Fig. S1 Cline estimates for all 77 markers used in this study and all three localities. **Fig. S2** Most likely number of genetic clusters *k* at the three localities. **Fig. S3** Locus specific ancestry (LSA) described by evidence for interspecific heterozygosity at the three localities for all potentially admixed plus eight of each of the parental-like individuals (*y*-axis), ordered by *Q* value as in Fig. 2, main article (at each locality: top, *P. alba* like; middle, admixed; bottom, *P. tremula* like), analyzed with 77 markers across 19 chromosomes in map position (*x*-axis). **Fig. S4** Locus specific ancestry (LSA) described by evidence for specific homozygosity at the three localities for all potentially admixed plus eight of each of the parental-like individuals (*y*-axis), ordered by *Q* value as in Fig. 2, main article (at each locality: top, *P. alba* like; middle, admixed; bottom, *P. tremula* like), analyzed with 77 markers across 19 chromosomes in map position (*x*-axis). **Fig. S5** Summary for 67 codominant loci for the Italy hybrid zone, including interspecific heterozygosity and excess ancestry in admixed individuals, and divergence- and diversity-based genome scans involving the two parental species *P. alba* and *P. tremula*. **Table S1** General information about 77 microsatellite markers used in this study, including published name, position on scaffold (in bp), distance used for structure linkage runs, marker type (codominant or dominant scoring), and primer sequences. **Table S2** Locus-specific measurements of genetic differentiation between *P. alba* and *P. tremula* for 67 codominant loci, including number of alleles (N_A), F_{ST}, Hedrick's G'_{ST}, and δ, for each of the three replicate localities. **Table S3** Locus-specific measurements of observed (H_O) and expected heterozygosity (H_E) for *P. alba* and *P. tremula* for 67 codominant loci at the three replicate localities. **Table S4** Results of outlier scans for 67 codominant markers at the three replicate localities. **Table S5** Fisher's exact tests (including *P*-value, odds ratio and 95% CI) and Phi coefficient φ for tests of association between loci identified as being under selection by genome scans involving the parental species using BayeScan, lositan, or Ewens-Watterson tests (EW; results for *P. alba* and *P. tremula* are pooled), and loci showing interspecific heterozygosities exceeding simulated 95% CIs in admixed individuals (see main article, Fig. 4, and Text S1 for more details on outlier scans) **Text S1** Genome scans for parental populations

SUPPORTING INFORMATION

Recombinant hybrids retain heterozygosity at many loci: new insights into the genomics of reproductive isolation in *Populus*

DOROTHEA LINDTKE[1], C. ALEX BUERKLE[2], THELMA BARBARÁ[1, 3], BERTHOLD HEINZE[4], STEFANO CASTIGLIONE[5], DENES BARTHA[6], and CHRISTIAN LEXER[1, 3]

[1]University of Fribourg, Department of Biology, Unit of Ecology and Evolution, 1700 Fribourg, Switzerland
[2]University of Wyoming, Department of Botany, Laramie, WY 82071, USA
[3]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3DS, UK
[4]Federal Research Centre for Forests, Department of Genetics, A-1140 Vienna, Austria
[5]University of Salerno, Department of Chemistry and Biology, 84084 Fisciano, Italy
[6]West-Hungarian University, Department of Botany, 9400 Sopron, Hungary

Table S1: General information about 77 microsatellite markers used in this study, including published name, position on scaffold (in bp), distance used for *structure* linkage runs, marker type (codominant or dominant scoring), and primer sequences. Microsatellite markers were obtained from http://www.ornl.gov/sci/ipgc/ssr_resource.htm; van der Schoot *et al.* (2000), Smulders *et al.* (2001), Tuskan *et al.* (2004); and De Carvalho *et al.* (2010).

