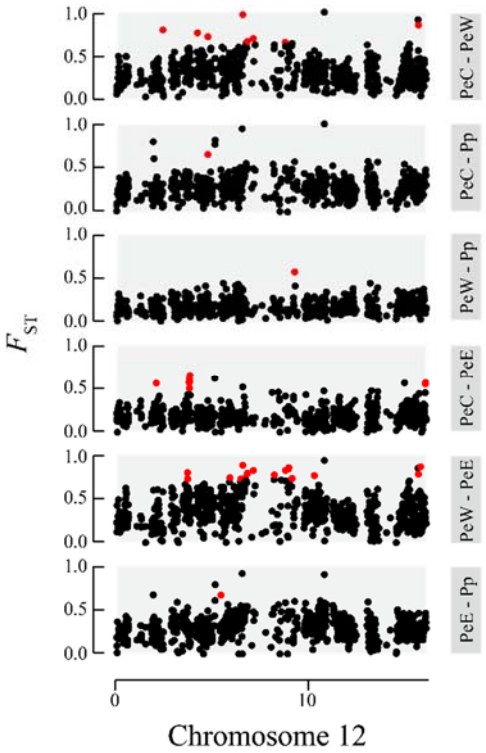

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 7

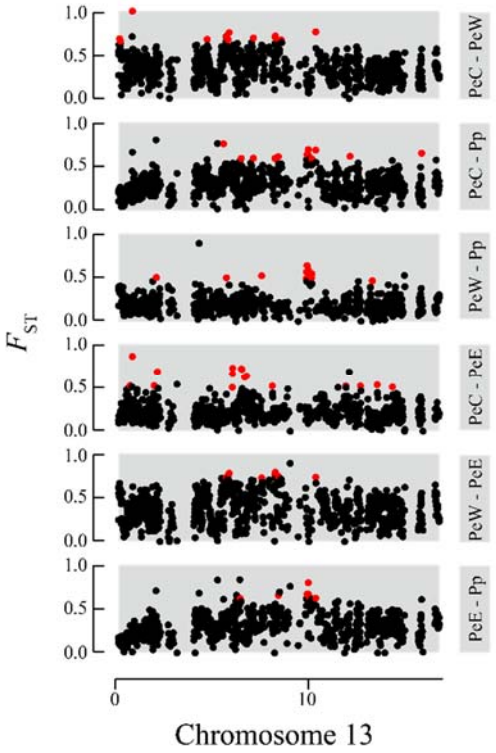

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20
Chromosome 8

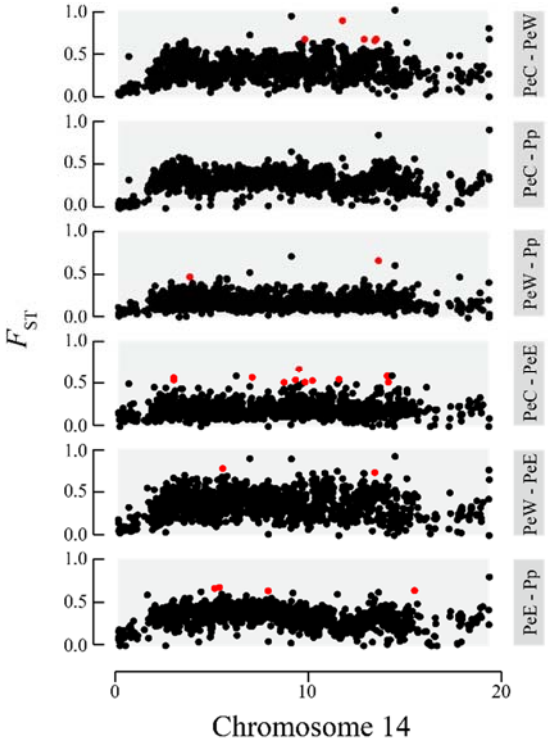

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 9

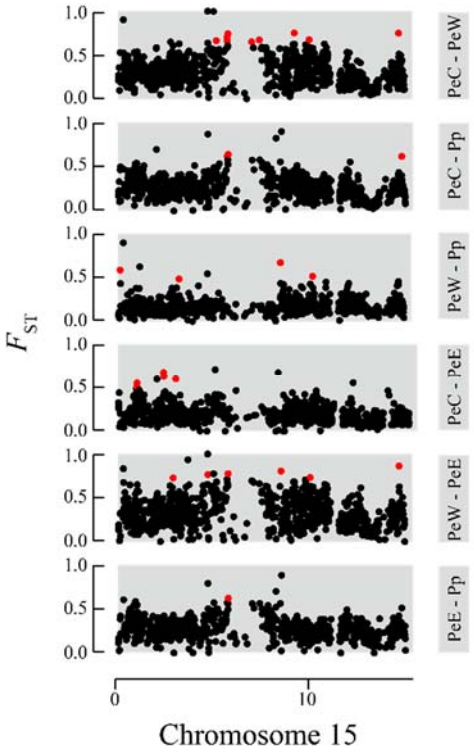

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20
Chromosome 10

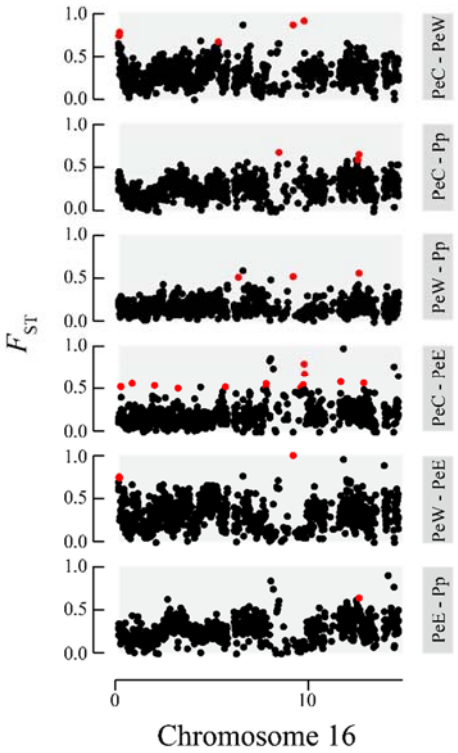

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 11

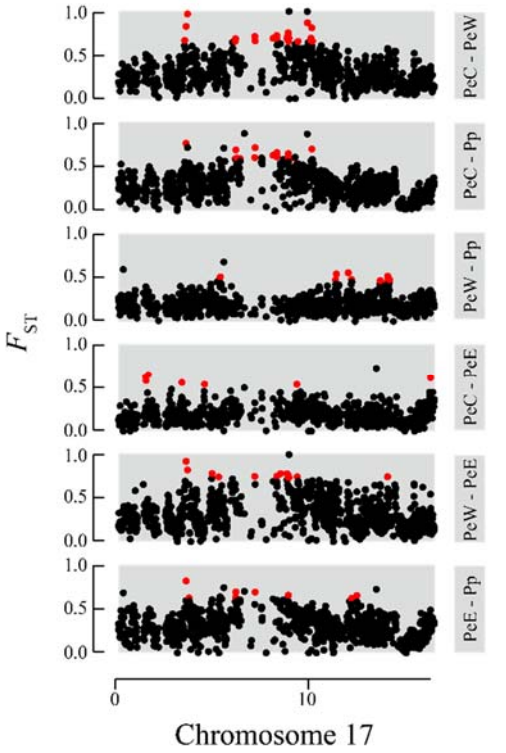

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 12

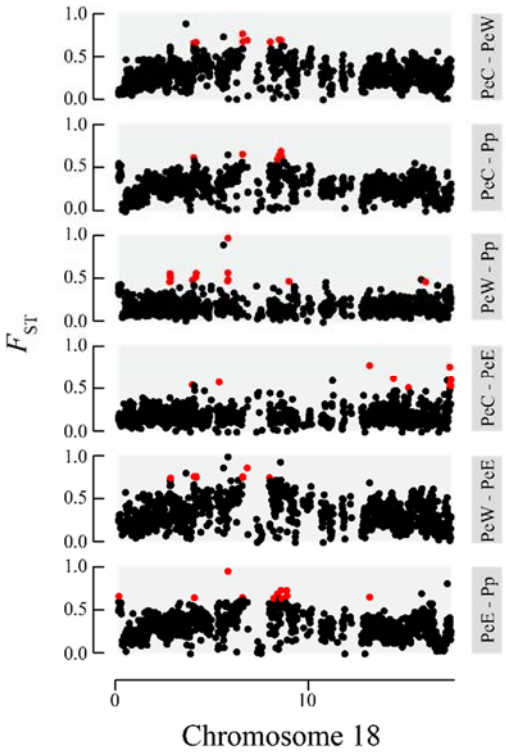

F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 13


F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
20
Chromosome 14


F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 15


F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 16


F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 17


F_{ST}
1.0
0.5
0.0
PeC - PeW
PeC - Pp
PeW - Pp
PeC - PeE
PeW - PeE
PeE - Pp
0
10
Chromosome 18

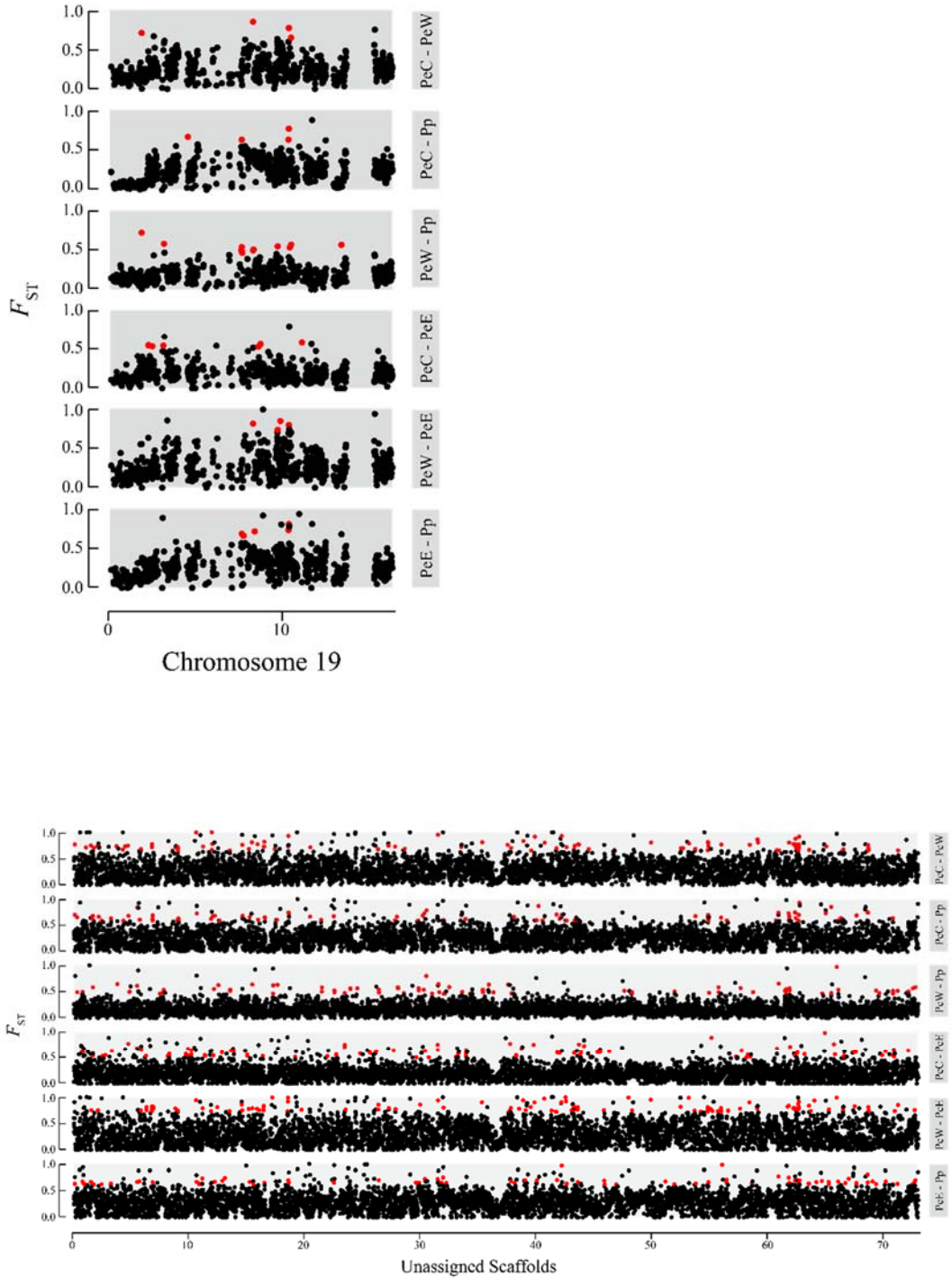


Fig. S12: Pairwise genetic divergence (F_{ST}) in 10-kb sliding windows across all chromosomes for all comparisons. Genomic islands of divergence are shown in red.

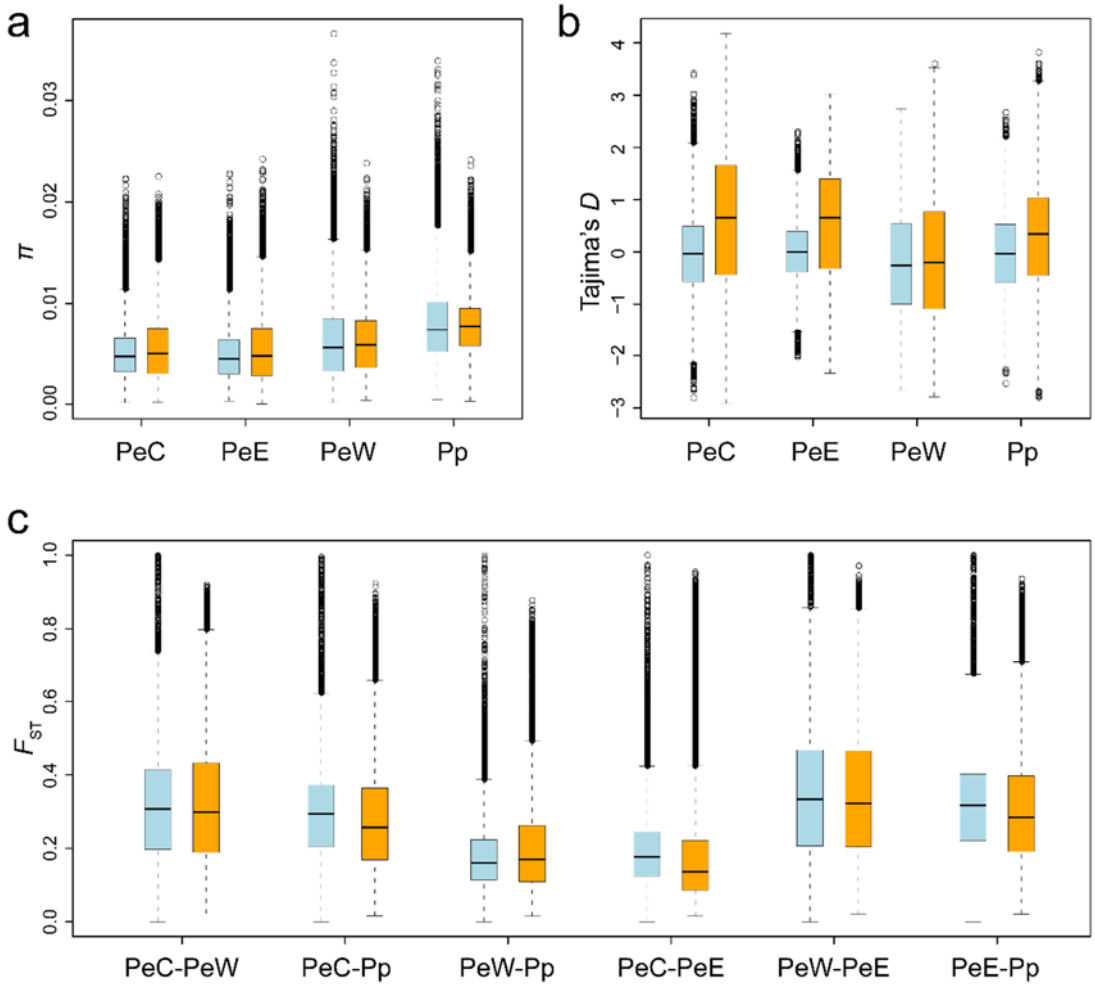


Fig. S13: Comparisons of (**a**) nucleotide diversity (π), (**b**) Tajima's D and (**c**) genetic differentiation (F_{ST}) between the observed (blue) and simulated (orange) data using *msms* under the best-fitting demographic model inferred by fastsimcoal2.

