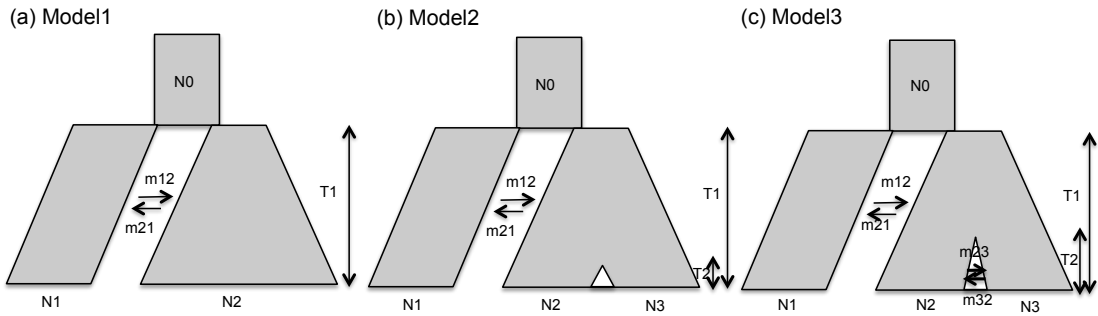


1 **Supplementary Figures S5-S13**

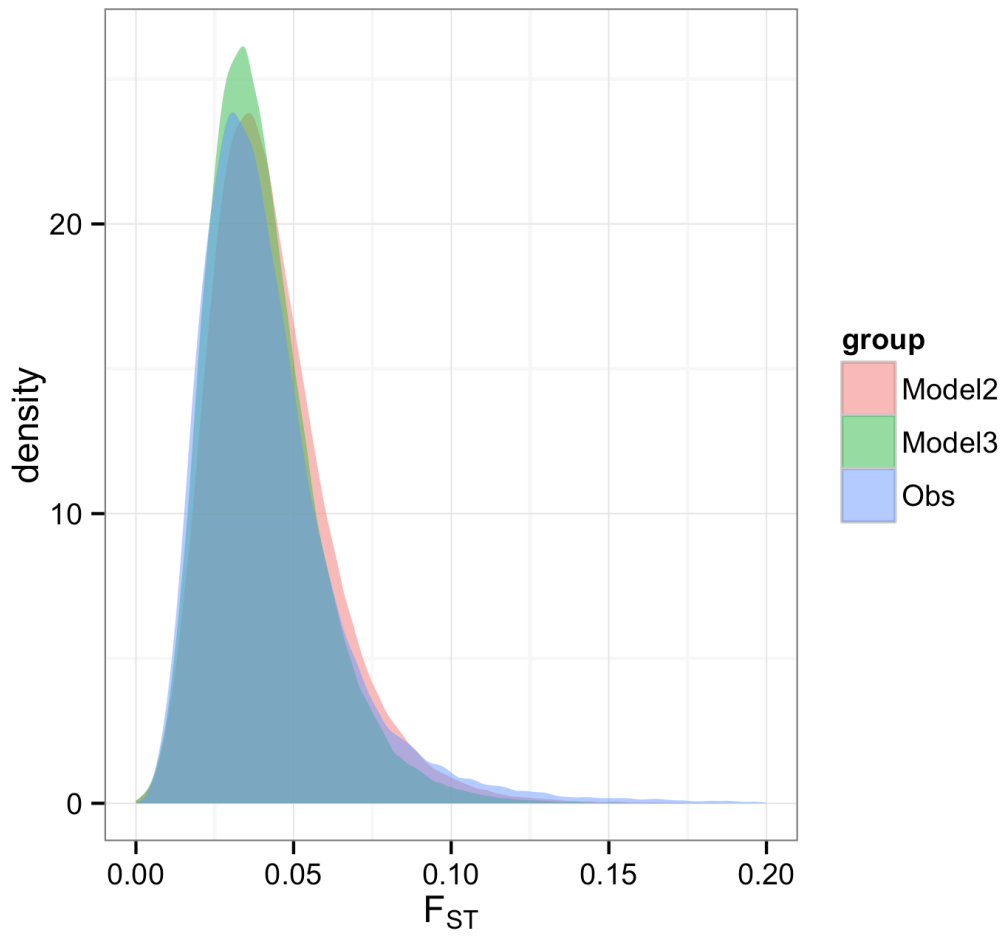
2
3
4
5



6

7 **Figure S5.** Demographic models used for the coalescent simulation. (a) Model1, the
8 best-fitting model inferred by *fastsimcoal2*. N1, N2 and N0 denote the effective
9 population size of *P. tremula*, *P. tremuloides* and their ancestral population. m12,
10 m21, T1 indicate the migration rates and divergence time between the two species,
11 respectively. (b) Model2, the best-fitting model inferred by *fastsimcoal2* but with the
12 population subdivision of *P. tremuloides*, where we assume no gene flow occurred
13 between the two subpopulations. (c) Model3, the best-fitting model inferred by
14 *fastsimcoal2* but with the population subdivision of *P. tremuloides*, where we assume
15 that gene flow occurred between the two subpopulations. All parameters in model 2
16 and 3 are the same as in model 1 except N2 and N3 represent the effective population
17 sizes of population Alberta and Wisconsin in *P. tremuloides*, respectively. m23, m32
18 and T2 indicate the migration rates and the time of divergence between the two
19 subpopulations of *P. tremuloides*. The demographic parameters used in each model
20 are shown in Table S5.