Marker	Published name	*	Chromo-some	Position (bp)	**	*structure* distance ***	Type	Forward primer	Reverse primer
O137	ORPM_137		1	9087453	#	-1	codom	CCGTGCATCTGCTCACTTTA	GCATTTGCAGATGAAATTGGT
G124	GCPM_124		1	9131303	#	0.219250	codom	TTTGAGCACTTCAACTACCA	TGTCTTCCCTTAGTCACCAC
O30_1	ORPM_30	*	1	12600000	##	17.343485	codom	ATGTCCACACCCAGATGACA	CCGGCTTCATTAAGAGTTGG
ASP376	ASP112376		1	13305812	#	3.529060	codom	GCTCGAGATCTATCGGCAAC	TCCTCTCCAGGAAACTCCAA
P2852	PMGC_2852	*	1	14548286	#	6.212370	codom	ATAATCTCCCTAGCTTAATTCC	GAATAACATGGATAATGTGTTTG
ASP302	ASP113302		1	30444558	#	79.481360	codom	GGCAGCGTCTTCATTCCTAT	AAAGCCATTGAAGTGGAGGA
G1719	GCPM_1719		1	35488312	#	25.218770	codom	AAGTGCTCATAACATCACCC	CTTTCCTCATTCCTGTTCTG
G1274	GCPM_1274		1	45729682	#	51.206850	dom	GCCTGATACTTGTGGACCTA	CCCGTATAATATGATGATCCA
G1158	GCPM_1158		2	2787112	#	-1	codom	ATGCACTTCCTTCCAAATTA	ATCAGTTCCTTCAGCTTCAA
G1376	GCPM_1376		2	23223025	#	102.179565	codom	TGTCAAATAGTAGCATCCCC	CCACCTTGACTTTTCTTCTG
G1133	GCPM_1133		3	4716293	#	-1	codom	TCCGATCGAGAATAGAAAGA	GCTTGAATGGATCAGATGTT
O30_2	ORPM_30		3	10738184	#	30.109455	codom	ATGTCCACACCCAGATGACA	CCGGCTTCATTAAGAGTTGG
G1887	GCPM_1887		3	11095921	#	1.788685	dom	TTTTTCCTTTGATGCACATT	AAAAGCATTGTTCCTGTGAT
G1629	GCPM_1629		3	11310185	#	1.071320	codom	ACAATGAAGGATATTGGCAC	TTGAGCTAATTACCAGTCCC
O203	ORPM_203		3	12009782	#	3.497985	codom	CCACCAGGCATGAGATATGA	TCAAACCGAAAGGTCAACAA
G1869	GCPM_1869		3	16805774	#	23.979960	codom	TAAACCTAATTGATGCCTGC	TTGTAGAAGGTTTTTGCCAT
G1688	GCPM_1688		3	17574314	#	3.842700	codom	CCCCTAAAAAGCCAGTTTAT	AAGAAAAGAGTGCTCCTCCT
O127	ORPM_127	*	4	6447171	#	-1	codom	TCAATGAGGGGTGCCATAAT	CTTTCCACTTTTGGCCCTTT
O220	ORPM_220	*	4	7778968	##	6.658985	codom	AGCTAGCCTGTCGTCAAGGA	CAAGGAAGCATTCTCGCAAT
G1809	GCPM_1809		4	9110875	#	6.659535	codom	TACAAATGCTAATTACCCCC	AATTAGCCAATCACATCTGC
G1255	GCPM_1255		5	1728199	#	-1	codom	GAACCTTAAAACCAGAACCC	GAGCCACAGAAATACTGCTC
G1192	GCPM_1192		5	4083941	#	11.778710	codom	CATGCATCATTAGAGAAGAGG	TAATTGGTGAATCAAAGCCT
G1838	GCPM_1838		5	8802231	#	23.591450	codom	GTTCAGCGAAAGCTAAAGAG	CACAGAATTACAGCTGATGC
G20	GCPM_20		5	23171490	#	71.846295	codom	TTTTCATCTCTGCCAAGTCT	TATACCAAGGGACTATGCGT
W15	WPMS_15	*	5	25424594	#	11.265520	codom	CAACAAACCATCAATGAAGAAGAC	AGAGGGTGTTGGGGGTGACTA
G139	GCPM_139		6	2281003	#	-1	codom	ATGACATGACATGATTGGAA	CTTCTGCTGGAAGAAGAAAA
G1831	GCPM_1831		6	3671141	#	6.950690	codom	CACGTAAACAGCTTCCAAGT	CAGATGGAAAATACGGAGAC
G1074	GCPM_1074		6	3989388	#	1.591235	codom	TCAAGTGAAAAGATTGAAAGTG	TTGAAAAGCAAATCTGAGGT
O26	ORPM_26		6	5786927	#	8.987695	codom	GCTGCAGTCAAATTCCAAAA	CGAGCGTCTTCTTCATGGAT
O167	ORPM_167	*	6	5821040	#	0.170565	codom	TGCACTATTTACTCGCAGTCTCTC	AAGCTTTTCCGAAACCGAAG
ASP933	ASP106933		6	13019647	#	35.993035	codom	AGCCAGTAGGCTTCTTCAACA	CATCCCTCCCTAATCCCATT
O190	ORPM_190		6	13718036	#	3.491945	codom	CCCTGGTTTTCTCTTCTTGG	CCAGATTGGACTTGGGATTC
W12	WPMS_12		6	19471676	#	28.768200	codom	TTTTTCGTATTCTTATCTATCC	CACTACTCTGACAAAACCATC
G2034	GCPM_2034		6	22219025	#	13.736745	codom	ACAAACTGCTTTGTTTGGTT	CTCCATTCATAAAATCGAGC
O369	ORPM_369		6	22796697	#	2.888360	codom	TGTTCGGGTTATATTGCCATT	TGATTGGGTGTCTCTGCTTG
O60	ORPM_60		6	23655126	#	4.292145	codom	ATAGCGCCAGAAGCAAAAAC	AAGCAGAAAGTCGTAGGTTCG
G1065	GCPM_1065		6	24119014	#	2.319440	codom	TGCAATCATATATTCCTCCC	ATAAAATTACTGCGTGCCAT
ASP322	ASP112322		6	25184620	#	5.328030	codom	CATTAACGCCCCATTTCAGT	GTGAGGCACCACCCTGATAG