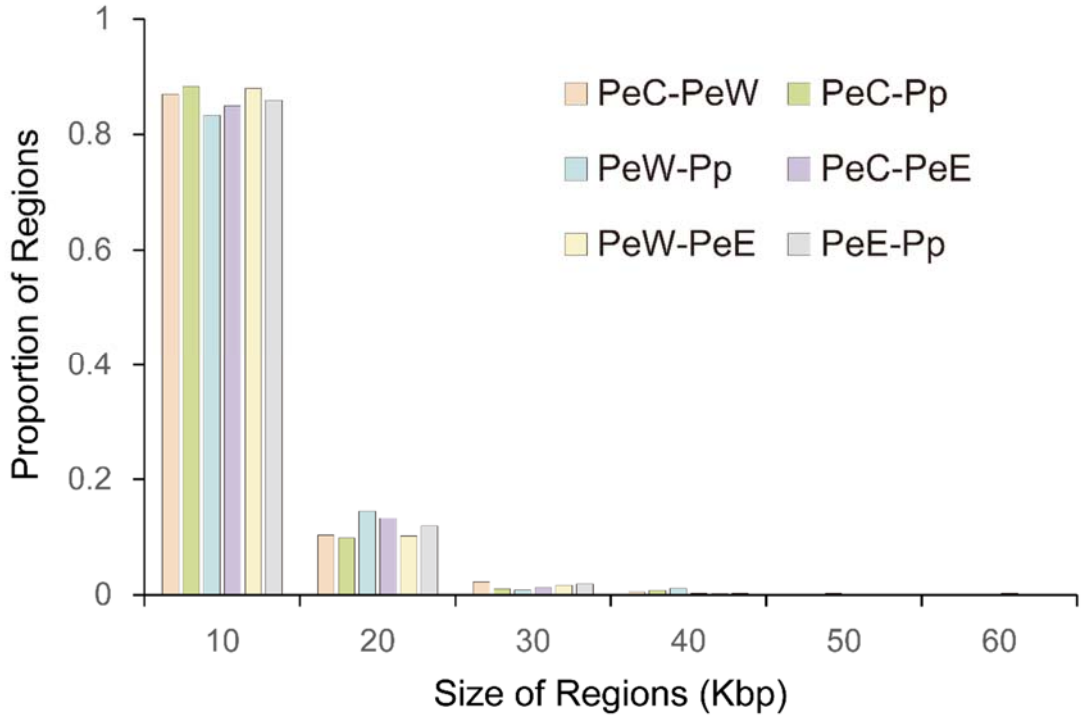


Fig. S14: The size distributions of genomic islands displaying significantly high genetic differentiation for all pairwise comparisons.

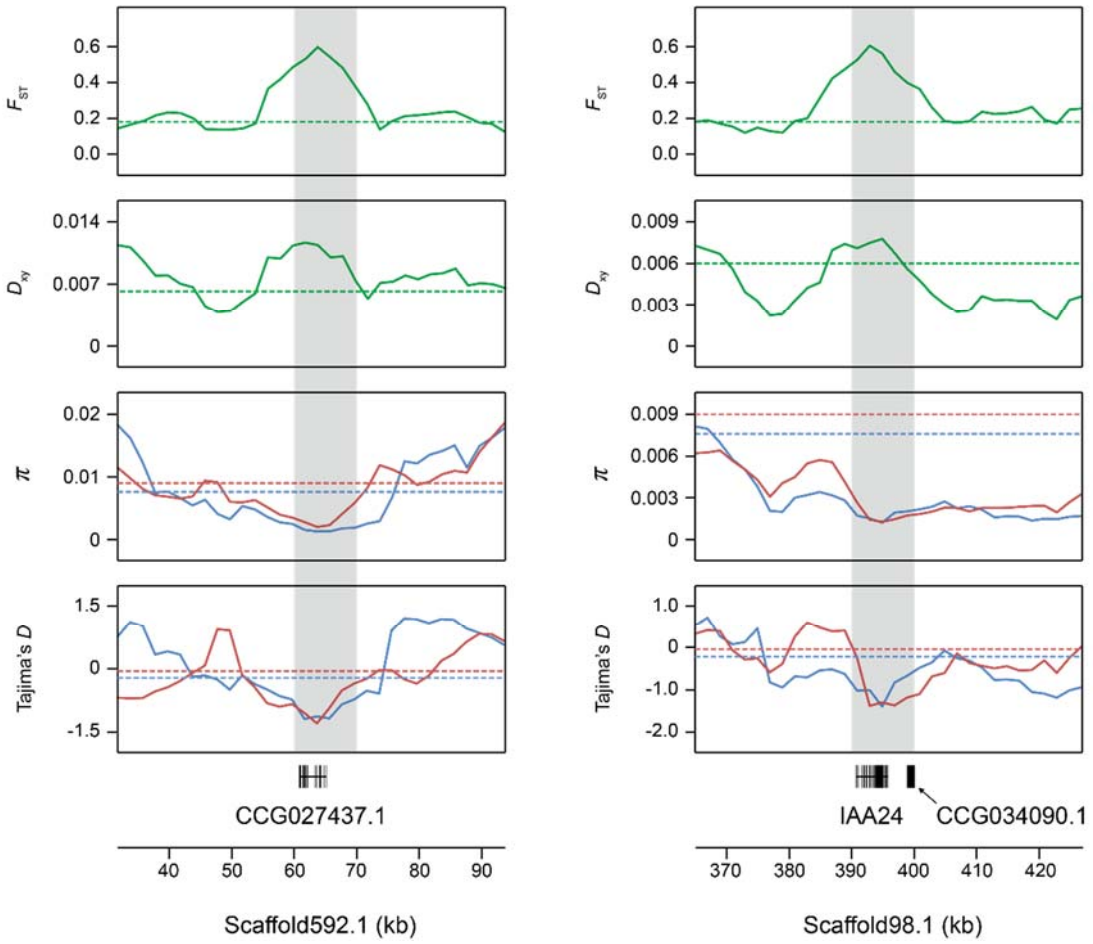


Fig. S15: Distribution of population genomic parameters around two examples of islands (gray bar) between lineages PeW (blue) and Pp (red). F_{ST}, D_{xy}, π and Tajima's *D* values are plotted using a 10 kb sliding window with 2 kb steps. Horizontal dashed lines represent mean whole-genome of corresponding values. Positively selected genes within the islands are shown at the bottom.

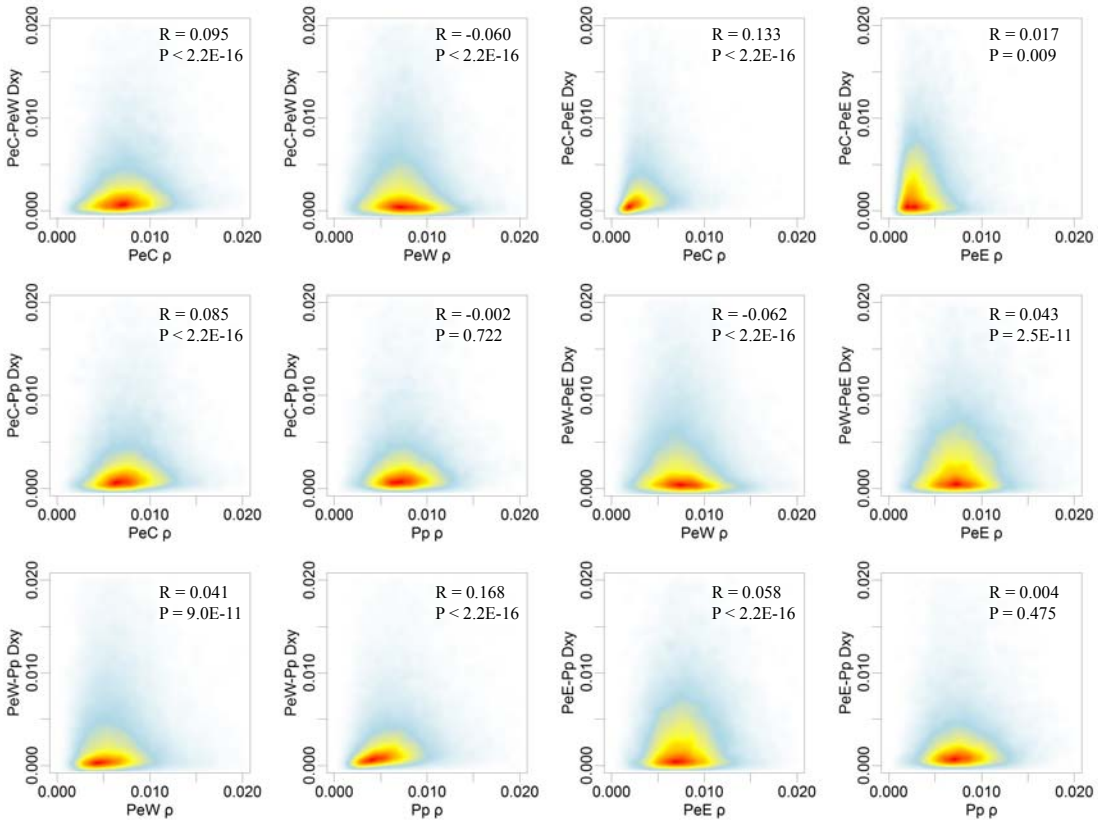


Fig. S16: Correlation of population-scaled recombination rates (ρ) and D_{xy}. Scatter plots display genome-wide values of two variables in dots over 10 Kbp non-overlapping windows. The red to yellow to blue gradient indicates decreased density of observed events at a given location in the graph. Correlation and significance are tested with Pearson's correlation test.

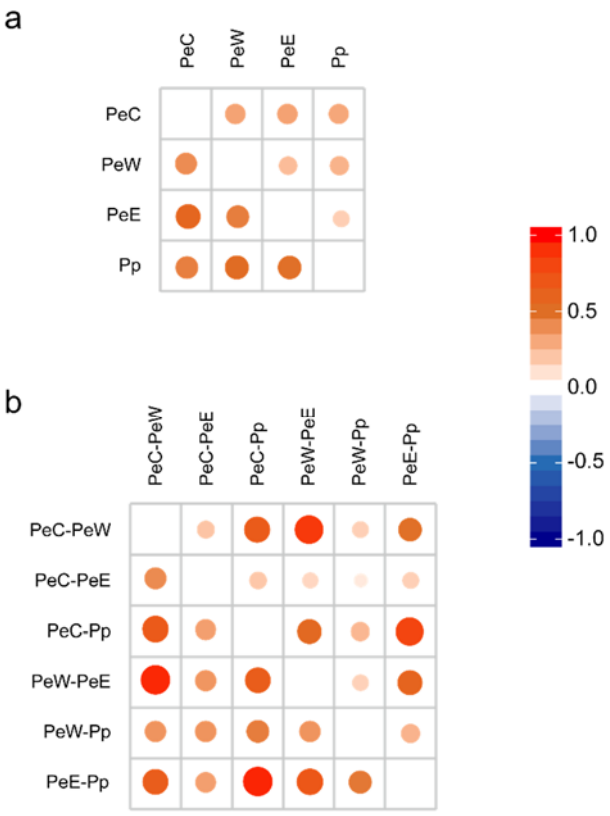


Fig. S17: Correlations of 10-kb window-based estimates of (**a**) intra-population summary statistics (above diagonal: population-scaled recombination rate ρ; below the diagonal: nucleotide diversity π) between populations and (**b**) inter-populations statistics (above diagonal: genetic divergence F_{ST}; below the diagonal: absolute divergence D_{xy}) of population pairs. Red circles indicate a positive correlation, blue a negative one. The color intensity and circle size are proportional to Spearman's correlation coefficient.

SI Tables:

Table S1: Overview of sample information and sequencing statistics.

Sample ID	Species	Country/Region	Latitude	Longitude	Clean Reads (Gbp)	Map Ratio (%)	Genome Coverage (%)	Effective Depth
PPr01	*Populus pruinosa*	Xinjiang, China	N40°17'	E80°21'	7.33	91.37	89.8	13.54
PPr02	*Populus pruinosa*	Xinjiang, China	N40°17'	E80°21'	3.64	94.97	82.48	7.08
PPr03	*Populus pruinosa*	Xinjiang, China	N40°19'	E79°55'	14.14	91.65	91.57	26.07
PPr04A	*Populus pruinosa*	Xinjiang, China	N40°19'	E79°55'	4.82	93.13	87.91	9.21
PPr05	*Populus pruinosa*	Xinjiang, China	N39°22'	E78°11'	7.98	86.16	90.27	13.78
PPr06	*Populus pruinosa*	Xinjiang, China	N39°22'	E78°11'	4.37	95.68	82.09	8.61
PPr07	*Populus pruinosa*	Xinjiang, China	N39°21'	E78°05'	8.54	83.1	90.47	14.23
PPr08	*Populus pruinosa*	Xinjiang, China	N39°21'	E78°05'	3.46	96.52	82.13	6.85
PPr09	*Populus pruinosa*	Xinjiang, China	N38°02'	E76°58'	11.25	92.98	91.24	21.09
PPr10	*Populus pruinosa*	Xinjiang, China	N38°02'	E76°58'	4.18	94.91	85.02	8.14
PPr11	*Populus pruinosa*	Xinjiang, China	N37°34'	E79°38'	5.81	95.5	87.87	11.38
PPr12	*Populus pruinosa*	Xinjiang, China	N37°11'	E82°47'	3.26	93.95	83.75	6.29
PPr13	*Populus pruinosa*	Xinjiang, China	N37°11'	E82°47'	7.93	93.33	89.89	14.83
PPr14	*Populus pruinosa*	Xinjiang, China	N41°15'	E84°12'	6.3	95.03	87.51	12.28
PPr15	*Populus pruinosa*	Xinjiang, China	N41°15'	E84°12'	4.01	94.05	86.32	7.74
PPr16	*Populus pruinosa*	Xinjiang, China	N41°00'	E83°20'	9.58	91.77	91.08	17.58
PPr17	*Populus pruinosa*	Xinjiang, China	N41°00'	E83°20'	5.59	95.17	86.01	10.9
PPr18	*Populus pruinosa*	Xinjiang, China	N41°00'	E83°20'	3.18	94.17	81.33	6.09
PPr19	*Populus pruinosa*	Xinjiang, China	N40°43'	E82°02'	8.66	93.1	90.09	16.15
PPr20	*Populus pruinosa*	Xinjiang, China	N38°02'	E76°58'	4.42	95.51	85.57	8.61
PPr21	*Populus pruinosa*	Xinjiang, China	N38°02'	E76°58'	2.82	94.8	81.86	5.44
PPr23	*Populus pruinosa*	Xinjiang, China	N40°43'	E82°02'	3.08	95.09	80.35	6
PPr24	*Populus pruinosa*	Xinjiang, China	N40°43'	E82°02'	2.92	95.78	75.4	5.76
PPr25	*Populus pruinosa*	Xinjiang, China	N40°43'	E82°02'	4.37	95.69	78.85	8.62
PPr26	*Populus pruinosa*	Xinjiang, China	N40°42′	E81° 40′	3.22	94.11	86.09	6.19
PPr27	*Populus pruinosa*	Xinjiang, China	N40°42′	E81° 40′	3.7	93.4	87.26	7.05
PPr28	*Populus pruinosa*	Xinjiang, China	N40°26′	E80°56′	3.55	93.38	87.18	6.78
PPr29	*Populus pruinosa*	Xinjiang, China	N40°26′	E80°56′	3.47	94.5	85.91	6.72
KY01	*Populus pruinosa*	Kyrgystan	N41°19'	E72°13'	13.17	92.29	86.93	24.05
KY02	*Populus pruinosa*	Kyrgystan	N41°19'	E72°13'	8.9	93.97	85.17	16.59
KY03	*Populus pruinosa*	Kyrgystan	N41°19'	E72°13'	6.83	93.11	80.86	12.83
KY04	*Populus pruinosa*	Kyrgystan	N41°19'	E72°13'	7.38	92.45	74.71	11.2
KY05	*Populus pruinosa*	Kyrgystan	N41°19'	E72°13'	14.64	93.52	86.66	26.8
KY06	*Populus pruinosa*	Kyrgystan	N41°19'	E72°13'	9.08	93.58	83.17	17.19
ALG01	*Populus euphratica*	Ghardaia, Algeria	N32°24'	E4°14'	14.74	90.58	88.22	27.26
ALG02	*Populus euphratica*	Ghardaia, Algeria	N32°24'	E4°14'	6.64	94.15	85.43	12.84
ALG03	*Populus euphratica*	Ghardaia, Algeria	N32°24'	E4°14'	6.53	94.66	85.82	12.67