21
22
23
24
25
26
27



28

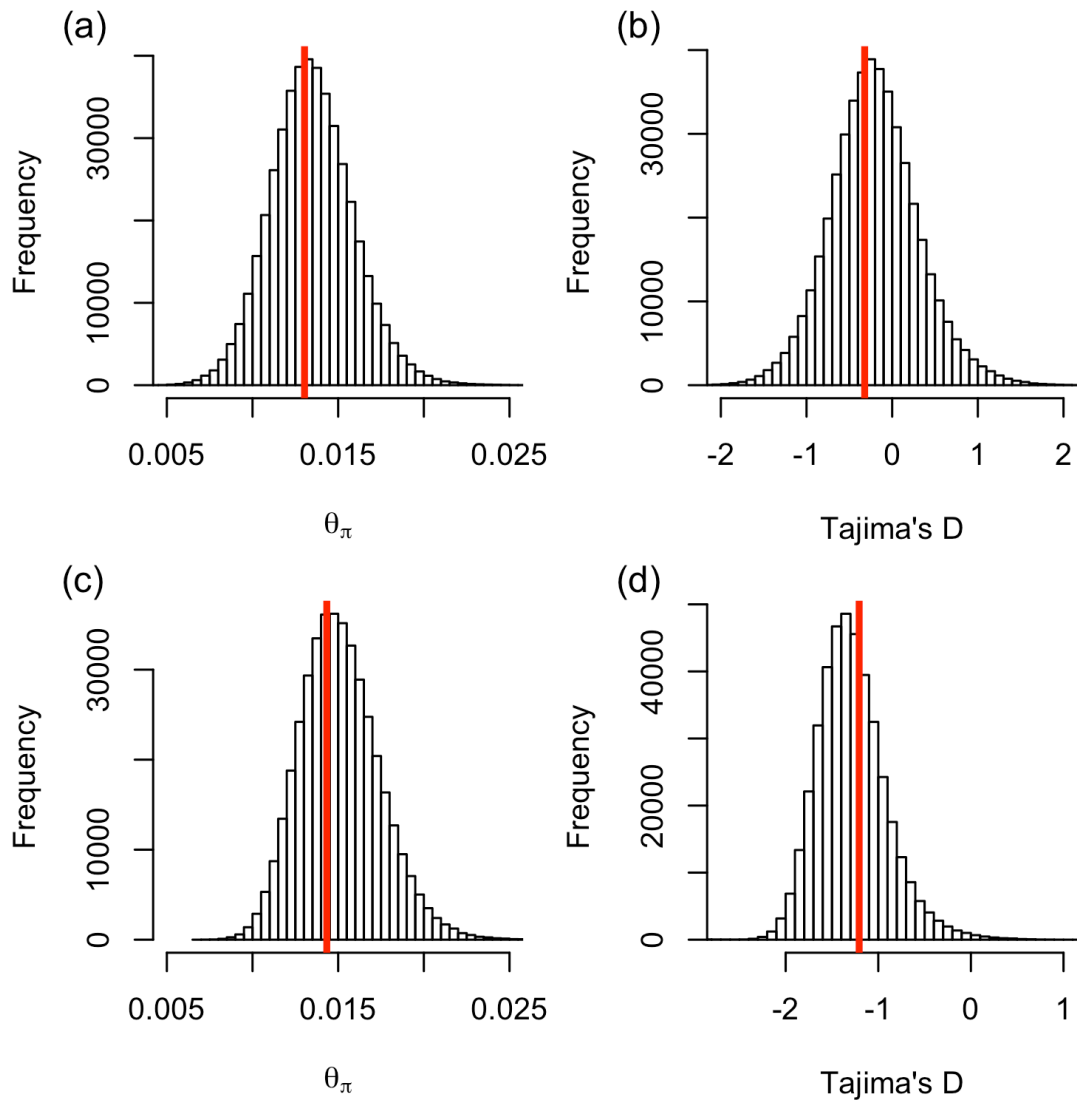
29

30 **Figure S6.** The distribution of F_{ST} values between the subpopulation Alberta and
31 Wisconsin of *P. tremuloides* in the observed data (blue) and in the simulated data
32 corresponding to model 2 (red) and model3 (green) as shown in Figure S5.

33

34

35



36

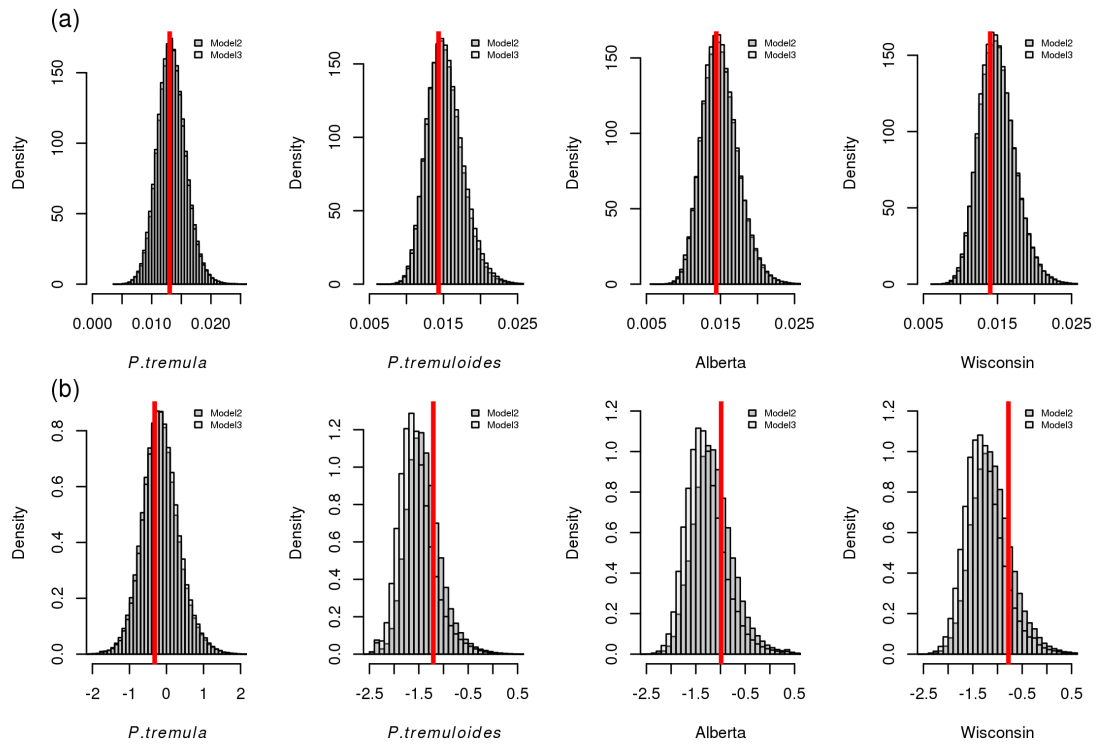
37

38 **Figure S7.** Predictive distribution of 100,000 simulated values of nucleotide diversity
 39 (θ_π) and Tajima's D under the best-fitting demographic model inferred by
 40 *fastsimcoal2* (histogram) (Model 1 as shown in Figure S5) in *P. tremula* (a,b) and *P.*
 41 *tremuloides* (c,d). Red horizontal bars represent the averages values obtained from the
 42 observed data. The length of the simulated region is 10 Kbp.