G1260	GCPM_1260		7	3438279	#	-1	codom	CACAGGAACCTGGTTATCAT	CTGGCATTCCTTCTAAGCTA
W17	WPMS_17		7	8696038	#	26.288795	codom	ACATCCGCCAATGCTTCGGTGTTT	GTGACGGTGGTGGCGGATTTTCTT
G1416	GCPM_1416		7	9247522	#	2.757420	codom	TCTTGAGAGGGCACTAGAAG	TGACAATTAGAATGGAACCC
G1295	GCPM_1295		7	11243952	#	9.982150	dom	TGATTAAACAGTGGCAGTGA	ATCCTCCTTCTACCCCTCTA
O312	ORPM_312	*	7	11625195	#	1.906215	codom	GTGGGGATCAATCCAAAAGA	CCCATATCAAACCATTTGAAAAA
G2062	GCPM_2062		8	5051212	#	-1	codom	TCTCTCCTGGTAAGTAAGTCTGT	CAGCATGTTTCTTCAGTCAA
O374	ORPM_374		8	6575467	#	7.621275	codom	TTTCAAAGAAAGGCTGCAGAA	TTAGCTAGGGTGCCGTATCG
O202	ORPM_202		8	13152189	#	32.883610	codom	TCGCAAAAGATTCTCCCAGT	TTCAAATCCCGGTAATGCTC
O268	ORPM_268		8	13427006	#	1.374085	codom	TTGCTGGGTACCCTATCTCA	AGCGTATTTGAAGCGATTTGA
G1949	GCPM_1949		9	1444490	#	-1	codom	TGGTTAGTACCAGCAAAACC	TTAAAGAGCCAGCCACTATC
O23	ORPM_23		9	4156696	#	13.561030	codom	ATTCCATTTGGCAATCAAGG	CCCTGAAAGTCACGTCTTCG
O21	ORPM_21		9	5179553	#	5.114285	codom	GGCTGCAGCACCAGAATAAT	TGCATCCAAAATTTTCCTCTTT
G1250	GCPM_1250		10	1067244	#[1]	-1	dom	GAAGACGAAGACGATAGCAG	AAGACGAGAGCAGAACAGAG
G2020	GCPM_2020		10	8741893	#	38.373245	codom	TAAAAATCCCCAAATTTCAA	CCAATAAATAGCTTTCCCTG
O344	ORPM_344		10	14738667	#	29.983870	codom	GGAGATTGTCGGAGAATGGA	TGGACGTTACGATAGGAGTGG
G1574	GCPM_1574		10	16527923	#	8.946280	codom	ATTGGTCTCGTACCCAAAA	ATTTCCAATGCATATGTTCC
O149	ORPM_149	*	10	16581540	#	0.268085	codom	GTCTCTGCCACATGATCCAA	CCCGAAATGGATCAAACAAG
G114	GCPM_114		10	20637529	#	20.279945	codom	TTAGCCATTGGATTTCATTT	CATTGCACTCTCACACATTC
G1037	GCPM_1037		11	5503115	#[1]	-1	codom	ATGAAATTCGCAAAGTCAGT	AAAAGAGGAAATTACGGTCC
G154	GCPM_154		12	8796449	#	-1	dom	CGTAGGAGCAAAAGAGATTG	TTTTGGAGACATTCCTTCAC
W05	WPMS_5		12	9208533	#	2.060420	codom	TTCTTTTTCAACTGCCTAACTT	TGATCCAATAACAGACAGAACA
G1186	GCPM_1186		12	13872383	#	23.319250	codom	TGTATTTGTGTGGGTTGAAA	AAGTAAGTGTTGGCTGCATT
O16	ORPM_16		13	282478	#[1]	-1	dom	GCAGAAACCACTGCTAGATGC	GCTTTGAGGAGGTGTGAGGA
G1353	GCPM_1353		13	817449	#	2.674855	codom	GAAAACTGATTCCTGATTCG	CAAGAATCAATGCATGTCTG
G1292	GCPM_1292		14	8054896	#	-1	dom	TTATTGCAGCTAGCCTCTTC	GCGTGTAATTTTTCTTCTGC
G1812	GCPM_1812		14	9862336	#	9.037200	codom	TGCTTCTCTATTTCTAGGCG	GCTGTTACTGTCTCTCCAGC
G2014	GCPM_2014		14	13260264	#	16.989640	dom	CAGGGAACTCTTTTCTTCCT	GGAAAGGGTAGTCACTCACA
G1894	GCPM_1894		15	809326	#	-1	codom	CTCTCGAACCATCAACTCTC	GACATGCACGCATAGAATTA
G1454	GCPM_1454		15	913028	#	0.518510	codom	ATTGCGCTGGTTGTAGTTAT	CATTTGAAAGAAGGGTTTTG
G1608	GCPM_1608		15	4979950	#	20.334610	codom	GCTCCTGGTTTTACCACAT	GAACAGCAGGATCATAGAGC
G430	ORPM_430		15	10362768	#	26.914090	dom	CCTTGGAAAAACCCCAAAAT	CAGCTCGACTCATTGCAAAA
O14	ORPM_14		16	1290594	#	-1	codom	GGGCTGCAGCAGATATTGA	CCAAAGGAACCCAAAGAAGA
G1381	GCPM_1381		17	6100168	#	-1	dom	CAATGTCAAGTGCTCAGAAA	GTATTGGGTGAAGGTTGAGA
O214	ORPM_214	*	18	4583970	#	-1	codom	TTTTCACAAGCCTCGAAGGA	TGGAAGACCCGAACTTTTTC
G1577	GCPM_1577		18	6074006	#	7.450180	codom	GAGAACATGTCAGCAGTTCA	GCTTAAACATTGAGAAAGCG
O28	ORPM_28	*	18	11993250	#	29.596220	codom	GGATCGACTTCCAACCCATA	AATTCCCAGATGAAGGCTCA
G162	GCPM_162		18	14881684	#	14.442170	codom	GCCCAAACTCTTATTTGATG	TGGTGGAGGCTAGGATAGTA
O276	ORPM_276		19	2850078	#[1]	-1	codom	GCAGGAGAAAACACCAGGAA	TCGCGAAAGAGAAGAAAAGC
O206	ORPM_206	*	19	3630075	#	3.899985	codom	CCGTGGCCATTGACTCTTTA	GAACCCATTTGGTGCAAGAT

* 11 loci used for inter-locality statistics in combined dataset

** Chromosomal positions were obtained by BLAST of the primer sequences against the *P. trichocarpa* genome assembly v2 (http://www.phytozome.net/poplar) (#), and from an interspecific cross *P. alba* x *P.* ×*canescens* (Macaya-Sanz, Alba, Gonzalez-Martinez, Lexer and coworkers, unpublished data) (##). In some cases the BLAST search resulted in several hits; then the most likely position according to motive type or fragment length, or the average position was used (#[1]).

*** *structure* distance gives inter-marker distances as needed for the *structure* input files. The physical distance in bp (computed from the column 'Position') was divided by 200,000 as a proxy for cM, assuming that 1cM corresponds roughly to 200kb (Tuskan *et al.* 2006, Supplementary information). The first marker on a chromosome is indicated by the value -1

Table S2: Locus-specific measurements of genetic differentiation between *P. alba* and *P. tremula* for 67 codominant loci, including number of alleles (N_A), F_{ST}, Hedrick's G'_{ST}, and δ, for each of the three replicate localities.