ALG04	*Populus euphratica*	Ghardaia, Algeria	N32°24'	E4°14'	6.53	94.32	86.29	12.62
ALG05	*Populus euphratica*	Ghardaia, Algeria	N32°24'	E4°14'	6.61	94.9	86.26	12.81
XBY02	*Populus euphratica*	Alicante, Spain	N38°17'	W0°42'	3.67	90.28	82.38	6.78
XBY04	*Populus euphratica*	Alicante, Spain	N38°17'	W0°42'	4.47	88.27	81.97	8.14
XBY05	*Populus euphratica*	Alicante, Spain	N38°17'	W0°42'	4.05	91.11	82.95	7.54
XBY06	*Populus euphratica*	Alicante, Spain	N38°17'	W0°42'	6.03	92.71	84.67	11.47
XBY08	*Populus euphratica*	Alicante, Spain	N38°17'	W0°42'	14.05	92.21	88.36	26.68
XBY10	*Populus euphratica*	Alicante, Spain	N38°17'	W0°42'	3.93	93.75	82.77	7.55
XBY11	*Populus euphratica*	Alicante, Spain	N38°17'	W0°42'	5.97	89.74	85.26	10.98
XBY13	*Populus euphratica*	Alicante, Spain	N38°17'	W0°42'	4.38	92.46	84.01	8.26
Peu20	*Populus euphratica*	Israel	N30°49'	E34°46'	6.24	77.3	86.59	9.86
Peu21	*Populus euphratica*	Israel	N30°49'	E34°46'	5.5	93.65	85.27	10.61
Peu22	*Populus euphratica*	Israel	N30°49'	E34°46'	7.41	88.55	90.02	13.51
Peu23	*Populus euphratica*	Israel	N30°49'	E34°45'	10.81	89.75	91.29	19.55
Peu24	*Populus euphratica*	Israel	N30°49'	E34°45'	7.4	90.39	90.18	13.74
Peu26	*Populus euphratica*	Israel	N30°49'	E34°45'	3.86	88.16	87.41	6.98
Peu27	*Populus euphratica*	Israel	N30°49'	E34°45'	5.14	88.66	88.83	9.35
Peu28	*Populus euphratica*	Israel	N30°49'	E34°45'	5.52	50.57	86.01	5.58
YiL11	*Populus euphratica*	Azarbayjan-e Sharqi, Iran	N38°50'	E46°8'	6.11	95.47	91.68	12.08
YiL12	*Populus euphratica*	Azarbayjan-e Sharqi, Iran	N38°52'	E46° 5'	4.55	95.67	87.3	9
YL058	*Populus euphratica*	Azarbayjan-e Sharqi, Iran	N38°50'	E46°7'	6.42	94.76	87.25	12.41
YL059	*Populus euphratica*	Azarbayjan-e Sharqi, Iran	N38°52'	E46°3'	5.83	94.09	84.86	11.05
YL060	*Populus euphratica*	Azarbayjan-e Sharqi, Iran	N38°50'	E46°9'	5.75	96.43	83.42	11.29
YiL03	*Populus euphratica*	Bushehr, Iran	N28°55'	E50°50'	6.11	94.68	87.2	11.9
YiL04	*Populus euphratica*	Bushehr, Iran	N28°55'	E50°51'	5.1	97.12	85.99	10.12
YL052	*Populus euphratica*	Bushehr, Iran	N28°55'	E50°50'	4.3	93.12	75.06	8.24
YL053	*Populus euphratica*	Bushehr, Iran	N28°55'	E50°51'	5.01	93.73	78.13	9.66
YL054	*Populus euphratica*	Bushehr, Iran	N28°55'	E50°51'	6.58	94.14	82.56	12.74
YiL07	*Populus euphratica*	Esfahan, Iran	N34°12'	E51°31'	6.01	93.31	88.72	11.52
YiL08	*Populus euphratica*	Esfahan, Iran	N34°12'	E51°31'	4.88	95.79	85.71	9.58
YL040	*Populus euphratica*	Esfahan, Iran	N34°7'	E51°29'	13.86	90.92	86.33	25.08
YL041	*Populus euphratica*	Esfahan, Iran	N34°12'	E51°31'	6.32	93.16	81.35	11.96
YL042	*Populus euphratica*	Esfahan, Iran	N34°12'	E51°31'	6.08	94.62	82.59	11.66
YiL01	*Populus euphratica*	Gilan, Iran	N36°40'	E49°25'	5.75	96.88	87.01	11.4
YiL02	*Populus euphratica*	Gilan, Iran	N36°40'	E49°25'	6	93.54	88.01	11.55
YL055	*Populus euphratica*	Gilan, Iran	N36°40'	E49°25'	4.67	96.38	78.36	9.1
YL056	*Populus euphratica*	Gilan, Iran	N36°14'	E49°26'	4.72	95.24	79.16	9.18
YL057	*Populus euphratica*	Gilan, Iran	N36°13'	E49°26'	6.1	96.4	82.61	11.97
YiL13	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°16'	E59°47'	5.95	96.54	92.2	12.03
YiL14	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°16'	E59°47'	6.13	95.14	91.81	12.23
YiL17	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°18'	E61°9'	5.96	95.49	91.66	11.94
YiL18	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°18'	E61° 8'	5.9	96.15	91.94	11.9
YL033	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°16'	E59°47'	14.91	92.98	91.57	28.57
YL034	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°16'	E59°47'	7.25	94.1	86.41	14.07

YL043	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°17'	E61°9'	13.63	91.58	86.75	21.81
YL044	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°17'	E61°9'	4.95	95.92	84.02	9.72
YL045	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°19'	E61°8'	5.81	95.23	86.85	11.36
YL046	*Populus euphratica*	Khorasan-e Razavi, Iran	N36°17'	E61°9'	4.99	95.53	85.92	9.99
YiL23	*Populus euphratica*	Khorasan-e Shemali, Iran	N37°59'	E56°18'	4.75	97.21	86.58	9.47
YiL24	*Populus euphratica*	Khorasan-e Shemali, Iran	N37°59'	E56°18'	5.95	93.14	89.6	11.44
YL029	*Populus euphratica*	Khorasan-e Shemali, Iran	N37°55'	E56°10'	4.99	96.53	83.04	9.72
YL030	*Populus euphratica*	Khorasan-e Shemali, Iran	N37°59'	E56°17'	6.98	82.4	84.38	12.23
YL031	*Populus euphratica*	Khorasan-e Shemali, Iran	N39°46'	E56°16'	6.34	86.37	81.84	11.24
YL032	*Populus euphratica*	Khorasan-e Shemali, Iran	N37°58'	E56°17'	5.27	93.11	77.49	9.96
YiL19	*Populus euphratica*	Khuzestan, Iran	N32°15'	E48°20'	4.75	96.97	85.24	9.42
YL035	*Populus euphratica*	Khuzestan, Iran	N32°15'	E48°20'	6.85	93.39	82.61	12.92
YL036	*Populus euphratica*	Khuzestan, Iran	N32°15'	E48°20'	7.31	95.86	83.46	13.83
YiL09	*Populus euphratica*	Lorestan, Iran	N33°35'	E47°52'	6	92.7	87.9	11.41
YiL10	*Populus euphratica*	Lorestan, Iran	N33°33'	E47°51'	4.86	96.76	85.64	9.62
YL037	*Populus euphratica*	Lorestan, Iran	N33°35'	E47°52'	11.4	94.78	81.53	21.82
YL038	*Populus euphratica*	Lorestan, Iran	N33°34'	E47°52'	5.62	96.52	80.99	10.96
YL039	*Populus euphratica*	Lorestan, Iran	N33°33'	E47°51'	6.56	96.03	79.86	12.54
YiL21	*Populus euphratica*	Semnan, Iran	N35°10'	E52°26'	6.05	93.88	87.28	11.7
YiL22	*Populus euphratica*	Semnan, Iran	N35° 8'	E52°27'	4.47	96.89	84.08	8.87
YL025	*Populus euphratica*	Semnan, Iran	N35°13'	E52°29'	8.88	95.38	85.05	16.61
YL026	*Populus euphratica*	Semnan, Iran	N35°16'	E52°23'	8.55	86.9	85.11	15.36
YL027	*Populus euphratica*	Semnan, Iran	N35°13'	E52°19'	4.92	96.15	81.36	9.49
YL028	*Populus euphratica*	Semnan, Iran	N35°11'	E52°30'	6.61	93.79	82.1	11.63
YiL05	*Populus euphratica*	Sistan va Baluchestan, Iran	N29°27'	E60°50'	5.98	94.35	85.98	11.63
YiL06	*Populus euphratica*	Sistan va Baluchestan, Iran	N29°27'	E60°50'	5.92	92.63	85.72	11.17
YL047	*Populus euphratica*	Sistan va Baluchestan, Iran	N29°27'	E60°55'	5.44	95.66	79.48	10.49
YL048	*Populus euphratica*	Sistan va Baluchestan, Iran	N29°27'	E60°55'	8.5	95.98	82.96	16.54
YiL15	*Populus euphratica*	Yazd, Iran	N32° 2'	E54°12'	5.29	97.08	90.89	10.64
YiL16	*Populus euphratica*	Yazd, Iran	N32° 2'	E54°12'	6.01	95.17	93.97	11.86
YL049	*Populus euphratica*	Yazd, Iran	N32°2'	E54°14'	6.06	92.82	85.56	11.27
YL050	*Populus euphratica*	Yazd, Iran	N32°2'	E54°12'	6.85	95.43	88.27	13.15
YL051	*Populus euphratica*	Yazd, Iran	N32°1'	E54°12'	7.28	93.33	85.58	13.61
Peu18	*Populus euphratica*	Inner Mongolia, China	N42°01′	E101°03'	3.15	96.86	91.32	6.4
Peu19A	*Populus euphratica*	Inner Mongolia, China	N42°01′	E101°03'	3.63	95.44	85.83	7.29
Peu44	*Populus euphratica*	Inner Mongolia, China	N42°01′	E101°03'	6.18	97.71	94.32	12.63
Peu45	*Populus euphratica*	Inner Mongolia, China	N42°01′	E101°03'	3.41	97.49	87.51	6.94
Peu01A	*Populus euphratica*	Xinjiang, China	N46°08'	E85°38'	3.6	96.16	92.41	7.26
Peu02	*Populus euphratica*	Xinjiang, China	N46°08'	E85°38'	6.11	95.7	94.19	12.22
Peu03A	*Populus euphratica*	Xinjiang, China	N41°39'	E80°46'	3.68	96.22	92.65	7.43
Peu04A	*Populus euphratica*	Xinjiang, China	N40°17'	E80°21'	3.85	96.35	92.76	7.79
Peu05	*Populus euphratica*	Xinjiang, China	N40°19'	E79°55'	4.33	97.09	92.3	8.85
Peu06A	*Populus euphratica*	Xinjiang, China	N39°22'	E78°11'	4.31	94.86	93.49	8.48
Peu07	*Populus euphratica*	Xinjiang, China	N39°22'	E78°11'	10.44	95.66	95.17	20.62

Peu08A	*Populus euphratica*	Xinjiang, China	N37°06'	E80°58'	5.11	96.32	93.62	10.34
Peu09	*Populus euphratica*	Xinjiang, China	N37°11'	E82°47'	5.71	94.68	93.97	11.25
Peu10A	*Populus euphratica*	Xinjiang, China	N38°11'	E85°00'	4.08	95.18	92.62	8.16
Peu11	*Populus euphratica*	Xinjiang, China	N38°41'	E86°55'	10.59	93.22	94.98	20.24
Peu13A	*Populus euphratica*	Xinjiang, China	N39°47'	E88°22'	4.31	95.77	93.1	8.66
Peu14	*Populus euphratica*	Xinjiang, China	N40°40'	E87°35'	3.38	97.12	90.31	6.92
Peu15A	*Populus euphratica*	Xinjiang, China	N41°05'	E86°33'	4.32	96.4	93.18	8.76
Peu16A	*Populus euphratica*	Xinjiang, China	N41°15'	E84°12'	3.19	95.03	90.42	6.38
Peu17A	*Populus euphratica*	Xinjiang, China	N41°00'	E83°20'	4.19	95.08	92.91	8.37
Peu31	*Populus euphratica*	Xinjiang, China	N46°08'	E85°38'	2.76	97.73	87.36	5.64
Peu32	*Populus euphratica*	Xinjiang, China	N41°39'	E80°46'	5.82	96.89	91.46	11.79
Peu33	*Populus euphratica*	Xinjiang, China	N41°39'	E80°46'	2.79	96.26	87.12	5.54
Peu34	*Populus euphratica*	Xinjiang, China	N40°17'	E80°21'	2.51	97.88	85.93	5.14
Peu35	*Populus euphratica*	Xinjiang, China	N40°19'	E79°55'	2.92	97.67	81.98	5.97
Peu37	*Populus euphratica*	Xinjiang, China	N37°34'	E79°38'	5.65	97.15	92.24	11.45
Peu38	*Populus euphratica*	Xinjiang, China	N37°34'	E79°38'	2.62	97.31	89.09	5.26
Peu39	*Populus euphratica*	Xinjiang, China	N37°06'	E80°58'	5.85	97.84	92.14	11.95
Peu40	*Populus euphratica*	Xinjiang, China	N37°11'	E82°47'	5.45	95.32	93.43	10.8
Peu41	*Populus euphratica*	Xinjiang, China	N39°47'	E88°22'	4.82	97.95	91.31	9.89
Peu42	*Populus euphratica*	Xinjiang, China	N40°40'	E87°35'	8.17	97.24	94.24	16.64
Peu43	*Populus euphratica*	Xinjiang, China	N41°05'	E86°33'	3.9	97.52	89.71	7.93
Peu46	*Populus euphratica*	Xinjiang, China	N46°08'	E85°38'	3.13	97.72	85.88	6.41
Peu47	*Populus euphratica*	Xinjiang, China	N41°39'	E80°46'	3.46	96.72	89.34	6.92
Peu48	*Populus euphratica*	Xinjiang, China	N39°38'	E78°24'	4.21	97.41	91.79	8.59
Peu49	*Populus euphratica*	Xinjiang, China	N38°11'	E85°00'	4.35	97.66	91.61	8.86
Peu50	*Populus euphratica*	Xinjiang, China	N41°15'	E84°12'	4.61	97.58	90.75	9.43
Peu51	*Populus euphratica*	Xinjiang, China	N41°00'	E83°20'	7.48	94.27	92.27	14.78
Peu52	*Populus euphratica*	Xinjiang, China	N40°40'	E81°44′	4.07	96.43	93	8.25
Peu53	*Populus euphratica*	Xinjiang, China	N40°40'	E81°44′	3.78	97.03	91.93	7.7
Peu54	*Populus euphratica*	Xinjiang, China	N40°27′	E80°59′	3.82	96.31	92.99	7.73
Peu56	*Populus euphratica*	Xinjiang, China	N38°23′	E77°22′	4.15	96.24	93.16	8.39

Table S2: Population genetic summary statistics. Mean (±standard deviation) values of nucleotide diversity π, Tajima's D statistic, pairwise relative measure of differentiation (F_{ST}, above the diagonal) and absolute divergence (D_{xy}, below the diagonal) between all four populations. The number of pairwise fixed differences is given in parentheses.