43

44

45



46

47 **Figure S8.** Predictive distribution of 100,000 simulated values of (a) nucleotide
 48 diversity (θ_{π}) and (b) Tajima's D, under the model 2 and model 3 as shown in Figure
 49 S5 (histogram), in *P. tremula*, pooled samples of *P. tremuloides*, subpopulations of
 50 Alberta and Wisconsin in *P. tremuloides*, respectively. Red horizontal bars represent
 51 the averages values obtained from the observed data. The length of the simulated
 52 region is 10 Kbp.

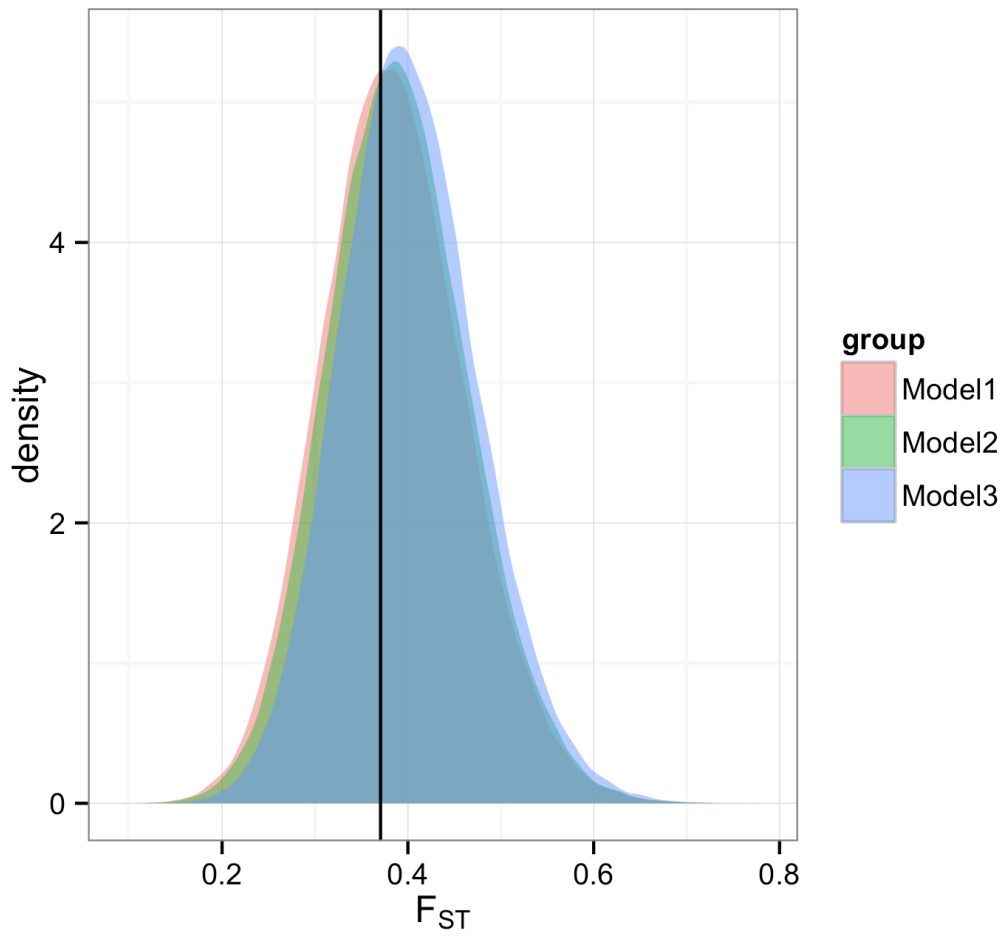
53

54

55

56

57



58

59

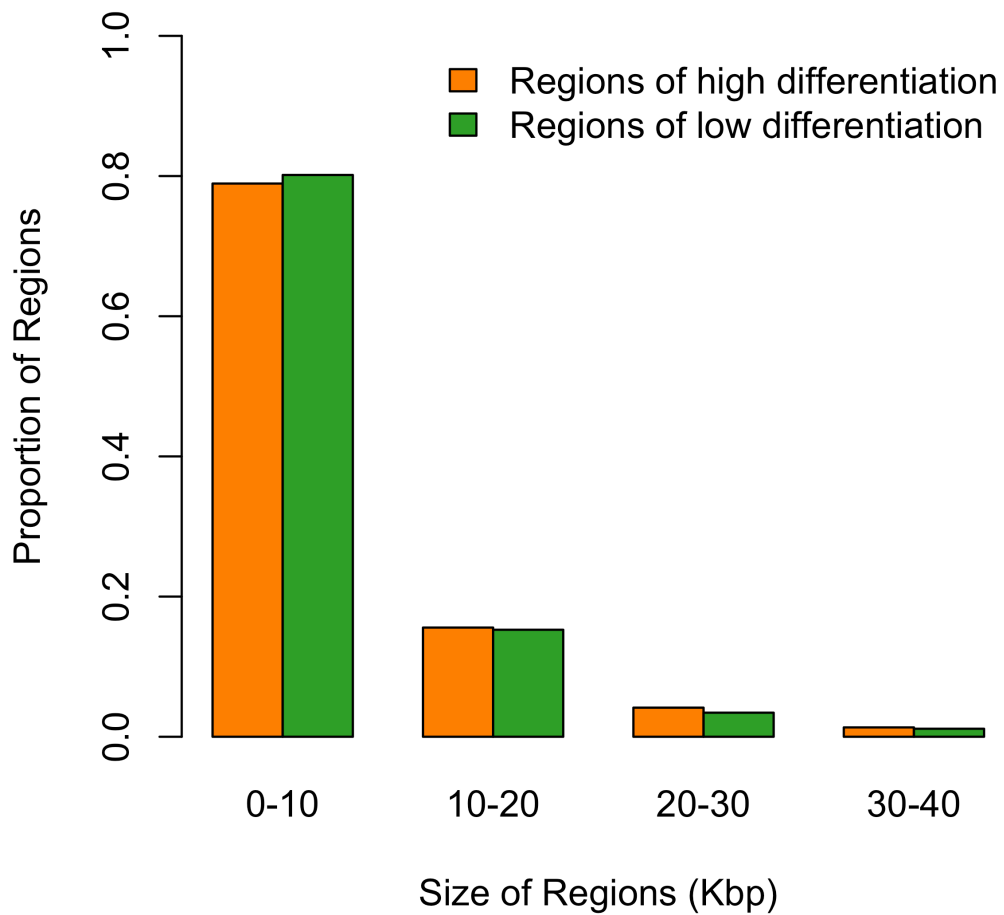
60 **Figure S9.** Predictive distributions of 100,000 simulated F_{ST} values between *P.*
 61 *tremula* and *P. tremuloides* under the model 1, 2, and 3 as shown in Figure S5. The
 62 length of the simulated region is 10 Kbp. The black horizontal bar represent the mean
 63 F_{ST} value obtained from the observed data.