Marker	*	Chromo-some	N_A			F_{ST}			Hedrick's G'_{ST}			δ		
			Italy	Austria	Hungary	Italy	Austria	Hungary	Italy	Austria	Hungary	Italy	Austria	Hungary
O137		1	10	8	11	0.72	0.55	0.36	1.00	0.98	0.97	0.98	0.96	0.90
G124		1	6	12	7	0.39	0.33	0.29	0.85	0.94	0.88	0.78	0.90	0.89
O30_1	*	1	3	4	6	0.02	0.15	0.12	0.02	0.12	0.10	0.05	0.25	0.26
ASP376		1	15	16	16	0.37	0.44	0.39	0.96	0.99	1.00	0.94	0.96	1.00
P2852	*	1	22	25	25	0.19	0.05	0.07	0.75	0.30	0.39	0.79	0.51	0.56
ASP302		1	9	12	13	0.10	0.09	0.13	0.49	0.45	0.63	0.57	0.52	0.64
G1719		1	7	13	11	0.11	0.16	0.15	0.42	0.65	0.56	0.56	0.60	0.63
G1158		2	7	8	6	0.86	0.93	0.88	1.00	1.00	1.00	1.00	1.00	0.98
G1376		2	17	16	5	0.54	0.81	0.85	1.00	1.00	1.00	1.00	0.99	1.00
G1133		3	12	19	8	0.24	0.25	0.28	0.89	0.80	0.64	0.81	0.77	0.57
O30_2		3	24	26	26	0.16	0.08	0.06	0.93	0.61	0.44	0.88	0.64	0.54
G1629		3	17	29	20	0.15	0.16	0.18	0.73	0.66	0.91	0.81	0.73	0.95
O203		3	6	10	6	0.15	0.23	0.23	0.26	0.25	0.33	0.43	0.44	0.51
G1869		3	20	24	21	0.20	0.17	0.20	1.00	1.00	1.00	1.00	1.00	1.00
G1688		3	6	16	14	0.68	0.35	0.35	0.86	0.73	0.64	0.80	0.68	0.65
O127	*	4	7	8	7	0.43	0.55	0.45	0.78	0.53	0.52	0.78	0.59	0.58
O220	*	4	6	5	5	0.27	0.84	0.40	0.35	0.99	0.79	0.43	0.97	0.70
G1809		4	7	9	8	0.30	0.60	0.45	0.54	0.70	0.63	0.63	0.75	0.69
G1255		5	4	3	2	0.90	0.99	1.00	1.00	1.00	1.00	1.00	0.99	1.00
G1192		5	6	13	6	0.21	0.11	0.16	0.34	0.21	0.38	0.51	0.44	0.53
G1838		5	20	16	15	0.16	0.20	0.13	0.78	0.80	0.66	0.81	0.82	0.69
G20		5	5	8	6	0.63	0.39	0.20	0.75	0.61	0.27	0.68	0.59	0.41
W15	*	5	10	9	10	0.09	0.16	0.11	0.48	0.57	0.44	0.52	0.59	0.53
G139		6	8	8	11	0.60	0.77	0.67	1.00	1.00	1.00	1.00	0.99	1.00
G1831		6	11	14	13	0.25	0.17	0.08	0.77	0.72	0.46	0.72	0.69	0.55
G1074		6	2	2	2	0.97	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00
O26		6	6	11	9	0.50	0.44	0.45	0.65	0.95	0.82	0.71	0.95	0.84
O167	*	6	4	4	5	0.92	0.90	0.86	0.99	1.00	1.00	0.96	0.99	0.99
ASP933		6	10	17	14	0.48	0.52	0.54	1.00	1.00	1.00	1.00	0.99	0.99
O190		6	3	4	3	0.79	0.68	0.72	0.98	0.92	0.96	0.95	0.86	0.91
W12		6	9	28	21	0.36	0.14	0.17	0.98	0.80	0.94	0.92	0.79	0.93
G2034		6	12	24	17	0.46	0.33	0.41	1.00	1.00	1.00	0.96	0.97	0.96
O369		6	5	8	6	0.11	0.04	0.04	0.17	0.08	0.08	0.33	0.24	0.24
O60		6	10	16	12	0.14	0.20	0.31	0.50	0.53	0.67	0.65	0.63	0.76
G1065		6	7	11	7	0.47	0.76	0.36	0.81	1.00	0.51	0.82	0.99	0.61
ASP322		6	15	17	18	0.29	0.19	0.16	1.00	0.93	0.89	1.00	0.91	0.86

G1260		7	7	8	8	0.14	0.15	0.24	0.20	0.29	0.55	0.35	0.38	0.58
W17		7	10	14	9	0.20	0.21	0.28	0.44	0.56	0.72	0.53	0.56	0.67
G1416		7	3	7	3	0.02	0.06	0.12	0.02	0.07	0.13	0.06	0.21	0.21
O312	*	7	9	12	11	0.32	0.16	0.15	0.95	0.64	0.57	0.89	0.65	0.59
G2062		8	19	23	19	0.18	0.08	0.16	0.94	0.55	0.82	0.86	0.59	0.76
O374		8	36	35	37	0.09	0.09	0.08	0.99	0.89	0.95	0.97	0.85	0.93
O202		8	5	5	6	0.24	0.32	0.33	0.52	0.68	0.75	0.51	0.63	0.66
O268		8	7	9	5	0.40	0.22	0.22	0.83	0.44	0.42	0.83	0.55	0.50
G1949		9	6	17	11	0.06	0.11	0.14	0.22	0.28	0.37	0.34	0.46	0.54
O23		9	23	36	26	0.20	0.22	0.21	1.00	1.00	0.99	0.98	1.00	0.95
O21		9	2	3	2	0.29	0.06	0.04	0.20	0.07	0.04	0.24	0.11	0.09
G2020		10	17	25	26	0.14	0.18	0.10	0.76	0.84	0.61	0.77	0.83	0.71
O344		10	9	8	9	0.50	0.71	0.61	0.93	1.00	0.99	0.93	0.98	0.96
G1574		10	8	15	14	0.62	0.40	0.25	1.00	0.88	0.74	0.98	0.78	0.66
O149	*	10	7	7	10	0.69	0.43	0.34	1.00	0.98	0.95	1.00	0.97	0.90
G114		10	11	9	8	0.36	0.70	0.71	0.88	0.97	0.96	0.86	0.93	0.93
G1037		11	3	5	3	0.10	0.00	0.03	0.13	0.00	0.04	0.25	0.03	0.14
W05		12	13	18	26	0.32	0.23	0.11	1.00	0.94	0.77	1.00	0.94	0.77
G1186		12	11	12	10	0.38	0.35	0.39	0.99	0.99	1.00	0.95	0.92	0.99
G1353		13	4	8	6	0.05	0.07	0.08	0.09	0.15	0.18	0.21	0.26	0.28
G1812		14	7	14	12	0.29	0.18	0.20	0.90	0.71	0.76	0.85	0.75	0.77
G1894		15	6	12	9	0.65	0.44	0.46	0.99	0.93	0.91	0.97	0.91	0.89
G1454		15	14	24	23	0.28	0.15	0.13	0.87	0.73	0.75	0.86	0.79	0.81
G1608		15	10	15	10	0.54	0.51	0.49	1.00	1.00	0.98	1.00	1.00	0.98
O14		16	2	2	2	0.00	0.15	0.15	0.01	0.05	0.08	0.02	0.08	0.13
O214	*	18	3	4	5	0.76	0.80	0.70	0.92	0.97	0.90	0.83	0.91	0.82
G1577		18	8	17	11	0.63	0.47	0.53	1.00	1.00	1.00	0.97	1.00	1.00
O28	*	18	6	6	7	0.31	0.11	0.11	0.39	0.14	0.14	0.49	0.29	0.30
G162		18	20	28	27	0.18	0.16	0.14	0.87	0.92	0.81	0.85	0.91	0.82
O276		19	6	10	7	0.07	0.01	0.01	0.07	0.02	0.02	0.17	0.13	0.12
O206	*	19	7	8	12	0.00	0.69	0.55	0.02	0.67	0.67	0.09	0.61	0.64

* 11 loci used for inter-locality statistics in combined dataset

Table S3: Locus-specific measurements of observed (H_O) and expected heterozygosity (H_E) for *P. alba* and *P. tremula* for 67 codominant loci at the three replicate localities.