Species	**Lineage**	**Number of individuals**	**π ($\times 10^{-3}$)**	**Tajima's *D***	**D_{xy} ($\times 10^{-3}$) / F_{ST} (fixed SNPs)**			
					PeC	**PeW**	**PeE**	**Pp**
P. euphratica	**PeC**	28	5.17±2.67	-0.04±0.80	-	0.33±0.13 (75,467)	0.19±0.09 (1,876)	0.31±0.11 (9,546)
	PeW	21	6.59±4.23	-0.22±0.99	7.73±2.85	-	0.36±0.15 (123,064)	0.19±0.08 (16,059)
	PeE	6	5.07±2.75	0.00±0.59	3.88±2.04	7.82±2.92	-	0.33±0.12 (26,454)
P. pruinosa	**Pp**	27	8.30±4.22	-0.02±0.78	7.82±2.71	6.34±2.79	8.07±2.80	-

Table S3: Relative likelihood of the different models shown in Fig. S6.

Model	Max(log_{10}(likelihood))[a]	No. of parameters	AIC[b]	δ[b]	Model normalized relative likelihood (w)[b]
mWP plus[c]	-118,029,395	34	543,545,518	0	1
mWP	-118,040,150	30	543,595,041	49,523	0
sWP	-118,063,925	30	543,704,528	159,010	0
sPW	-118,068,276	30	543,724,564	179,046	0
sPC	-118,284,097	30	544,718,455	1,172,937	0
sCP	-118,287,563	30	544,734,420	1,188,902	0
mCE	-118,287,632	30	544,734,736	1,189,218	0
sEC	-118,289,911	30	544,745,230	1,199,712	0
sCE	-118,294,500	30	544,766,366	1,220,848	0
mCW	-118,311,173	30	544,843,148	1,297,630	0
sCW	-118,313,080	30	544,851,928	1,306,411	0
sWC	-118,315,292	30	544,862,114	1,316,596	0
mCP	-118,325,432	30	544,908,814	1,363,296	0
sEW	-118,349,493	30	545,019,616	1,474,098	0
mWE	-118,350,792	30	545,025,597	1,480,080	0
sEP	-118,356,446	30	545,051,636	1,506,118	0
sWE	-118,357,088	30	545,054,591	1,509,073	0
mEP	-118,360,647	30	545,070,983	1,525,465	0
sPE	-118,363,153	30	545,082,522	1,537,004	0

[a]Based on the best likelihood among the 50 independent runs for each model

[b]The calculation of AIC, δ and w are according to the methods shown in [3]

[c]The best-fitting model is a modified version of model 'mWP' with exponential change in population size following their divergence.

Table S4: Inferred parameters estimates with 95% confidence intervals for the best-fitting demographic scenario modelled in fastsimcoal2, with parameter tags corresponding to the models shown in Fig. S7. Estimates of gene flow between populations are given in the migration fraction per generation.

Parameters	Point estimation	95% confidence intervals	
		Lower bound	Upper bound
NPeC	11,421	10,517	12,350
NPeW	21,749	20,874	22,233
NPeE	5,360	4,876	6,035
NPp	18,281	17,625	18,669
NancCE	44,222	34,370	48,058
NancWP	27,240	21,762	31,272
Nanc	9,608	5,803	16,798
NancC	7,743	7,227	9,481
NancW	6,985	5,643	9,579
NancE	37,468	28,970	40,612
NancP	44,242	34,667	46,778
TDIV	1,856,330	1,736,560	1,902,350
TWP	877,700	800,930	934,850
TCE	80,490	75,080	83,700
mWC	1.11E-08	2.94E-09	2.39E-07
mEC	1.06E-05	7.20E-06	1.21E-05
mPC	3.08E-06	2.71E-06	3.41E-06
mCW	5.68E-09	2.96E-09	1.39E-06
mEW	5.33E-06	3.85E-06	5.63E-06
mPW	2.04E-06	1.69E-06	2.55E-06
mCE	4.68E-07	4.55E-09	2.54E-06
mWE	2.30E-07	4.36E-09	6.06E-07
mPE	1.24E-06	8.10E-07	1.78E-06
mCP	1.83E-05	1.76E-05	1.94E-05
mWP	4.95E-09	2.32E-09	2.20E-07
mEP	9.30E-09	2.27E-09	2.51E-07
maPaW	1.93E-06	7.87E-07	2.75E-06
maWaP	4.98E-06	4.31E-06	5.59E-06
maWaCE	1.45E-08	3.19E-09	3.99E-07
maCEaW	4.07E-06	3.78E-06	4.36E-06
maPaCE	7.01E-07	3.55E-07	9.18E-07
maCEaP	6.72E-09	2.55E-09	4.01E-07
maWPaCE	4.11E-09	1.81E-09	3.95E-07
maCEaWP	3.13E-06	1.18E-06	3.92E-06

Table S5: Number of outlier windows and genomic islands after combining consecutive windows for all pairwise comparisons. The number of genes and Positively Selected Genes (PSGs) within these islands are also counted. Asterisks designate significantly higher proportions of PSGs in these islands by Chi-squared test (***P-value < 1E-8).

Pop1	Pop2	# of outlier windows	# of genomic islands	# of genes in these islands	# of PSGs in these islands
PeC	**PeW**	464	399	388	37***
PeC	**Pp**	422	369	385	28***
PeW	**Pp**	426	353	345	48***
PeC	**PeE**	454	385	410	6
PeW	**PeE**	464	406	319	19***
PeE	**Pp**	430	369	362	37***
Total		1,869	1,533	1,537	108***

Table S6: Number of shared outlier windows between paired comparisons.

	PeC-PeW	PeC-Pp	PeW-Pp	PeC-PeE	PeW-PeE
PeC-Pp	151				
PeW-Pp	44	49			
PeC-PeE	18	20	11		
PeW-PeE	228	91	41	7	
PeE-Pp	73	187	58	18	73

Table S7: Comparison of population genomic parameters (the mean ± standard deviation values) of genomic islands with the rest of the genomic regions for all pairwise comparisons by Mann–Whitney U test. Note that the linkage disequilibrium was not measured for lineage PeE due to limited number of samples.

Parameter	Lineage	Genomic background	Genomic islands	*P*-value
PeC-PeW				
F_{ST}		0.3074 ± 0.1449	0.7194 ± 0.0739	2.20E-16
D_{xy}		0.0076 ± 0.0028	0.0090 ± 0.0030	5.96E-10
π	PeC	0.0052 ± 0.0029	0.0030 ± 0.0020	2.20E-16
	PeW	0.0076 ± 0.0055	0.0027 ± 0.0025	2.20E-16
Tajima's *D*	PeC	-0.0784 ± 0.7947	-0.3626 ± 0.8447	1.80E-09
	PeW	-0.2143 ± 0.9452	-1.0453 ± 0.6735	2.20E-16
LD (r^2)	PeC	0.2980 ± 0.1342	0.3841 ± 0.2429	8.24E-09
	PeW	0.3098 ± 0.1598	0.2951 ± 0.2550	7.63E-08
ρ	PeC	0.0014 ± 0.0016	0.0004 ± 0.0007	2.20E-16
	PeW	0.0014 ± 0.0020	0.0008 ± 0.0013	1.26E-06
PeC-Pp				
F_{ST}		0.2867 ± 0.1275	0.6466 ± 0.0697	2.20E-16
D_{xy}		0.0077 ± 0.0026	0.0099 ± 0.0030	2.20E-16
π	PeC	0.0052 ± 0.0029	0.0028 ± 0.0020	2.20E-16
	Pp	0.0091 ± 0.0051	0.0049 ± 0.0029	2.20E-16
Tajima's *D*	PeC	-0.0776 ± 0.7942	-0.4729 ± 0.8707	2.67E-16
	Pp	-0.0359 ± 0.7543	-0.8221 ± 0.7257	2.20E-16
LD (r^2)	PeC	0.2981 ± 0.1343	0.3790 ± 0.2300	5.69E-09
	Pp	0.2898 ± 0.1372	0.4905 ± 0.2865	2.20E-16
ρ	PeC	0.0014 ± 0.0016	0.0004 ± 0.0007	2.20E-16
	Pp	0.0014 ± 0.0014	0.0002 ± 0.0005	2.20E-16
PeW-Pp				
F_{ST}		0.1752 ± 0.0889	0.5287 ± 0.0726	2.20E-16
D_{xy}		0.0062 ± 0.0027	0.0105 ± 0.0036	2.20E-16
π	PeW	0.0076 ± 0.0055	0.0043 ± 0.0036	2.20E-16
	Pp	0.0090 ± 0.0051	0.0063 ± 0.0040	2.20E-16
Tajima's *D*	PeW	-0.2162 ± 0.9458	-0.8443 ± 0.7864	2.20E-16
	Pp	-0.0382 ± 0.7558	-0.5227 ± 0.7899	2.20E-16
LD (r^2)	PeW	0.3095 ± 0.1595	0.3406 ± 0.2505	4.84E-01
	Pp	0.2911 ± 0.1392	0.3797 ± 0.2426	8.80E-08
ρ	PeW	0.0014 ± 0.0020	0.0008 ± 0.0016	7.41E-09
	Pp	0.0014 ± 0.0014	0.0006 ± 0.0008	2.20E-16
PeC-PeE				
F_{ST}		0.1886 ± 0.1000	0.5856 ± 0.0943	2.20E-16
D_{xy}		0.0037 ± 0.0019	0.0076 ± 0.0025	2.20E-16

π	PeC	0.0052 ± 0.0029	0.0034 ± 0.0019	2.20E-16
	PeE	0.0054 ± 0.0030	0.0048 ± 0.0029	1.12E-04
Tajima's *D*	PeC	-0.0760 ± 0.7943	-0.6019 ± 0.7647	2.20E-16
	PeE	-0.0096 ± 0.5700	-0.0914 ± 0.8144	3.00E-04
LD (r^2)	PeC	0.2973 ± 0.1335	0.4234 ± 0.2227	2.20E-16
	PeE	-	-	-
ρ	PeC	0.0014 ± 0.0016	0.0007 ± 0.0010	2.20E-16
	PeE	0.0013 ± 0.0017	0.0004 ± 0.0008	2.20E-16
PeW-PeE				
F_{ST}		0.3362 ± 0.1703	0.7878 ± 0.0698	2.20E-16
D_{xy}		0.0077 ± 0.0028	0.0083 ± 0.0027	2.15E-03
π	PeW	0.0076 ± 0.0055	0.0025 ± 0.0023	2.20E-16
	PeE	0.0054 ± 0.0030	0.0037 ± 0.0021	2.20E-16
Tajima's *D*	PeW	-0.2143 ± 0.9454	-1.0374 ± 0.6512	2.20E-16
	PeE	-0.0098 ± 0.5724	-0.0694 ± 0.6179	9.37E-03
LD (r^2)	PeW	0.3098 ± 0.1598	0.2788 ± 0.2531	3.23E-10
	PeE	-	-	-
ρ	PeW	0.0014 ± 0.0020	0.0009 ± 0.0012	1.49E-01
	PeE	0.0013 ± 0.0017	0.0004 ± 0.0008	1.05E-15
PeE-Pp				
F_{ST}		0.3089 ± 0.1362	0.6808 ± 0.0761	2.20E-16
D_{xy}		0.0079 ± 0.0027	0.0083 ± 0.0035	2.04E-01
π	PeE	0.0054 ± 0.0030	0.0034 ± 0.0020	2.20E-16
	Pp	0.0091 ± 0.0051	0.0044 ± 0.0032	2.20E-16
Tajima's *D*	PeE	-0.0090 ± 0.5725	-0.1658 ± 0.5987	1.11E-07
	Pp	-0.0340 ± 0.7532	-0.9748 ± 0.6427	2.20E-16
LD (r^2)	PeE	-	-	-
	Pp	0.2909 ± 0.1387	0.3904 ± 0.2805	2.53E-04
ρ	PeE	0.0013 ± 0.0017	0.0004 ± 0.0006	3.10E-15
	Pp	0.0014 ± 0.0014	0.0005 ± 0.0009	2.20E-16

Table S8: Lists of positively selected genes (PSGs) with a significant nominal *P*-value (<0.01) for the HKA test and 95th percentile population branch statistic (PBS) values for lineages PeC, PeW and Pp. Note that some PSGs shared between lineages are listed twice.

Gene ID	Best Arabidopsis hit	Gene Name	Description	PBS	HKA	Genomic islands
PeC						
CCG000110.1	AT3G12560.1	TRP3	Telomere repeat-binding protein 3	0.386	2.20E-16	-
CCG000133.2	AT3G53830.1			0.382	2.20E-16	-
CCG000211.1	AT2G40140.1		Zinc finger CCCH domain-containing protein 29	0.365	2.20E-16	-
CCG000279.1	AT2G31600.1			0.508	2.20E-16	PeC-Pp; PeE-Pp
CCG000305.2	AT1G07950.1	MED22B	Mediator of RNA polymerase II transcription subunit 22b	0.5	8.15E-03	-
CCG000359.1	-			0.353	2.20E-16	PeW-PeE
CCG000371.1	AT1G26761.1			0.376	2.20E-16	-
CCG000375.1	AT2G03250.1			0.376	2.20E-16	-
CCG000627.1	AT1G25425.1	CLE43	CLAVATA3/ESR (CLE)-related protein 43	0.531	1.26E-07	-
CCG001025.1	AT3G10910.1	ATL72	RING-H2 finger protein ATL72	0.409	5.75E-04	-
CCG001718.1	AT2G24350.1			0.433	1.87E-09	-
CCG002106.1	AT5G11380.1			0.431	4.65E-15	-
CCG002491.1	AT2G34540.2			1.098	2.20E-16	-
CCG002528.1	AT5G57620.1	MYB36	Transcription factor MYB36	0.506	2.20E-16	-
CCG002866.1	AT1G09700.1	DRB1	Double-stranded RNA-binding protein 1	0.528	1.87E-09	PeC-Pp
CCG003726.1	AT5G22090.1		Protein FAF-like, chloroplastic	0.374	2.20E-16	PeE-Pp
CCG003769.1	AT1G70740.1			0.37	7.43E-05	PeC-PeW; PeW-PeE
CCG003781.1	AT2G01505.1	CLE16	CLAVATA3/ESR (CLE)-related protein 16	0.559	1.26E-07	-
CCG003907.1	AT5G06570.1	CXE15	Probable carboxylesterase 15	0.39	8.18E-12	-
CCG003911.2	AT3G54750.3			0.378	2.20E-16	-
CCG005859.1	AT3G28690.2			0.413	2.20E-16	-
CCG006332.1	AT5G57015.1	CKL12	Casein kinase 1-like protein 12	1.025	2.20E-16	PeC-Pp; PeE-Pp
CCG007424.1	AT5G48820.2	KRP3	Cyclin-dependent kinase inhibitor 3	0.488	3.50E-14	-
CCG007602.1	AT3G12890.1			0.461	1.87E-09	PeE-Pp
CCG007604.1	AT5G60570.1		F-box/kelch-repeat protein At5g60570	0.495	2.20E-16	-
CCG007697.1	AT4G36910.1	CBSX1	CBS domain-containing protein CBSX1, chloroplastic	0.586	1.87E-09	-
CCG008358.1	AT1G32700.1			1.848	2.52E-07	-
CCG008722.1	AT1G76940.1	NSRA	Nuclear speckle RNA-binding protein A	0.845	2.52E-07	-
CCG009176.2	AT2G26070.1	RTE1	Protein REVERSION-TO-ETHYLENE SENSITIVITY1	0.444	5.75E-04	-
CCG009591.1	AT1G15550.1	GA3OX1	Gibberellin 3-beta-dioxygenase 1	0.862	5.75E-04	-
CCG010076.1	AT4G10650.1	DGP2	DAR GTPase 2, mitochondrial	0.484	1.87E-09	-
CCG010372.1	AT1G75100.1	JAC1		0.384	2.20E-16	-
CCG010374.1	AT4G16580.1		Probable protein phosphatase 2C 55	0.391	8.18E-12	-
CCG010383.1	AT2G33750.1	PUP2	Purine permease 2	0.373	4.65E-15	-
CCG010761.1	AT2G46310.1	CRF5	Ethylene-responsive transcription factor CRF5	0.382	4.26E-05	-