64

65

66

67



68

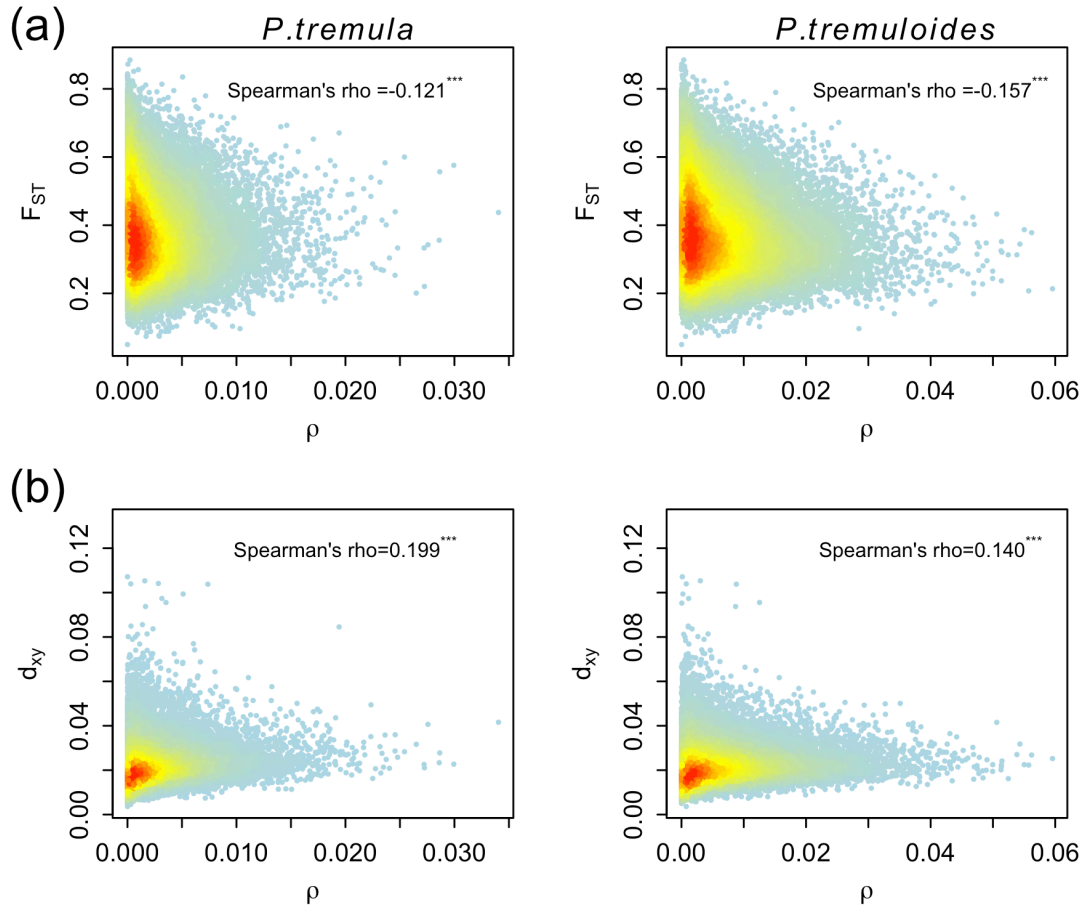
69 **Figure S10.** The physical size distributions (Kbp) of regions displaying significantly
 70 high (orange bars) and low (green bars) genetic differentiation between *P. tremula*
 71 and *P. tremuloides*. Here, all consecutive windows are considered as a single larger
 72 region.

73

74

75

76



77

78

79 **Figure S11.** Relationships between population-scaled recombination rates (ρ) and F_{ST}
 80 (a), d_{xy} (b) in both *P. tremula* (left panel) and *P. tremuloides* (right panel). Scatter
 81 plots display genome-wide values of two variables in dots over 10 Kbp non-
 82 overlapping windows. The red to yellow to blue gradient indicates decreased density
 83 of observed events at a give location in the graph.