Marker	*	Chromo-some	Italy *P. tremula*		Italy *P. alba*		Austria *P. tremula*		Austria *P. alba*		Hungary *P. tremula*		Hungary *P. alba*	
			H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E	H_O	H_E
O137		1	0.23	0.32	0.23	0.22	0.21	0.43	0.28	0.45	0.31	0.67	0.47	0.60
G124		1	0.71	0.77	0.32	0.28	0.68	0.72	0.57	0.60	0.76	0.72	0.60	0.65
O30_1	*	1	0.10	0.10	0.03	0.03	0.30	0.41	0.16	0.16	0.34	0.42	0.14	0.15
ASP376		1	0.53	0.68	0.48	0.53	0.56	0.80	0.37	0.42	0.43	0.69	0.46	0.55
P2852	*	1	0.69	0.90	0.33	0.59	0.80	0.86	0.68	0.78	0.82	0.84	0.74	0.81
ASP302		1	0.83	0.82	0.62	0.77	0.76	0.78	0.83	0.83	0.78	0.78	0.89	0.82
G1719		1	0.64	0.67	0.82	0.83	0.56	0.63	0.85	0.86	0.63	0.65	0.89	0.83
G1158		2	0.22	0.23	0.00	0.00	0.14	0.13	0.04	0.05	0.07	0.07	0.11	0.14
G1376		2	0.52	0.81	0.00	0.00	0.36	0.64	0.02	0.02	0.05	0.39	0.01	0.01
G1133		3	0.11	0.77	0.21	0.67	0.11	0.72	0.23	0.68	0.22	0.69	0.07	0.49
O30_2		3	0.75	0.73	0.91	0.93	0.87	0.80	0.86	0.91	0.79	0.83	0.85	0.89
G1629		3	0.26	0.79	0.51	0.80	0.38	0.79	0.51	0.73	0.20	0.76	0.33	0.83
O203		3	0.54	0.65	0.20	0.32	0.55	0.63	0.14	0.21	0.60	0.67	0.14	0.25
G1869		3	0.82	0.88	0.26	0.71	0.74	0.85	0.54	0.80	0.72	0.84	0.68	0.77
G1688		3	0.03	0.03	0.46	0.58	0.24	0.30	0.64	0.68	0.17	0.26	0.52	0.68
O127	*	4	0.30	0.74	0.02	0.12	0.38	0.71	0.04	0.05	0.37	0.72	0.03	0.03
O220	*	4	0.58	0.58	0.00	0.00	0.50	0.50	0.02	0.02	0.62	0.60	0.09	0.44
G1809		4	0.68	0.80	0.07	0.07	0.71	0.84	0.04	0.04	0.77	0.83	0.02	0.05
G1255		5	0.01	0.01	0.27	0.24	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00
G1192		5	0.00	0.34	0.55	0.63	0.17	0.37	0.64	0.64	0.00	0.47	0.73	0.68
G1838		5	0.80	0.92	0.69	0.67	0.71	0.80	0.65	0.71	0.90	0.82	0.81	0.79
G20		5	0.10	0.10	0.31	0.48	0.04	0.04	0.54	0.58	0.04	0.04	0.56	0.61
W15	*	5	0.64	0.79	0.83	0.84	0.65	0.67	0.74	0.78	0.80	0.78	0.70	0.74
G139		6	0.29	0.69	0.00	0.00	0.40	0.78	0.01	0.01	0.42	0.77	0.02	0.02
G1831		6	0.83	0.76	0.59	0.61	0.76	0.75	0.74	0.78	0.88	0.80	0.67	0.84
G1074		6	0.04	0.04	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
O26		6	0.00	0.00	0.93	0.82	0.00	0.00	0.90	0.87	0.00	0.00	0.80	0.83
O167	*	6	0.07	0.07	0.08	0.07	0.00	0.00	0.15	0.14	0.02	0.02	0.22	0.23
ASP933		6	0.56	0.65	0.37	0.37	0.53	0.53	0.47	0.45	0.56	0.52	0.47	0.42
O190		6	0.28	0.28	0.10	0.09	0.41	0.39	0.28	0.24	0.20	0.39	0.14	0.17
W12		6	0.31	0.76	0.14	0.46	0.43	0.83	0.43	0.82	0.34	0.78	0.54	0.85
G2034		6	0.23	0.31	0.93	0.82	0.30	0.32	0.89	0.91	0.15	0.23	0.93	0.91
O369		6	0.49	0.58	0.25	0.27	0.54	0.55	0.54	0.54	0.58	0.60	0.50	0.48
O60		6	0.18	0.65	0.80	0.81	0.17	0.47	0.64	0.75	0.18	0.31	0.70	0.76
G1065		6	0.67	0.78	0.00	0.00	0.62	0.74	0.02	0.04	0.80	0.75	0.17	0.16
ASP322		6	0.85	0.83	0.45	0.56	0.78	0.81	0.78	0.79	0.82	0.79	0.86	0.85