CCG010829.1	AT2G46600.1	KIC	Calcium-binding protein KIC	0.801	2.52E-07	-
CCG011200.1	AT3G10700.1	GALAK	Galacturonokinase	0.512	2.20E-16	-
CCG011257.1	AT2G11520.1	CRCK3	Calmodulin-binding receptor-like cytoplasmic kinase 3	0.363	2.20E-16	-
CCG011362.1	-			1.025	2.52E-07	-
CCG012315.1	AT1G04940.1	TIC20-I	Protein TIC 20-I, chloroplastic	0.433	2.20E-16	-
CCG012317.2	AT1G35180.1			0.377	2.20E-16	-
CCG012319.1	AT1G75450.1	CKX5	Cytokinin dehydrogenase 5	0.372	2.20E-16	PeC-Pp; PeE-Pp
CCG012334.1	AT2G02040.1	NPF8.3	Protein NRT1/ PTR FAMILY 8.3	0.399	2.20E-16	-
CCG012587.2	AT1G73970.1			0.385	5.75E-04	-
CCG013195.1	AT3G22830.1	HSFA6B	Heat stress transcription factor A-6b	0.371	4.06E-10	PeC-Pp; PeE-Pp
CCG013358.1	AT5G16770.1			0.43	2.20E-16	-
CCG013426.1	AT3G50620.1			0.549	2.20E-16	-
CCG013427.1	AT5G66815.1			0.448	1.72E-12	-
CCG013429.1	AT2G23440.1			0.573	1.87E-09	-
CCG013430.1	AT2G23440.1			0.46	8.18E-12	PeE-Pp
CCG013431.1	AT5G66820.1			0.359	2.20E-16	-
CCG013531.1	AT2G16750.1			0.449	2.20E-16	PeC-Pp; PeE-Pp
CCG013539.1	AT1G75950.1	SKP1A	SKP1-like protein 1A	0.518	8.18E-12	-
CCG013576.1	AT4G30210.1	ATR2	NADPH--cytochrome P450 reductase 2	0.557	1.29E-12	PeC-PeW; PeC-Pp
CCG014164.1	AT1G09910.1			0.397	8.15E-03	PeC-PeW; PeC-Pp; PeW-PeE
CCG014366.1	AT5G45550.1	MOB1A	MOB kinase activator-like 1A	1.115	2.52E-07	-
CCG014403.1	AT5G41380.1			0.455	5.75E-04	-
CCG014621.1	AT3G21140.1			0.392	3.20E-14	-
CCG014824.1	AT3G49940.1	LBD38	LOB domain-containing protein 38	0.68	1.87E-09	-
CCG015265.1	AT5G47870.1			0.444	2.13E-06	-
CCG016281.1	AT1G45688.1			0.402	4.26E-05	-
CCG016353.2	-			0.782	1.72E-12	PeC-Pp
CCG016373.1	AT2G28070.1	ABCG3	ABC transporter G family member 3	0.431	2.11E-06	-
CCG016375.1	AT3G45040.1	DOK1	Dolichol kinase EVAN	0.383	4.70E-03	-
CCG016662.1	AT2G20990.3	SYT1	Synaptotagmin-1	0.364	2.20E-16	-
CCG016825.1	AT1G58250.2			0.466	2.20E-16	-
CCG017164.1	AT2G20940.1			0.382	2.20E-16	-
CCG017184.1	AT4G35000.1	APX3	L-ascorbate peroxidase 3, peroxisomal	1.806	2.20E-16	-
CCG017317.1	AT4G34530.1	CIB1	Transcription factor bHLH63	0.411	2.20E-16	-
CCG017334.1	AT4G34460.1	GB1	Guanine nucleotide-binding protein subunit beta	0.459	2.20E-16	-
CCG017482.1	AT4G34150.1			0.358	7.64E-09	PeC-PeW
CCG017483.2	AT4G34160.1	CYCD3-1	Cyclin-D3-1	0.452	2.20E-16	-
CCG017532.1	AT2G14880.1			0.65	2.20E-16	-
CCG017536.1	AT5G43400.1			0.395	2.20E-16	-
CCG017558.1	AT4G33880.1	RSL2	Transcription factor bHLH85	0.489	2.13E-06	-
CCG017572.1	AT3G49870.1			0.455	4.26E-05	-
CCG017594.2	AT4G08170.2	ITPK3	Inositol-tetrakisphosphate 1-kinase 3	0.356	1.29E-12	-
CCG017598.1	AT5G60440.1	AGL62	Agamous-like MADS-box protein AGL62	0.459	2.20E-16	PeC-PeE
CCG017854.1	AT1G03170.1	FAF2	Protein FANTASTIC FOUR 2	0.385	6.36E-04	-

CCG018006.1	AT3G14067.1	SBT1.4	Subtilisin-like protease SBT1.4	0.676	2.20E-16	-
CCG018140.2	AT5G24910.1	CYP714A1	Cytochrome P450 714A1	0.37	2.20E-16	-
CCG018730.1	AT5G40490.1	HLP1		0.537	3.50E-14	-
CCG019053.1	AT2G46040.1	ARID1	AT-rich interactive domain-containing protein 1	0.361	2.20E-16	-
CCG019459.1	AT1G74960.2	KAS2	3-oxoacyl-[acyl-carrier-protein] synthase II, chloroplastic	0.394	2.20E-16	-
CCG019610.2	AT2G14120.3	DRP3B	Dynamin-related protein 3B	0.767	2.20E-16	PeC-PeW; PeC-Pp; PeW-PeE; PeE-Pp
CCG019658.2	AT5G09590.1	HSP70	Heat shock 70 kDa protein 10, mitochondrial	0.376	7.64E-09	PeC-PeW
CCG019662.1	AT5G64420.1			0.453	2.20E-16	PeC-PeW; PeC-Pp
CCG019705.1	-			0.535	2.20E-16	-
CCG019725.1	AT4G25210.1	GEBPL	GLABROUS1 enhancer-binding protein-like	0.382	2.20E-16	-
CCG019726.1	AT5G67500.2	VDAC2	Mitochondrial outer membrane protein porin 2	0.354	3.50E-14	-
CCG019881.1	AT3G21670.1	NPF6.4	Protein NRT1/ PTR FAMILY 6.4	0.409	2.20E-16	PeC-PeW
CCG019959.1	AT5G05610.1	AL1	PHD finger protein ALFIN-LIKE 1	0.616	4.26E-05	-
CCG019960.1	AT2G01060.1	PHL7	Myb family transcription factor PHL7	0.354	1.26E-07	-
CCG020025.1	AT2G22420.1	PER17	Peroxidase 17	0.366	1.62E-05	-
CCG020253.1	AT3G14450.1	CID9	Polyadenylate-binding protein-interacting protein 9	0.382	1.87E-09	-
CCG020366.1	AT1G17330.1			0.468	2.60E-16	-
CCG020458.1	AT2G40060.1	CLC2	Clathrin light chain 2	0.404	2.20E-16	-
CCG020475.1	AT3G11260.1	WOX5	WUSCHEL-related homeobox 5	1.451	2.20E-16	-
CCG020560.1	AT4G36920.1	AP2	Floral homeotic protein APETALA 2	0.409	3.52E-16	-
CCG020695.1	AT1G06330.1			0.531	1.72E-12	-
CCG020711.1	AT1G27340.1	LCR	F-box only protein 6	0.44	2.20E-16	-
CCG020739.1	AT5G13140.1			0.65	4.26E-05	-
CCG020746.1	AT5G13180.1	NAC083	NAC domain-containing protein 83	0.612	3.50E-14	-
CCG020747.1	AT5G50850.1	PDH2		0.465	2.20E-16	-
CCG020749.1	AT5G21090.1			0.36	1.26E-07	-
CCG020776.1	AT3G26618.1	ERF1-3	Eukaryotic peptide chain release factor subunit 1-3	0.558	2.20E-16	-
CCG020778.1	AT1G66050.1	ORTH5	E3 ubiquitin-protein ligase ORTHRUS 5	0.49	2.20E-16	PeC-Pp; PeE-Pp
CCG021050.1	-			1.649	1.72E-12	-
CCG021060.1	AT5G40990.1	GLIP1	GDSL esterase/lipase 1	0.376	2.48E-09	-
CCG021426.1	AT5G63870.2	PP7	Serine/threonine-protein phosphatase 7	0.351	2.20E-16	PeC-Pp; PeE-Pp
CCG021427.1	AT3G13780.1			0.589	2.20E-16	PeC-Pp; PeE-Pp
CCG021453.1	-			0.469	1.87E-09	-
CCG021668.1	AT4G10750.1			0.435	7.43E-05	-
CCG021702.1	AT5G42050.1			0.798	2.52E-07	-
CCG022019.1	AT1G51720.1			0.417	3.50E-14	-
CCG022442.1	AT1G51760.1	ILL4	IAA-amino acid hydrolase ILR1-like 4	2.048	2.52E-07	PeC-Pp; PeE-Pp
CCG022605.1	AT2G26170.1	CYP711A1	Cytochrome P450 711A1	0.502	2.20E-16	-
CCG022606.1	AT2G26180.1			0.386	4.65E-15	PeE-Pp
CCG022649.1	AT5G16320.1	FRL1	FRIGIDA-like protein 1	0.366	2.20E-16	-
CCG023204.1	-			0.361	8.18E-12	-
CCG023330.1	AT1G30840.1	PUP4	Probable purine permease 4	0.561	2.20E-16	-
CCG023690.1	AT1G07010.3	SLP1	Shewanella-like protein phosphatase 1	0.72	2.20E-16	-

CCG023830.1	AT5G01020.1		Serine/threonine-protein kinase At5g01020	0.559	2.20E-16	-
CCG024145.1	AT1G56170.1	NFYC2	Nuclear transcription factor Y subunit C-2	0.513	2.20E-16	-
CCG024158.1	AT5G11890.1			0.349	2.20E-16	-
CCG024228.1	-			0.409	2.20E-16	-
CCG024379.1	AT1G62850.2			0.352	1.87E-09	-
CCG024388.2	AT4G22760.1	PCMP-E6	Pentatricopeptide repeat-containing protein At4g22760	0.389	2.80E-03	-
CCG025109.1	AT1G68620.1	CXE6	Probable carboxylesterase 6	0.683	2.20E-16	-
CCG025151.2	AT1G68690.1	PERK9	Proline-rich receptor-like protein kinase PERK9	0.409	2.20E-16	-
CCG025153.1	AT3G61460.1			0.359	2.13E-06	-
CCG025162.1	AT1G13570.1		F-box/FBD/LRR-repeat protein At1g13570	0.369	8.18E-12	-
CCG025241.3	AT3G66654.1	CYP21-4	Peptidyl-prolyl cis-trans isomerase CYP21-4	0.506	2.20E-16	-
CCG025261.1	AT5G49480.1			0.409	3.50E-14	-
CCG025277.1	AT1G47980.1			0.493	2.20E-16	-
CCG025658.1	-			0.493	2.20E-16	-
CCG026007.2	AT2G01870.1			1.22	1.72E-12	PeC-PeW; PeC-Pp; PeE-Pp
CCG026259.1	AT1G72490.1			0.473	2.20E-16	-
CCG026474.1	AT5G39210.1	CRR7	Protein CHLORORESPIRATORY REDUCTION 7	2.031	1.72E-12	-
CCG026817.1	AT2G41540.1	GPDHC1	Glycerol-3-phosphate dehydrogenase [NAD(+)] GPDHC1	0.754	2.52E-07	-
CCG027462.1	AT1G72760.1			0.402	2.20E-16	PeE-Pp
CCG027485.2	AT1G36050.1			0.445	3.20E-14	-
CCG027486.1	AT5G65140.1	TPPJ	Probable trehalose-phosphate phosphatase J	0.63	1.26E-07	-
CCG027500.1	AT1G78170.1			0.417	1.26E-07	-
CCG027579.1	AT5G49700.1	AHL17	AT-hook motif nuclear-localized protein 17	0.592	2.20E-16	-
CCG027827.1	AT4G35600.2		Probable serine/threonine-protein kinase Cx32, chloroplastic	0.489	2.20E-16	-
CCG028167.2	AT2G17410.1	ARID3	AT-rich interactive domain-containing protein 3	0.353	2.20E-16	-
CCG028333.1	AT4G10590.2	UBP9	Ubiquitin carboxyl-terminal hydrolase 9	0.364	2.20E-16	-
CCG028575.1	AT3G50240.1	KIN4B	Kinesin-like protein KIN-4B	0.379	6.99E-09	PeC-PeW; PeC-Pp
CCG029163.1	AT2G25800.1			0.535	1.40E-03	-
CCG029166.1	AT4G32810.1	CCD8	Carotenoid cleavage dioxygenase 8, chloroplastic	0.387	7.43E-05	-
CCG029660.2	AT1G64550.1	ABCF3	ABC transporter F family member 3	0.353	2.20E-16	-
CCG029692.1	AT4G23910.1			0.372	4.26E-05	-
CCG029702.1	AT4G23840.1			0.378	2.20E-16	-
CCG029732.1	AT4G11160.1			0.4	2.20E-16	-
CCG029977.1	AT4G29230.1			0.522	2.20E-16	-
CCG030309.1	AT5G15210.1	ZHD8	Zinc-finger homeodomain protein 8	0.487	2.80E-03	-
CCG030670.2	AT1G33490.1			0.425	2.52E-07	-
CCG030915.1	AT3G48140.1			1.425	1.87E-09	-
CCG031189.1	AT3G26935.1	PAT07	Probable protein S-acyltransferase 7	0.375	2.20E-16	-
CCG031270.1	AT2G06050.2	OPR3	12-oxophytodienoate reductase 3	0.395	2.20E-16	PeC-PeW; PeC-Pp; PeE-Pp
CCG031367.1	AT5G07900.1			0.368	4.06E-10	-
CCG031432.1	AT5G53090.1			0.742	2.20E-16	PeE-Pp
CCG031437.1	AT1G80810.2			0.429	2.20E-16	PeC-PeE; PeC-Pp
CCG031640.1	AT4G32430.1			0.481	2.20E-16	PeC-Pp; PeE-Pp
CCG032127.1	AT4G32915.1	GATC		1.961	1.72E-12	-