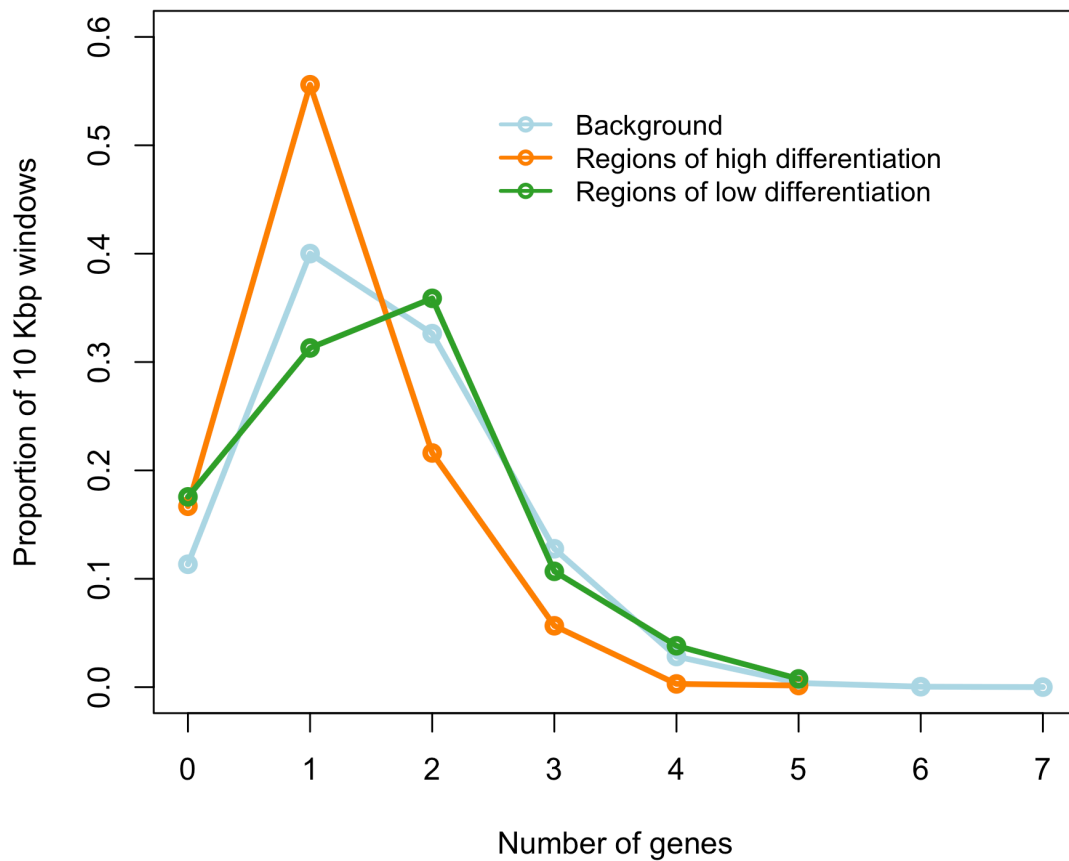
84

85

86

87

88



89

90 **Figure S12.** Comparison of gene density between regions displaying significantly
 91 high (orange line) and low (green line) genetic differentiation with the genomic
 92 background.