G1260		7	0.51	0.48	0.39	0.34	0.42	0.49	0.55	0.60	0.43	0.43	0.71	0.72
W17		7	0.84	0.76	0.34	0.38	0.53	0.59	0.67	0.68	0.53	0.58	0.57	0.66
G1416		7	0.03	0.06	0.12	0.12	0.09	0.40	0.19	0.37	0.04	0.04	0.36	0.37
O312	*	7	0.65	0.65	0.74	0.68	0.74	0.75	0.73	0.78	0.66	0.74	0.75	0.77
G2062		8	0.87	0.84	0.73	0.77	0.81	0.85	0.84	0.85	0.73	0.83	0.76	0.78
O374		8	0.69	0.89	0.64	0.92	0.61	0.89	0.63	0.92	0.64	0.91	0.72	0.92
O202		8	0.52	0.54	0.53	0.63	0.52	0.55	0.59	0.56	0.63	0.56	0.59	0.60
O268		8	0.22	0.48	0.59	0.60	0.16	0.34	0.64	0.65	0.21	0.48	0.44	0.56
G1949		9	0.38	0.73	0.38	0.66	0.49	0.79	0.46	0.53	0.56	0.78	0.53	0.53
O23		9	0.30	0.81	0.57	0.77	0.35	0.68	0.81	0.86	0.40	0.71	0.80	0.84
O21		9	0.00	0.00	0.24	0.37	0.00	0.00	0.21	0.19	0.00	0.07	0.16	0.22
G2020		10	0.82	0.82	0.78	0.81	0.65	0.67	0.84	0.89	0.78	0.75	0.89	0.90
O344		10	0.53	0.82	0.02	0.02	0.52	0.58	0.15	0.16	0.55	0.66	0.17	0.18
G1574		10	0.54	0.64	0.05	0.05	0.65	0.77	0.31	0.44	0.62	0.81	0.57	0.57
O149	*	10	0.12	0.17	0.18	0.51	0.54	0.52	0.42	0.58	0.39	0.63	0.30	0.65
G114		10	0.56	0.78	0.36	0.38	0.39	0.74	0.11	0.11	0.32	0.55	0.09	0.09
G1037		11	0.31	0.27	0.62	0.56	0.41	0.41	0.55	0.42	0.30	0.32	0.57	0.48
W05		12	0.30	0.73	0.46	0.62	0.47	0.80	0.69	0.73	0.53	0.87	0.76	0.84
G1186		12	0.62	0.55	0.48	0.70	0.44	0.48	0.59	0.75	0.33	0.46	0.56	0.74
G1353		13	0.41	0.52	0.52	0.49	0.38	0.64	0.51	0.53	0.49	0.64	0.44	0.53
G1812		14	0.61	0.64	0.82	0.73	0.70	0.64	0.67	0.83	0.70	0.68	0.71	0.79
G1894		15	0.06	0.06	0.62	0.74	0.06	0.06	0.87	0.81	0.11	0.10	0.78	0.81
G1454		15	0.59	0.65	0.74	0.72	0.65	0.77	0.80	0.82	0.82	0.80	0.86	0.86
G1608		15	0.22	0.26	0.51	0.73	0.11	0.23	0.46	0.64	0.10	0.32	0.50	0.63
O14		16	0.01	0.04	0.00	0.00	0.13	0.15	0.00	0.00	0.21	0.22	0.00	0.00
O214	*	18	0.12	0.11	0.31	0.33	0.18	0.16	0.19	0.20	0.30	0.31	0.19	0.21
G1577		18	0.09	0.09	0.69	0.75	0.04	0.06	0.82	0.83	0.02	0.02	0.76	0.83
O28	*	18	0.25	0.23	0.41	0.56	0.26	0.24	0.41	0.44	0.29	0.32	0.45	0.44
G162		18	0.76	0.82	0.86	0.78	0.78	0.85	0.75	0.80	0.89	0.87	0.84	0.79
O276		19	0.42	0.37	0.07	0.07	0.19	0.21	0.13	0.28	0.35	0.34	0.19	0.32
O206	*	19	0.40	0.59	0.40	0.58	0.40	0.55	0.03	0.03	0.55	0.60	0.04	0.06

* 11 loci used for inter-locality statistics in combined dataset

Table S4: Results of outlier scans for 67 codominant markers at the three replicate localities. Het: Interspecific heterozygosity. Loci with values exceeding simulated 95% CIs are indicated (+, excess of heterozygotes; -, deficiency of heterozygotes; see Fig. 4, main article). Hom: Specific homozygosity (excess ancestry). Loci with values exceeding simulated 95% CIs are indicated (+, excess of *P. alba* alleles; -, excess of *P. tremula* alleles; see Fig. 5, main article). BS: *BayeScan* results. Losi: *lositan* results. EWalba: Ewens-Watterson test for *P. alba*. EWtrem: Ewens-Watterson test for *P. tremula*. Loci identified as under balancing (bal) or directional selection (dir) are indicated. For the Ewens-Watterson tests, loci identified after controlling for the false discovery rate (FDR) at FDR = 0.1 are marked with bal+ or dir+. See Text S1 for more details on outlier scans.

Marker	*	Chromo-some	Italy						Austria						Hungary					
			Het	Hom	BS	Losi	EWalba	EWtrem	Het	Hom	BS	Losi	EWalba	EWtrem	Het	Hom	BS	Losi	EWalba	EWtrem
O137		1	+			dir		dir		+					+					
G124		1	+					bal							+					
O30_1	*	1					dir+													
ASP376		1	+												+					
P2852	*	1			bal						bal	bal					bal	bal		
ASP302		1				bal	bal+	bal				bal	bal		+					
G1719		1				bal	bal+						bal						bal	
G1158		2	+					dir			dir		dir		+			dir		
G1376		2	+						-			dir	dir	dir			dir		dir+	
G1133		3																		
O30_2		3	+		bal						bal	bal					bal	bal		bal
G1629		3									bal									
O203		3																		
G1869		3	+						+						+					
G1688		3						dir+												dir
O127	*	4	+										dir						dir	
O220	*	4							+						+					
G1809		4	-	-			dir	bal	-	-			dir	bal	-	-				bal
G1255		5	+		dir			dir+	+		dir	dir	dir		+		dir	dir		
G1192		5												dir						
G1838		5	+		bal			bal+	+											
G20		5	+																	
W15	*	5			bal	bal	bal+													
G139		6		+									dir+			+			dir+	
G1831		6	+															bal	bal	bal
G1074		6	+		dir	dir			+		dir	dir	dir+		+		dir	dir		
O26		6					bal+						bal						bal	
O167	*	6	-	-		dir			+		dir	dir				-	dir	dir		dir+
ASP933		6	+												+					