CCG032293.2	AT1G22860.1			0.777	2.20E-16	PeC-PeW; PeC-Pp; PeE-Pp
CCG032297.1	AT5G54160.1	OMT1	Flavone 3'-O-methyltransferase 1	0.349	1.87E-09	PeC-Pp; PeW-Pp
CCG032301.1	AT1G10020.1			0.512	2.20E-16	-
CCG032484.1	AT3G23940.1	DHAD	Dihydroxy-acid dehydratase, chloroplastic	0.37	2.20E-16	-
CCG032488.1	AT1G04780.1			0.393	1.29E-12	PeC-PeW
CCG032840.2	AT4G32940.1		Vacuolar-processing enzyme gamma-isozyme	0.374	2.20E-16	-
CCG032890.1	AT3G03860.1	APRL5	5'-adenylylsulfate reductase-like 5	0.38	2.20E-16	-
CCG032891.1	AT5G18130.1			0.66	2.20E-16	-
CCG032986.1	AT1G68660.1	CPLS1	ATP-dependent Clp protease adapter protein CLPS1	0.647	2.20E-16	PeC-PeW; PeC-Pp; PeE-Pp
CCG032987.1	AT4G21200.1	GA2OX8	Gibberellin 2-beta-dioxygenase 8	0.378	3.20E-14	PeE-Pp
CCG033061.1	AT1G24764.1	MAP70.2	Microtubule-associated protein 70-2	0.398	2.20E-16	-
CCG033066.1	AT2G01690.2	VAC14	Protein VAC14 homolog	0.405	2.20E-16	-
CCG033619.1	AT5G17060.1			0.367	3.50E-14	-
CCG033620.1	AT3G03110.1	XPO1B	Protein EXPORTIN 1B	0.541	2.20E-16	-
CCG033885.1	AT2G33490.1		Uncharacterized protein At2g33490	0.508	2.20E-16	PeC-PeW; PeC-Pp
CCG033939.1	AT1G09730.1	ULP2B	Probable ubiquitin-like-specific protease 2B	0.458	2.56E-12	PeC-Pp; PeE-Pp
CCG034082.1	AT5G06300.1	LOG7		1.201	2.20E-16	-
CCG034093.1	AT5G39590.1			0.353	2.20E-16	-
CCG034094.1	AT1G19860.1		Zinc finger CCCH domain-containing protein 6	0.389	2.20E-16	-
CCG034114.1	AT1G75520.1	SRS5	Protein SHI RELATED SEQUENCE 5	0.528	2.20E-16	-
CCG034190.1	AT1G67080.1	ABA4	Protein ABA DEFICIENT 4, chloroplastic	0.544	3.50E-14	-
CCG034275.1	AT1G77170.1			0.497	2.20E-16	-
PeW						
CCG000298.1	AT5G36950.1	DEGP10	Protease Do-like 10, mitochondrial	0.166	3.24E-03	-
CCG000327.3	AT3G12760.1			0.389	2.20E-16	-
CCG000332.1	AT1G16280.1	RH36	DEAD-box ATP-dependent RNA helicase 36	0.168	2.20E-16	-
CCG000356.1	AT5G39660.1	CDF2	Cyclic dof factor 2	0.206	5.56E-04	-
CCG000860.1	-			0.265	1.17E-03	-
CCG001040.1	AT5G50750.1	RGP4	Probable UDP-arabinopyranose mutase 4	0.192	2.45E-08	-
CCG001041.1	AT5G15810.1		Probable tRNA (guanine(26)-N(2))-dimethyltransferase 1	0.366	2.20E-16	PeW-Pp
CCG001174.1	AT5G16150.1		Plastidic glucose transporter 4	0.236	2.20E-16	-
CCG001734.1	AT2G21385.4			0.212	2.20E-16	-
CCG001797.2	AT2G21120.1		Probable magnesium transporter NIPA6	0.309	2.20E-16	-
CCG001935.1	AT2G20820.2			0.43	2.74E-12	-
CCG001942.1	AT1G03530.1			0.186	2.20E-16	-
CCG001946.1	AT2G20840.1	SCAMP1	Secretory carrier-associated membrane protein 1	0.281	2.67E-16	-
CCG002048.1	-			0.192	2.20E-16	-
CCG002483.1	ATMG00980.1			0.398	2.67E-16	-
CCG002491.1	AT2G34540.2			0.78	1.17E-03	-
CCG002562.1	AT5G62790.2	DXR	1-deoxy-D-xylulose 5-phosphate reductoisomerase	0.196	2.20E-16	-
CCG002566.1	AT5G14750.1	MYB66	Transcription factor MYB66	0.428	2.20E-16	PeW-Pp
CCG002653.1	AT2G47140.1	SDR3B	Short-chain dehydrogenase reductase 3b	0.315	1.90E-08	-
CCG003228.1	AT5G60440.1	AGL62	Agamous-like MADS-box protein AGL62	0.233	1.90E-08	-

CCG003691.1	AT1G63640.1			0.181	2.20E-16	-
CCG003769.1	AT1G70740.1			0.23	1.65E-06	PeC-PeW; PeW-PeE
CCG003889.3	AT3G58490.1	LPPD	Lipid phosphate phosphatase delta	0.2	2.74E-12	-
CCG004099.1	AT2G32990.1		Endoglucanase 11	0.167	1.93E-12	-
CCG004499.1	AT2G21050.1	LAX2	Auxin transporter-like protein 2	0.188	1.51E-04	-
CCG004578.1	AT5G60910.1	AGL8	Agamous-like MADS-box protein AGL8	0.286	2.74E-12	PeC-PeW; PeW-PeE
CCG004846.1	AT1G13980.1	GN	ARF guanine-nucleotide exchange factor GNOM	0.27	2.20E-16	-
CCG005158.1	AT1G15460.1	BOR4	Boron transporter 4	0.202	2.20E-16	PeW-PeE; PeE-Pp
CCG005453.1	AT4G30220.2		Probable small nuclear ribonucleoprotein F	1.267	1.17E-03	-
CCG005511.1	AT5G10770.1		Aspartyl protease family protein At5g10770	0.215	5.31E-11	PeC-PeW; PeW-PeE
CCG005643.1	AT5G54730.1	ATG18F	Autophagy-related protein 18f	0.23	2.20E-16	-
CCG005753.1	-			0.251	2.20E-16	-
CCG005795.1	AT3G01170.1			0.217	2.74E-12	-
CCG006205.1	AT4G39530.1			0.211	2.20E-16	-
CCG006206.1	AT1G28380.1	NSL1	MACPF domain-containing protein NSL1	0.283	2.20E-16	-
CCG006922.1	AT3G26040.1			0.219	2.20E-16	-
CCG007363.1	AT1G70980.1	SYNC3	Asparagine--tRNA ligase, cytoplasmic 3	0.477	1.17E-03	-
CCG007415.1	AT1G02170.1	AMC1	Metacaspase-1	0.245	1.50E-15	-
CCG007498.1	AT1G21410.1	SKP2A	F-box protein SKP2A	0.281	1.90E-08	-
CCG007698.1	AT1G74160.1			0.209	2.20E-16	-
CCG007699.1	AT3G02160.1			0.374	2.20E-16	-
CCG007711.1	AT3G29400.1			0.214	2.20E-16	PeW-Pp
CCG007712.1	AT3G29400.1			0.433	2.20E-16	PeW-Pp
CCG007713.1	AT5G50380.1			0.277	2.20E-16	PeW-Pp
CCG007716.1	-			0.196	1.17E-03	-
CCG007717.1	AT5G15550.1			0.268	2.20E-16	PeW-Pp
CCG007720.1	AT5G39240.1			0.551	2.20E-16	-
CCG007806.1	AT4G14490.1		FHA domain-containing protein At4g14490	0.2	2.20E-16	-
CCG007936.1	AT3G56370.1	IRK		0.21	2.20E-16	-
CCG008318.1	AT5G65685.1			0.18	1.17E-03	-
CCG008365.1	AT5G60850.1	DOF5.4	Dof zinc finger protein DOF5.4	0.169	1.17E-03	PeC-PeW; PeW-PeE
CCG008620.1	AT1G04360.1	ATL1	RING-H2 finger protein ATL1	0.167	2.20E-16	-
CCG009373.1	AT2G19560.1	EER5	Enhanced ethylene response protein 5	0.337	2.20E-16	PeC-PeW; PeW-PeE; PeW-Pp
CCG009434.1	AT2G26110.1			0.205	2.20E-16	-
CCG009539.1	AT4G32760.2			0.169	2.20E-16	-
CCG009656.1	AT2G42000.2	MT4A	Metallothionein-like protein 4A	0.43	1.17E-03	-
CCG009657.3	AT1G20670.1			0.22	2.20E-16	-
CCG009658.1	AT4G38350.2			0.222	2.20E-16	-
CCG010335.1	AT1G55200.1			0.221	1.44E-04	-
CCG011362.1	-			0.94	2.74E-12	-
CCG011363.1	AT5G15720.1	GLIP7	GDSL esterase/lipase 7	0.24	2.20E-16	-
CCG011521.1	AT5G66290.1			0.319	1.17E-03	-
CCG011542.1	AT2G26140.1	FTSH4	ATP-dependent zinc metalloprotease FTSH 4	0.207	2.20E-16	-
CCG011554.1	AT1G66920.2			0.23	2.20E-16	-

CCG011700.2	AT1G54070.1		Dormancy-associated protein homolog 4	0.425	2.20E-16	-
CCG011708.1	AT1G72370.1	RPSAA	40S ribosomal protein Sa-1	0.255	2.45E-05	-
CCG011727.1	AT1G72310.1	ATL3	RING-H2 finger protein ATL3	0.335	1.92E-07	-
CCG012574.1	AT5G16010.1			0.675	2.20E-16	-
CCG012589.1	AT5G20230.1	BCB	Blue copper protein	0.196	2.74E-12	-
CCG012591.1	AT3G44110.1	ATJ3	Chaperone protein dnaJ 3	0.452	2.20E-16	-
CCG012592.1	AT5G06490.1	ATL71	Putative RING-H2 finger protein ATL71	1.96	2.20E-16	-
CCG013069.3	AT2G19880.2			0.606	2.20E-16	-
CCG013362.3	AT5G36290.1		GDT1-like protein 3	0.168	1.17E-03	PeC-PeW
CCG013393.1	AT5G18700.1	RUK	Serine/threonine-protein kinase RUNKEL	0.239	1.49E-03	PeC-PeW; PeW-PeE
CCG014245.1	AT4G12230.1			0.206	2.20E-16	-
CCG014279.1	AT5G47570.1			0.287	1.17E-03	-
CCG014352.2	AT3G49010.1	RPL13B	60S ribosomal protein L13-1	0.376	2.67E-16	-
CCG014357.1	AT1G31880.1	BRX	Protein BREVIS RADIX	0.21	1.65E-06	-
CCG014441.1	AT4G17920.1	ATL29	RING-H2 finger protein ATL29	0.206	1.65E-06	-
CCG014442.1	AT5G46690.1	BHLH71	Transcription factor bHLH71	0.193	2.20E-16	-
CCG014445.1	AT5G46700.1	TRN2	Protein TORNADO 2	0.171	1.65E-06	-
CCG014585.1	AT3G16260.1			0.173	2.20E-16	-
CCG014590.1	AT1G10865.1			0.214	1.17E-03	-
CCG014621.1	AT3G21140.1			0.277	1.17E-03	-
CCG014628.1	AT1G50430.1	DWF5	7-dehydrocholesterol reductase	0.197	3.99E-13	-
CCG014629.1	AT1G16810.1			0.363	1.90E-08	PeW-Pp
CCG014749.1	AT5G37310.1	TMN4	Transmembrane 9 superfamily member 4	0.342	2.20E-16	-
CCG014792.1	AT1G30640.1			0.212	2.20E-16	-
CCG015096.3	AT1G53310.1	PPC1	Phosphoenolpyruvate carboxylase 1	0.247	2.20E-16	-
CCG015315.1	AT4G10430.1			0.508	2.74E-12	-
CCG016124.1	AT5G27700.1	RPS21C	40S ribosomal protein S21-2	1.643	1.17E-03	-
CCG016185.1	AT4G25800.1	CBP60D	Calmodulin-binding protein 60 D	0.373	2.45E-05	-
CCG016633.1	AT3G26040.1			0.172	3.10E-05	-
CCG016647.1	AT3G54840.1	RABF1	Ras-related protein RABF1	0.333	2.67E-16	-
CCG017091.1	AT4G02750.1			0.467	1.90E-08	-
CCG017093.1	AT5G59810.1	SBT5.4	Subtilisin-like protease SBT5.4	0.179	3.99E-13	-
CCG017110.1	AT4G00710.1			0.413	2.20E-16	PeC-PeW; PeW-PeE; PeW-Pp
CCG017143.1	-			0.261	1.17E-03	-
CCG017312.1	AT3G17600.1	IAA31	Auxin-responsive protein IAA31	0.22	1.17E-03	-
CCG017318.1	AT1G75250.1	RL6	Protein RADIALIS-like 6	0.257	2.20E-16	-
CCG017411.1	AT2G26520.1			1.83	2.74E-12	-
CCG017482.1	AT4G34150.1			0.216	2.20E-16	PeC-PeW
CCG018329.1	AT4G32150.1	VAMP711	Vesicle-associated membrane protein 711	0.188	2.45E-05	-
CCG019515.1	AT2G35930.1	PUB23	E3 ubiquitin-protein ligase PUB23	0.238	1.65E-06	-
CCG019936.1	AT1G10780.1		F-box protein At1g10780	0.242	1.76E-09	-
CCG020355.1	AT5G07720.1	XXT3	Probable xyloglucan 6-xylosyltransferase 3	0.21	1.17E-03	PeC-PeW; PeC-Pp; PeW-PeE
CCG020363.1	AT5G23130.1			0.18	1.51E-04	-
CCG020471.1	AT3G55830.1	EPC1	Glycosyltransferase family 64 protein C4	0.171	1.50E-15	-

CCG020474.1	AT2G39990.1	TIF3F1	Eukaryotic translation initiation factor 3 subunit F	0.2	1.50E-15	-
CCG020487.1	AT2G39920.1		Uncharacterized protein At2g39920	0.182	2.45E-05	-
CCG020942.1	AT3G03130.1			0.268	2.20E-16	PeW-PeE; PeW-Pp
CCG020944.2	AT5G17165.1			0.367	1.90E-08	PeW-PeE; PeW-Pp
CCG020951.2	AT5G16870.1			0.171	1.17E-03	-
CCG021011.1	AT1G72570.1	AIL1	AP2-like ethylene-responsive transcription factor AIL1	0.193	2.20E-16	-
CCG021041.1	AT1G72450.1	TIFY11B	Protein TIFY 11B	0.437	2.20E-16	-
CCG021043.1	AT3G14150.1	GLO3	Peroxisomal (S)-2-hydroxy-acid oxidase GLO3	0.272	1.17E-03	-
CCG021158.1	AT3G20940.1			0.173	1.51E-04	-
CCG021638.1	AT1G13210.1	ALA11	Probable phospholipid-transporting ATPase 11	0.174	2.20E-16	PeC-PeW
CCG021829.1	AT1G63660.1			0.278	2.20E-16	-
CCG022132.1	AT5G27110.1			0.184	2.20E-16	PeW-Pp
CCG022350.1	AT1G04990.2		Zinc finger CCCH domain-containing protein 3	0.184	2.20E-16	PeC-PeW
CCG023064.1	AT5G16420.1			0.207	1.50E-15	-
CCG023502.1	AT2G03380.1			0.204	2.20E-16	-
CCG023504.1	AT1G22400.1	UGT85A1	UDP-glycosyltransferase 85A1	0.177	2.20E-16	PeC-PeE
CCG023560.1	AT3G57710.1			0.177	2.20E-16	-
CCG023669.2	AT2G32400.1	GLR5		0.202	2.20E-16	-
CCG024240.1	AT5G22090.1		Protein FAF-like, chloroplastic	0.217	1.17E-03	PeC-PeW; PeW-PeE
CCG024383.1	AT1G12330.1			0.183	2.80E-08	-
CCG024550.1	AT5G61370.1			0.178	2.20E-16	-
CCG025286.1	AT1G47990.1	GA2OX4	Gibberellin 2-beta-dioxygenase 4	0.174	1.90E-08	-
CCG025460.1	AT1G16390.1	43011	Organic cation/carnitine transporter 3	0.193	1.65E-06	-
CCG025514.1	AT3G53180.1			0.215	1.17E-03	-
CCG025638.1	AT5G11320.1	YUC4	Probable indole-3-pyruvate monooxygenase YUCCA4	0.232	2.20E-16	-
CCG025886.1	AT3G49500.1	RDR6	RNA-dependent RNA polymerase 6	0.19	2.20E-16	PeW-Pp; PeE-Pp
CCG025980.1	AT2G03500.1	EFM	Myb family transcription factor EFM	0.185	2.20E-16	-
CCG026125.2	AT5G56180.1	ARP8	Actin-related protein 8	0.364	1.65E-06	-
CCG026226.1	AT1G61700.1			0.283	2.20E-16	-
CCG026337.2	-			0.258	1.90E-08	PeC-PeW; PeW-PeE; PeE-Pp
CCG026699.1	AT1G68020.2	TPS6	Alpha,alpha-trehalose-phosphate synthase [UDP-forming] 6	0.239	1.65E-06	-
CCG026803.1	AT5G19040.1	IPT5	Adenylate isopentenyltransferase 5, chloroplastic	1.189	1.17E-03	PeC-PeW; PeC-Pp; PeW-PeE
CCG026828.1	AT4G28390.1	AAC3	ADP,ATP carrier protein 3, mitochondrial	1.722	1.17E-03	-
CCG026862.1	AT5G42260.1	BGLU12	Beta-glucosidase 12	0.516	2.20E-16	PeC-PeW; PeW-Pp
CCG027253.6	AT3G26730.1			0.312	1.51E-04	-
CCG027260.1	AT3G02130.1	RPK2	LRR receptor-like serine/threonine-protein kinase RPK2	0.171	3.27E-16	-
CCG027408.2	-			0.329	1.90E-08	-
CCG027437.1	AT1G19270.1	DA1	Protein DA1	0.353	2.20E-16	PeW-Pp
CCG027486.1	AT5G65140.1	TPPJ	Probable trehalose-phosphate phosphatase J	0.174	2.74E-12	-
CCG027495.1	AT1G78160.1	APUM7	Putative pumilio homolog 7, chloroplastic	0.181	2.90E-14	-
CCG027528.1	AT5G10360.1	RPS6B	40S ribosomal protein S6-2	0.517	2.20E-16	PeW-Pp
CCG027560.1	AT1G28200.1	FIP1	GEM-like protein 1	0.215	1.65E-06	PeE-Pp
CCG027597.1	AT1G09960.1	SUC4	Sucrose transport protein SUC4	0.242	2.20E-16	PeW-Pp
CCG027598.1	AT4G29140.1	DTX51	Protein DETOXIFICATION 51	0.165	3.99E-13	PeW-Pp