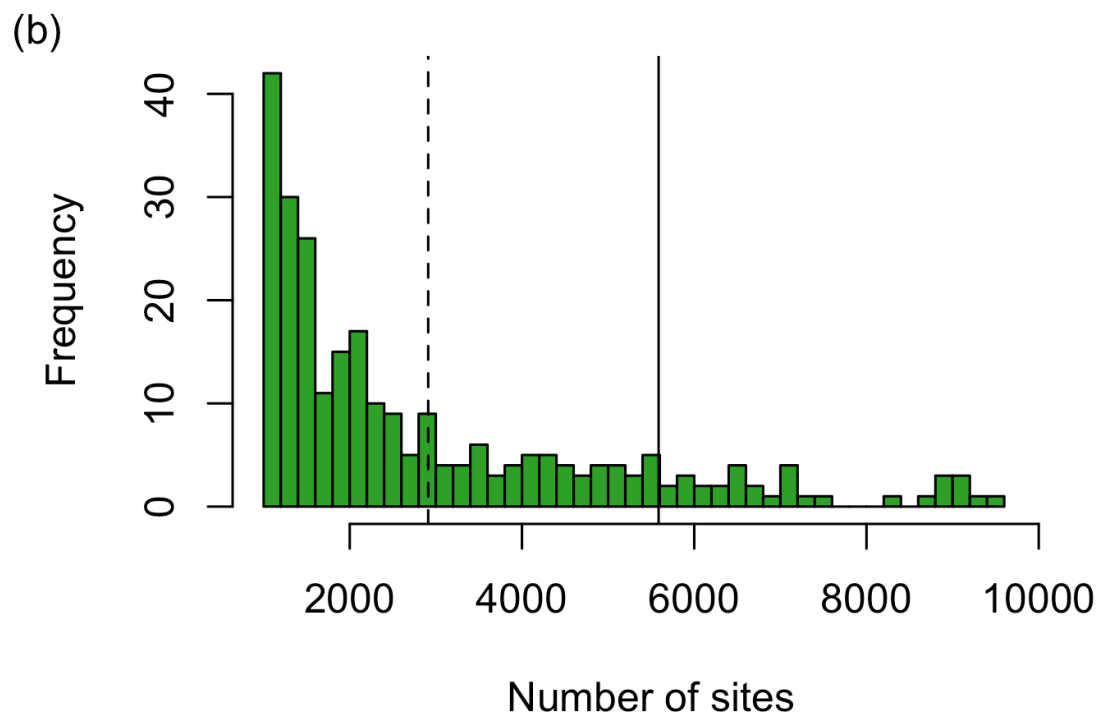
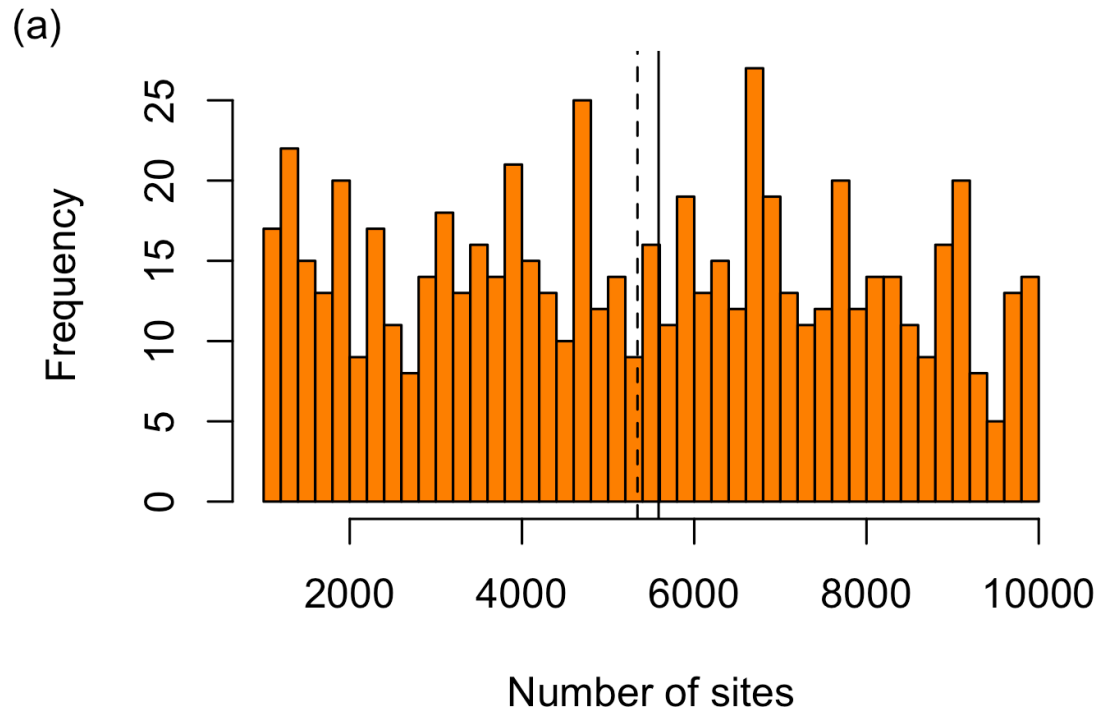
93

94

95

96

97



98

99 **Figure S13.** Frequency distribution of coverage breadth (the number of sites left after
 100 a number of filtering steps listed in Methods and Materials) in outlier windows
 101 displaying exceptionally high (a) and low (b) genetic differentiation between *P.*
 102 *tremula* and *P. tremuloides*, with mean values of these regions being represented by
 103 the vertical dashed line. Mean coverage breadth across the genome is represented by
 104 the vertical solid line in both (a) and (b).