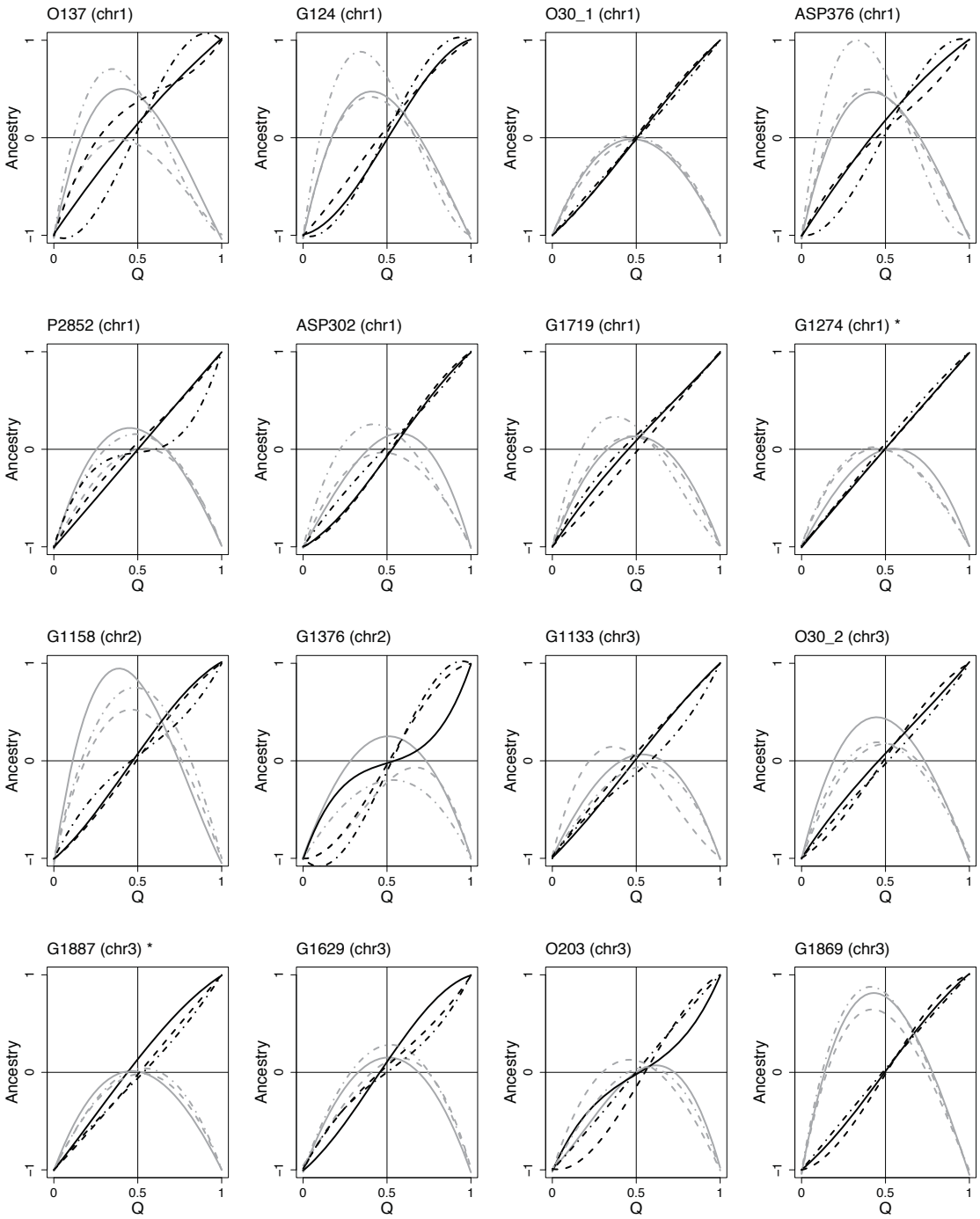
O190		6																													
W12		6									bal					bal									bal					
G2034		6		+					bal									bal				+					bal+			
O369		6																													
O60		6							bal																						
G1065		6		-		-						-		-				dir				-		-							
ASP322		6		+							bal	+										+								bal	
G1260		7																													
W17		7																													
G1416		7																													
O312	*	7		+																											
G2062		8					bal				bal					bal		bal							bal						
O374		8			-		bal		bal							bal									bal		bal				
O202		8							bal																						
O268		8			+																										
G1949		9			-			bal	bal		bal			-																	
O23		9		+			bal									bal						+			bal						
O21		9		-								-																			
G2020		10		+			bal									bal						+			bal	bal	bal				
O344		10		+					dir+		bal						dir	dir				+									
G1574		10		+				dir	dir+													+									
O149	*	10			-																	+									
G114		10		+																		+									
G1037		11										-					bal														
W05		12		+																					bal	bal					
G1186		12		+								+										+									
G1353		13			+							-																			
G1812		14		+					bal+																						
G1894		15		+					bal			+							dir			+									
G1454		15		+									+									+					bal				
G1608		15		-	+							-	+										+								
O14		16			+							-																			
O214	*	18		+	+							+										+									
G1577		18		+															dir			+					bal		dir+		
O28	*	18																													
G162		18		+			bal		bal+							bal						+			bal						
O276		19							dir			-						dir													
O206	*	19						bal										dir+									dir+				