CCG027795.1	AT5G42680.1			0.177	1.65E-06	PeW-Pp
CCG027918.1	AT3G60540.1			0.266	2.74E-12	-
CCG027939.1	AT3G60740.1	TFCD	Tubulin-folding cofactor D	0.395	2.67E-16	-
CCG027964.1	AT2G45420.1	LBD18	LOB domain-containing protein 18	0.169	1.90E-08	-
CCG028113.1	AT1G31930.1	XLG3	Extra-large guanine nucleotide-binding protein 3	0.183	2.20E-16	-
CCG028159.1	AT4G38220.2			0.167	3.24E-03	-
CCG028227.1	AT4G16447.1			0.166	1.17E-03	-
CCG028236.1	AT1G68020.2	TPS6	Alpha,alpha-trehalose-phosphate synthase [UDP-forming] 6	0.193	1.65E-06	-
CCG028331.1	AT4G10610.1	CID12	Polyadenylate-binding protein-interacting protein 12	0.255	2.20E-16	-
CCG028336.1	AT5G42110.1			0.239	5.31E-11	PeC-PeW; PeW-PeE; PeW-Pp
CCG028337.1	AT3G15850.1			0.184	2.20E-16	-
CCG028353.1	AT4G10490.1	DLO2	Protein DMR6-LIKE OXYGENASE 2	0.504	2.74E-12	PeW-Pp
CCG028396.1	AT3G07580.2			0.186	1.17E-03	-
CCG028680.1	AT3G04750.1			0.283	2.74E-12	PeC-PeW; PeC-Pp
CCG028825.1	AT2G17080.1			0.422	2.20E-16	PeC-PeW; PeW-Pp
CCG028874.1	-			2.514	1.17E-03	-
CCG028875.1	AT3G29160.1	KIN11	SNF1-related protein kinase catalytic subunit alpha KIN11	0.323	2.74E-12	-
CCG029063.1	AT2G46410.1	CPC	Transcription factor CPC	0.608	1.17E-03	-
CCG029166.1	AT4G32810.1	CCD8	Carotenoid cleavage dioxygenase 8, chloroplastic	0.17	2.45E-08	-
CCG029264.1	AT3G51490.2	MSSP3	Monosaccharide-sensing protein 3	0.548	2.20E-16	PeC-PeW; PeC-PeE; PeW-Pp; PeE-Pp
CCG029265.1	AT1G20830.1	MCD1	Protein MULTIPLE CHLOROPLAST DIVISION SITE 1	0.351	1.90E-08	PeC-PeE; PeW-Pp; PeE-Pp
CCG029726.1	AT4G23760.1			0.254	1.17E-03	-
CCG030167.1	AT1G01320.1			1.372	2.74E-12	PeW-Pp
CCG030251.1	AT5G25220.1	KNAT3	Homeobox protein knotted-1-like 3	0.25	5.31E-11	-
CCG030280.1	AT3G06035.1		Uncharacterized GPI-anchored protein At3g06035	0.341	1.17E-03	-
CCG030328.1	AT5G06600.1	UBP12	Ubiquitin carboxyl-terminal hydrolase 12	0.178	2.45E-05	-
CCG030361.1	AT2G35910.1	ATL70	RING-H2 finger protein ATL70	0.187	1.17E-03	-
CCG030376.1	AT2G35940.1	BLH1	BEL1-like homeodomain protein 1	0.229	2.20E-16	-
CCG030814.1	AT3G46780.1	PTAC16	Protein plastid transcriptionally active 16, chloroplastic	0.169	2.90E-14	-
CCG031053.1	AT5G12870.1	MYB46	Transcription factor MYB46	0.217	1.65E-06	-
CCG031057.1	AT1G51350.1			0.176	1.92E-07	PeC-PeW
CCG031175.1	AT3G62980.1	TIR1	Protein TRANSPORT INHIBITOR RESPONSE 1	0.173	6.05E-03	-
CCG031258.1	AT3G47600.1			0.274	2.67E-16	PeC-PeW
CCG031341.2	AT1G74320.1		Probable choline kinase 2	0.17	7.99E-14	-
CCG032085.1	AT1G06490.1	CALS7	Callose synthase 7	0.196	2.20E-16	-
CCG032086.1	AT4G21700.1			0.339	2.20E-16	-
CCG032087.1	AT2G33770.1	UBC24	Probable ubiquitin-conjugating enzyme E2 24	0.167	2.20E-16	-
CCG032313.1	AT4G27280.1			0.17	1.65E-06	PeC-PeW; PeW-PeE; PeW-Pp
CCG032560.1	-			0.211	2.20E-16	PeW-Pp
CCG032579.1	AT1G79660.1			0.166	1.17E-03	-
CCG032642.1	AT3G18830.1	PLT5	Polyol transporter 5	0.189	1.65E-06	-
CCG032688.1	AT5G43580.1			0.227	2.45E-05	-
CCG032846.1	AT5G57340.2			1.75	1.17E-03	-

CCG032914.1	AT3G22430.1			0.202	2.20E-16	-
CCG032916.1	AT4G14910.2	HISN5B	Imidazoleglycerol-phosphate dehydratase 2, chloroplastic	0.224	2.20E-16	-
CCG032917.2	AT4G14920.1			0.201	2.20E-16	PeW-Pp
CCG033044.2	AT4G17720.1			0.173	1.65E-06	-
CCG033165.1	AT4G37370.1			0.393	2.20E-16	-
CCG033437.1	AT3G07390.1			0.427	2.20E-16	-
CCG033549.1	AT1G77700.1			2.016	1.17E-03	-
CCG033671.1	AT3G20640.1	BHLH123	Transcription factor bHLH123	0.205	2.20E-16	PeC-PeW
CCG034214.1	AT3G29590.1	5MAT		0.17	2.20E-16	-
Pp						
CCG000166.1	AT3G55800.1		Sedoheptulose-1,7-bisphosphatase, chloroplastic	0.348	3.56E-06	-
CCG000168.1	AT5G58910.1	LAC16	Laccase-16	0.234	2.20E-16	-
CCG000169.1	AT3G11280.1			0.276	2.20E-16	-
CCG000223.2	AT2G40200.1	BHLH51	Transcription factor bHLH51	0.218	2.20E-16	-
CCG000368.4	AT1G26760.1			0.255	2.20E-16	PeW-Pp; PeE-Pp
CCG000387.1	AT2G03340.1	WRKY3	Probable WRKY transcription factor 3	0.267	2.20E-16	-
CCG000388.1	AT1G13950.1	ELF5A-1	Eukaryotic translation initiation factor 5A-1	0.321	3.34E-11	-
CCG000389.1	AT1G69400.1	BUB3.3	Mitotic checkpoint protein BUB3.3	0.241	2.20E-16	PeW-Pp
CCG000391.2	AT1G69390.1	MINE1	Cell division topological specificity factor homolog	0.227	2.20E-16	PeW-Pp
CCG001766.1	AT4G34760.1			1.022	3.34E-11	-
CCG002048.1	-			0.177	2.20E-16	-
CCG002483.1	ATMG00980.1			0.308	1.83E-04	-
CCG003880.1	AT3G58070.1	GIS	Zinc finger protein GIS	0.215	1.00E-15	-
CCG003977.1	AT3G28857.1	PRE5	Transcription factor PRE5	0.234	3.56E-06	-
CCG004079.1	AT1G52560.1	HSP26.5	26.5 kDa heat shock protein, mitochondrial	0.281	2.20E-16	-
CCG004882.1	AT5G15460.1	MUB2	Membrane-anchored ubiquitin-fold protein 2	0.398	2.20E-16	-
CCG007600.1	AT1G52565.1			0.183	1.83E-04	-
CCG007687.1	AT3G29290.1	EMB2076	Pentatricopeptide repeat-containing protein At3g29290	0.392	2.20E-16	PeW-Pp
CCG009705.1	AT3G16480.1	MPPA2		0.181	2.20E-16	-
CCG009892.1	AT5G45600.1	GAS41		0.818	3.34E-11	PeW-Pp
CCG012322.1	AT5G56870.1	BGAL4	Beta-galactosidase 4	0.171	2.20E-16	-
CCG013906.1	AT5G57030.1	LUT2	Lycopene epsilon cyclase, chloroplastic	0.178	2.20E-16	PeC-PeE
CCG014183.1	AT5G18190.1			0.18	2.20E-16	-
CCG014368.1	AT5G45560.1	EDR2L	Protein ENHANCED DISEASE RESISTANCE 2-like	0.209	2.20E-16	PeW-Pp
CCG015757.2	AT1G08070.1			0.245	2.20E-16	PeW-Pp; PeE-Pp
CCG015909.2	AT2G19480.1	NAP1;2	Nucleosome assembly protein 1;2	0.245	3.56E-06	-
CCG017318.1	AT1G75250.1	RL6	Protein RADIALIS-like 6	0.175	2.20E-16	-
CCG017962.1	AT1G71240.1			0.294	2.20E-16	PeW-Pp; PeE-Pp
CCG017971.1	AT3G56960.1	PIP5K4	Phosphatidylinositol 4-phosphate 5-kinase 4	0.198	2.20E-16	-
CCG018082.1	AT5G50340.1			0.19	2.20E-16	-
CCG019008.1	AT3G61590.1		F-box/kelch-repeat protein At3g61590	0.356	2.20E-16	-
CCG019576.1	-			0.205	3.56E-06	-
CCG020063.2	AT1G21610.3			0.206	1.10E-11	-

CCG020528.1	AT3G53850.1		CASP-like protein 5B2	0.262	2.20E-16	-
CCG020756.3	AT5G13230.1			0.434	2.20E-16	PeW-Pp
CCG023682.1	-			0.537	3.34E-11	PeE-Pp
CCG025143.1	AT1G71960.1	ABCG25	ABC transporter G family member 25	0.196	2.20E-16	-
CCG025144.1	-			0.354	2.20E-16	-
CCG025162.1	AT1G13570.1		F-box/FBD/LRR-repeat protein At1g13570	0.336	2.20E-16	-
CCG025167.1	AT1G13580.2		LAG1 longevity assurance homolog 3	0.537	2.20E-16	-
CCG025169.1	AT3G13470.1	CPN60B2	Chaperonin 60 subunit beta 2	0.273	2.20E-16	PeW-Pp
CCG025226.1	AT3G06670.2			0.188	2.20E-16	PeW-Pp
CCG025227.2	AT4G16160.2	OEP162	Outer envelope pore protein 16-2	0.29	2.20E-16	-
CCG025229.1	AT4G16265.1	NRPB9B	DNA-directed RNA polymerases II, IV and V subunit 9B	0.279	3.34E-11	PeE-Pp
CCG025230.1	AT2G29360.1	SDR	Tropinone reductase homolog At2g29360	0.765	2.20E-16	PeE-Pp
CCG025245.1	AT3G17040.1	HCF107	Protein high chlorophyll fluorescent 107	0.194	2.20E-16	PeW-Pp
CCG025251.1	AT1G09570.1	PHYA	Phytochrome A	0.293	2.20E-16	-
CCG025286.1	AT1G47990.1	GA2OX4	Gibberellin 2-beta-dioxygenase 4	0.19	2.20E-16	-
CCG025287.1	AT5G24460.1			0.238	2.20E-16	-
CCG026334.1	AT3G02150.2	TCP13	Transcription factor TCP13	0.217	2.20E-16	-
CCG026974.1	AT1G22030.1			0.235	2.20E-16	-
CCG027524.1	AT4G08960.1			0.234	2.20E-16	-
CCG027598.1	AT4G29140.1	DTX51	Protein DETOXIFICATION 51	0.277	2.20E-16	PeW-Pp
CCG028345.1	AT3G61460.1			0.278	3.34E-11	-
CCG029543.1	AT1G73250.1	GER1	GDP-L-fucose synthase 1	0.221	2.20E-16	-
CCG029619.1	AT4G24150.1	GRF8	Growth-regulating factor 8	0.205	2.20E-16	-
CCG029626.1	AT1G14430.1			0.252	2.20E-16	PeW-Pp
CCG029627.1	AT2G46680.1	ATHB-7	Homeobox-leucine zipper protein ATHB-7	0.312	2.20E-16	PeW-Pp
CCG029628.1	AT1G26665.2	MED10B	Mediator of RNA polymerase II transcription subunit 10b	0.439	2.20E-16	PeW-Pp
CCG029646.1	-			1.041	3.34E-11	-
CCG029648.1	AT1G64625.1	BHLH157	Transcription factor bHLH157	0.222	2.20E-16	-
CCG029649.1	AT3G05675.2		BTB/POZ domain-containing protein At3g05675	0.198	2.20E-16	-
CCG030167.1	AT1G01320.1			1.363	3.34E-11	PeW-Pp
CCG030227.1	AT5G58030.1			0.271	2.20E-16	-
CCG030296.1	AT5G56770.1	ET3	Protein EFFECTOR OF TRANSCRIPTION 3	0.204	2.20E-16	PeW-Pp
CCG030309.1	AT5G15210.1	ZHD8	Zinc-finger homeodomain protein 8	0.26	2.20E-16	-
CCG032519.1	AT5G10970.1			0.243	2.20E-16	-
CCG032524.1	AT1G80760.1	NIP6-1	Aquaporin NIP6-1	0.75	2.20E-16	PeW-Pp
CCG033618.1	AT3G21330.1	BHLH87	Transcription factor bHLH87	0.242	2.20E-16	-
CCG033877.1	AT5G15100.1	PIN8	Auxin efflux carrier component 8	0.199	2.20E-16	-
CCG033880.1	AT3G28720.1			0.2	2.20E-16	-
CCG033936.1	AT5G25370.1		Phospholipase D alpha 3	0.249	2.20E-16	-
CCG034089.1	AT1G19850.1	IAA24	Auxin response factor 5	0.295	2.20E-16	PeW-Pp; PeE-Pp
CCG034090.1	AT3G27230.1			0.248	2.20E-16	PeW-Pp; PeE-Pp

Table S9: Results from Gene Ontology enrichment analysis for PSGs.

GO ID	GO name	Ontology	PSGs	Annotated	*P*-value	FDR
PeC						
GO:0060776	simple leaf morphogenesis	BP	1	3	4.39E-11	1.22E-09
GO:0060968	regulation of gene silencing	BP	1	8	1.96E-05	2.21E-04
GO:0009688	abscisic acid biosynthetic process	BP	1	11	2.80E-04	2.38E-03
GO:0048467	gynoecium development	BP	1	13	8.88E-04	6.76E-03
GO:0016114	terpenoid biosynthetic process	BP	3	102	1.43E-03	1.05E-02
GO:0048438	floral whorl development	BP	1	15	2.13E-03	1.44E-02
GO:0010073	meristem maintenance	BP	1	17	4.22E-03	2.74E-02
GO:0009631	cold acclimation	BP	1	19	7.34E-03	4.36E-02
GO:0015205	nucleobase transmembrane transporter activity	MF	2	15	4.60E-10	1.24E-08
GO:0019207	kinase regulator activity	MF	2	42	3.54E-04	2.98E-03
GO:0016653	oxidoreductase activity	MF	1	15	2.13E-03	1.44E-02
PeW						
GO:0048765	root hair cell differentiation	BP	3	26	2.88E-11	1.18E-09
GO:0048764	trichoblast maturation	BP	3	26	2.88E-11	1.18E-09
GO:2000024	regulation of leaf development	BP	1	4	3.73E-08	5.77E-07
GO:0021700	developmental maturation	BP	4	68	1.02E-07	1.49E-06
GO:0010087	phloem or xylem histogenesis	BP	4	72	2.68E-07	3.81E-06
GO:0048364	root development	BP	7	223	2.83E-06	3.43E-05
GO:0048856	anatomical structure development	BP	19	1276	7.42E-05	5.46E-04
GO:0010118	stomatal movement	BP	2	36	2.78E-04	1.88E-03
GO:0070887	cellular response to chemical stimulus	BP	11	677	9.83E-04	5.25E-03
GO:0010229	inflorescence development	BP	1	13	1.94E-03	9.43E-03
GO:0048440	carpel development	BP	1	16	5.88E-03	2.52E-02
GO:0005769	early endosome	CC	2	17	4.16E-08	6.34E-07
GO:0090406	pollen tube	CC	2	31	7.21E-05	5.34E-04
Pp						
GO:0003156	regulation of organ formation	BP	1	2	2.20E-16	8.19E-15
GO:0010026	trichome differentiation	BP	1	18	1.75E-05	1.48E-04
GO:0048510	regulation of timing of transition from vegetative to reproductive phase	BP	1	21	7.52E-05	5.79E-04
GO:0055072	iron ion homeostasis	BP	1	22	1.13E-04	8.48E-04
GO:0009880	embryonic pattern specification	BP	1	26	4.23E-04	2.87E-03
GO:0048507	meristem development	BP	1	28	7.18E-04	4.54E-03
GO:0009583	detection of light stimulus	BP	1	29	9.12E-04	5.41E-03
GO:0010305	leaf vascular tissue pattern formation	BP	1	36	3.41E-03	1.57E-02
GO:0009881	photoreceptor activity	MF	1	8	1.47E-10	2.66E-09
GO:0016679	oxidoreductase activity	MF	2	100	6.98E-04	4.54E-03
GO:0008135	translation factor activity	MF	3	206	7.72E-04	4.79E-03

Table S10: List of PSGs with putative function associated with ecological and morphological divergence between lineages.

Gene ID	Best Arabidopsis hit	Gene Name	References
Genes related to trichome differentiation and branching			
CCG003880.1	AT3G58070.1	GIS	(35, 36)
CCG002566.1	AT5G14750.1	MYB66	(37-39)
CCG029063.1	AT2G46410.1	CPC	（40, 41）
CCG017558.1	AT4G33880.1	RSL2	(42)
Genes involved in regulation of leaf development and morphology			
CCG026334.1	AT3G02150.2	TCP13	(43, 44)
CCG020711.1	AT1G27340.1	LCR	(45, 46)
CCG025886.1	AT3G49500.1	RDR6	(47)
CCG010761.1	AT2G46310.1	CRF5	(48)
CCG025638.1	AT5G11320.1	YUC4	(49)
CCG014445.1	AT5G46700.1	TRN2	(50)
CCG020746.1	AT5G13180.1	NAC083	(51, 52)
CCG034089.1	AT1G19850.1	IAA24	(53)
Genes associated with flowering time			
CCG025980.1	AT2G03500.1	EFM	(54)
CCG018730.1	AT5G40490.1	HLP1	(55)
CCG017317.1	AT4G34530.1	CIB1	(56)
CCG009892.1	AT5G45600.1	GAS41	(57, 58)
Genes associated with anther development			
CCG027260.1	AT3G02130.1	RPK2	(59)
Genes associated with pollen tube growth and guidance			
CCG023669.2	AT2G32400.1	GLR5	(60)
CCG017971.1	AT3G56960.1	PIP5K4	(61)
CCG016375.1	AT3G45040.1	DOK1	(62, 63)
Genes involved in the development of male and female gametophytes			
CCG019053.1	AT2G46040.1	ARID1	(64, 65)
CCG033877.1	AT5G15100.1	PIN8	(66)
CCG032484.1	AT3G23940.1	DHAD	(67)
CCG000332.1	AT1G16280.1	RH36	(68)

SI References

1. Wang J, et al. (2014) Speciation of two desert poplar species triggered by Pleistocene climatic oscillations. *Heredity* 112(2):156–164.
2. Wang J, et al. (2011) Genetic differentiation and delimitation between ecologically diverged *Populus euphratica* and *P. pruinosa*. *PLOS ONE* 6(10):e26530.
3. Ma T, et al. (2013) Genomic insights into salt adaptation in a desert poplar. *Nat Commun* 4:2797.
4. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
5. Li H, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
6. DePristo MA, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5):491–498.
7. Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27(21):2987–2993.
8. Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of Next Generation Sequencing Data. *BMC bioinform* 15:356.
9. Kim SY, et al. (2011) Estimation of allele frequency and association mapping using next-generation sequencing data. *BMC bioinform* 12.
10. Nielsen R, Korneliussen T, Albrechtsen A, Li Y, Wang J (2012) SNP calling, genotype calling, and sample allele frequency estimation from New-Generation Sequencing data. *PLOS ONE* 7(7):e37558.
11. Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 81(5):1084–1097.
12. Browning BL, Browning SR (2013) Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics* 194(2):459–471.
13. Manichaikul A, et al. (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics* 26(22):2867–2873.
14. Bryant D, Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol* 21(2):255–265.
15. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23(2):254–267.
16. Tamura K, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731–2739.
17. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLOS Genet* 2(12):e190.
18. Skotte L, Korneliussen TS, Albrechtsen A (2013) Estimating individual admixture proportions from next generation sequencing data. *Genetics* 195(3):693–702.
19. Pickrell JK, Pritchard JK (2012) Inference of population splits and mixtures from genome-wide allele frequency data. *PLOS Genet* 8(11):e1002967.
20. Li H, Durbin R (2011) Inference of human population history from individual whole-genome

sequences. *Nature* 475(7357):493–496.

21. Wang J, Street NR, Scofield DG, Ingvarsson PK (2016) Variation in linked selection and recombination drive genomic divergence during allopatric speciation of european and american aspens. *Mol Biol Evol* 33(7):1754–1767.
22. Excoffier L, Dupanloup I, Huerta-Sanchez E, Sousa VC, Foll M (2013) Robust demographic inference from genomic and SNP data. *PLOS Genet* 9(10):e1003905.
23. Barrett JC (2009) Haploview: Visualization and analysis of SNP genotype data. *Cold Spring Harbor protocols* 2009(10):pdb ip71.
24. McVean GA, et al. (2004) The fine-scale structure of recombination rate variation in the human genome. *Science* 304(5670):581–584.
25. Korneliussen TS, Moltke I, Albrechtsen A, Nielsen R (2013) Calculation of Tajima's *D* and other neutrality test statistics from low depth next-generation sequencing data. *BMC bioinform* 14.
26. Fumagalli M, et al. (2013) Quantifying population genetic differentiation from next-generation sequencing data. *Genetics* 195(3):979–992.
27. Fumagalli M, Vieira FG, Linderoth T, Nielsen R (2014) ngsTools: methods for population genetics analyses from next-generation sequencing data. *Bioinformatics* 30(10):1486–1487.
28. Ewing G, Hermisson J (2010) *MSMS*: a coalescent simulation program including recombination, demographic structure and selection at a single locus. *Bioinformatics* 26(16):2064–2065.
29. Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics* 116(1):153–159.
30. Yi X, et al. (2010) Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329(5987):75–78.
31. Liu S, et al. (2014) Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. *Cell* 157(4):785–794.
32. Conesa A, Gotz S (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics* 2008:619832.
33. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*:289–300.
34. Liu X, Wang Z, Shao W, Ye Z, Zhang J (2016) Phylogenetic and taxonomic status analyses of the Abaso section from multiple nuclear genes and plastid fragments reveal new insights into the north america origin of *Populus* (Salicaceae). *Frontiers Plant Sci* 7:2022.
35. Gan Y, et al. (2006) GLABROUS INFLORESCENCE STEMS modulates the regulation by gibberellins of epidermal differentiation and shoot maturation in *Arabidopsis*. *Plant Cell* 18(6):1383–1395.
36. Sun L, et al. (2015) GLABROUS INFLORESCENCE STEMS3 (GIS3) regulates trichome initiation and development in *Arabidopsis*. *New Phytologist* 206(1):220-230.
37. Ishida T, et al. (2007) *Arabidopsis* TRANSPARENT TESTA GLABRA2 is directly regulated by R2R3 MYB transcription factors and is involved in regulation of GLABRA2 transcription in epidermal differentiation. *Plant Cell* 19(8):2531-2543.
38. Seo E, et al. (2011) WEREWOLF, a regulator of root hair pattern formation, controls flowering time through the regulation of FT mRNA stability. *Plant Physiology* 156(4):1867-1877.
39. Pesch M, Hülskamp M. (2004) Creating a two-dimensional pattern *de novo* during *Arabidopsis*

trichome and root hair initiation. *Current opinion in genetics & development* 14(4):422-427.

40. Schellmann S, et al. (2002) TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in *Arabidopsis*. *The EMBO journal* 21(19):5036-5046.
41. Digiuni S, et al. (2008) A competitive complex formation mechanism underlies trichome patterning on *Arabidopsis* leaves. *Molecular Systems Biology* 4(1):217.
42. Yi K, et al. (2010) A basic helix-loop-helix transcription factor controls cell growth and size in root hairs. *Nature Genetics* 42(3):264-267.
43. Koyama T, Furutani M, Tasaka M, Ohme-Takagi M (2007) TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in *Arabidopsis*. *Plant Cell* 19(2):473–484.
44. Koyama T, Mitsuda N, Seki M, Shinozaki K, Ohme-Takagi M (2010) TCP transcription factors regulate the activities of ASYMMETRIC LEAVES1 and miR164, as well as the auxin response, during differentiation of leaves in *Arabidopsis*. *Plant Cell* 22(11):3574–3588.
45. Song JB, Huang SQ, Dalmay T, Yang ZM (2012) Regulation of leaf morphology by microRNA394 and its target LEAF CURLING RESPONSIVENESS. *Plant Cell Physiol* 53(7):1283–1294.
46. Song JB, et al. (2013) miR394 and LCR are involved in *Arabidopsis* salt and drought stress responses in an abscisic acid-dependent manner. *BMC Plant Biol* 13:210.
47. Adenot X, et al. (2006) DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. *Current Biology* 16(9):927-932.
48. Rashotte A, et al. (2006) A subset of *Arabidopsis* AP2 transcription factors mediates cytokinin responses in concert with a two-component pathway. *Proceedings of the National Academy of Sciences* 103(29):11081-11085.
49. Cheng Y, Dai X, Zhao Y. (2007) Auxin synthesized by the YUCCA flavin monooxygenases is essential for embryogenesis and leaf formation in *Arabidopsis*. *Plant Cell* 19(8):2430-2439.
50. Cnops G, et al. (2006) The TORNADO1 and TORNADO2 genes function in several patterning processes during early leaf development in *Arabidopsis thaliana*. *Plant Cell* 18(4):852-866.
51. Seo P, Park C. (2011) Signaling linkage between environmental stress resistance and leaf senescence in *Arabidopsis*. *Plant signaling & behavior* 6(10):1564-1566.
52. Yang S, et al. (2011) The *Arabidopsis* NAC transcription factor VNI2 integrates abscisic acid signals into leaf senescence via the COR/RD genes. *Plant Cell* 23(6):2155-2168.
53. Krogan N, et al. Deletion of MP/ARF5 domains III and IV reveals a requirement for Aux/IAA regulation in *Arabidopsis* leaf vascular patterning. *New Phytologist* 194(2):391-401.
54. Yan Y, et al. (2014) A MYB-domain protein EFM mediates flowering responses to environmental cues in *Arabidopsis*. *Developmental cell* 30(4):437-448.
55. Zhang Y, et al. (2015) Integrative genome-wide analysis reveals HLP1, a novel RNA-binding protein, regulates plant flowering by targeting alternative polyadenylation. *Cell research* 25(7):864.
56. Liu Y, et al. (2013) Multiple bHLH proteins form heterodimers to mediate CRY2-dependent regulation of flowering-time in *Arabidopsis*. *PLoS genetics* 9(10):e1003861.
57. Su Y, et al. (2017) Phosphorylation of Histone H2A at Serine 95: A Plant-specific Mark Involved in Flowering Time Regulation and H2A. Z Deposition. *Plant Cell*: tpc.00266.2017.
58. Zacharaki V, et al. (2012) The *Arabidopsis* ortholog of the YEATS domain containing protein YAF9a regulates flowering by controlling H4 acetylation levels at the FLC locus. *Plant science*

196:44-52.

59. Zhao D. (2009) Control of anther cell differentiation: a teamwork of receptor-like kinases. *Sexual plant reproduction* 22(4):221-228.
60. Michard E, et al. (2011) Glutamate receptor–like genes form Ca^{2+} channels in pollen tubes and are regulated by pistil D-serine. *Science* 332(6028):434-437.
61. Sousa E, Kost B, Malhó R. (2008) *Arabidopsis* phosphatidylinositol-4-monophosphate 5-kinase 4 regulates pollen tube growth and polarity by modulating membrane recycling. *Plant Cell* 20(11):3050-3064.
62. Cho Y, et al. (2017) *Arabidopsis* dolichol kinase AtDOK1 is involved in flowering time control. *Journal of Experimental Botany* erx095.
63. Kanehara K, et al. (2015) *Arabidopsis* DOK1 encodes a functional dolichol kinase involved in reproduction. *Plant Journal* 81(2):292-303.
64. Li L, et al. (2017) A reciprocal inhibition between ARID1 and MET1 in male and female gametes in *Arabidopsis*. *Journal of Integrative Plant Biology*.
65. Zheng B, et al. (2014) An ARID domain-containing protein within nuclear bodies is required for sperm cell formation in *Arabidopsis thaliana*. *PLoS genetics* 10(7):e1004421.
66. Ding Z, et al. (2012) ER-localized auxin transporter PIN8 regulates auxin homeostasis and male gametophyte development in *Arabidopsis*. *Nature communications* 3:941.
67. Zhang C, et al. (2014) Dihydroxyacid dehydratase is important for gametophyte development and disruption causes increased susceptibility to salinity stress in *Arabidopsis*. *Journal of experimental botany* 66(3):879-888.
68. Huang C K, et al. (2010) A DEAD-box protein, AtRH36, is essential for female gametophyte development and is involved in rRNA biogenesis in *Arabidopsis*. *Plant and cell physiology* 51(5):694-706.