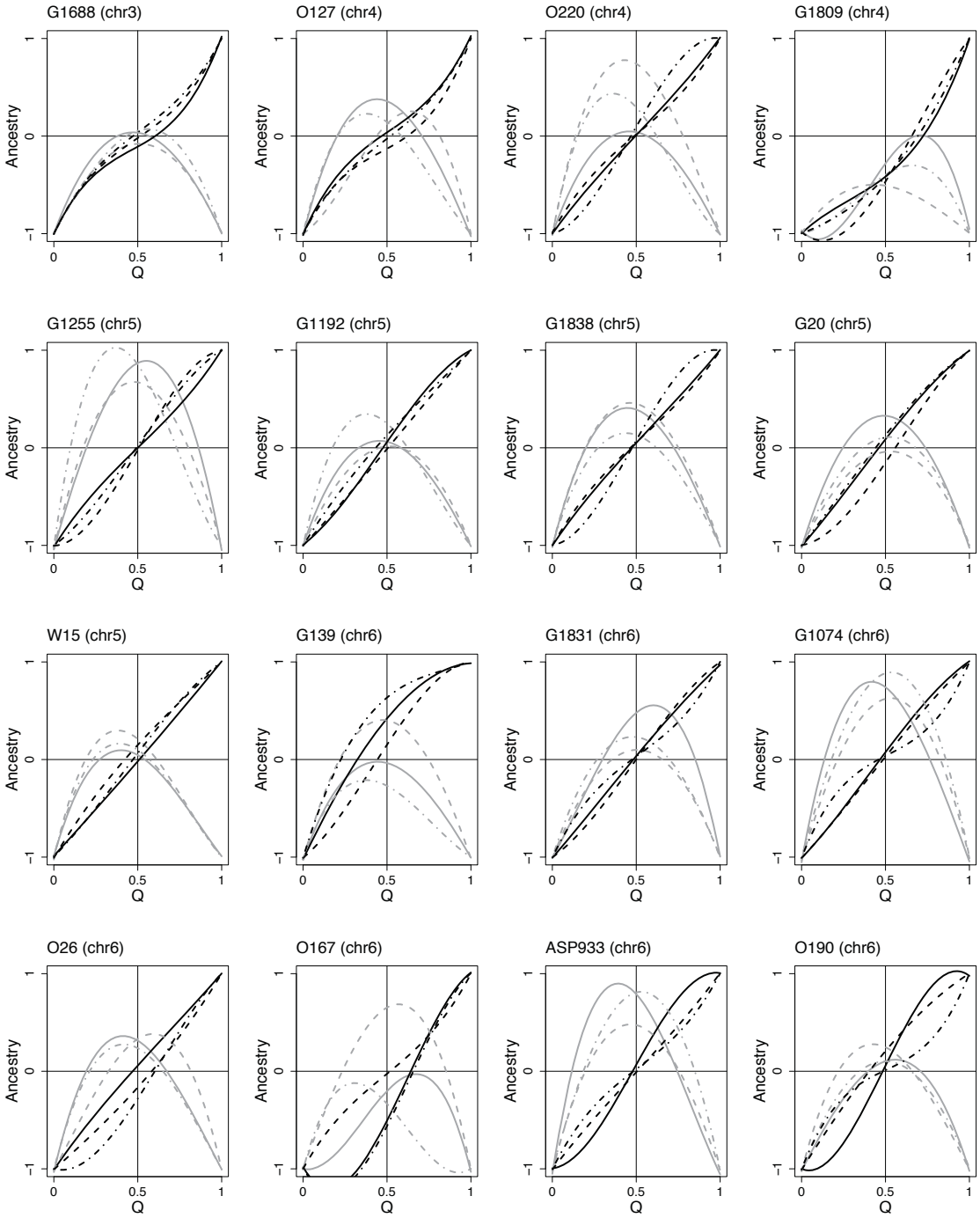
* 11 loci used for inter-locality statistics in combined dataset

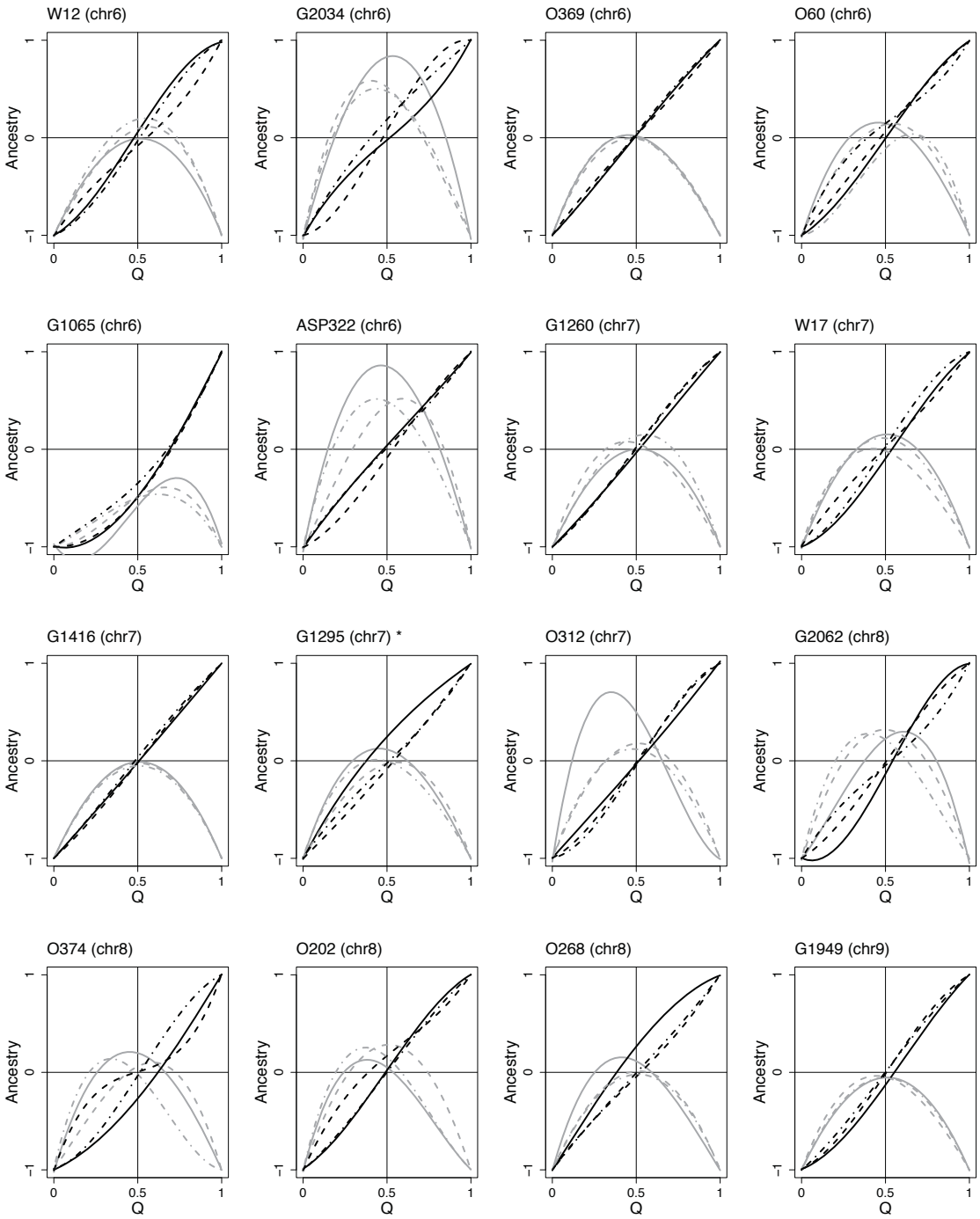
Table S5: Fisher's exact tests (including *p*-value, odds ratio and 95% CI) and Phi coefficient φ for tests of association between loci identified as being under selection by genome scans involving the parental species using *BayeScan*, *lositan*, or Ewens-Watterson tests (EW; results for *P. alba* and *P. tremula* are pooled), and loci showing interspecific heterozygosities exceeding simulated 95% CIs in admixed individuals (see main article, Fig. 4, and Text S1 for more details on outlier scans) *.

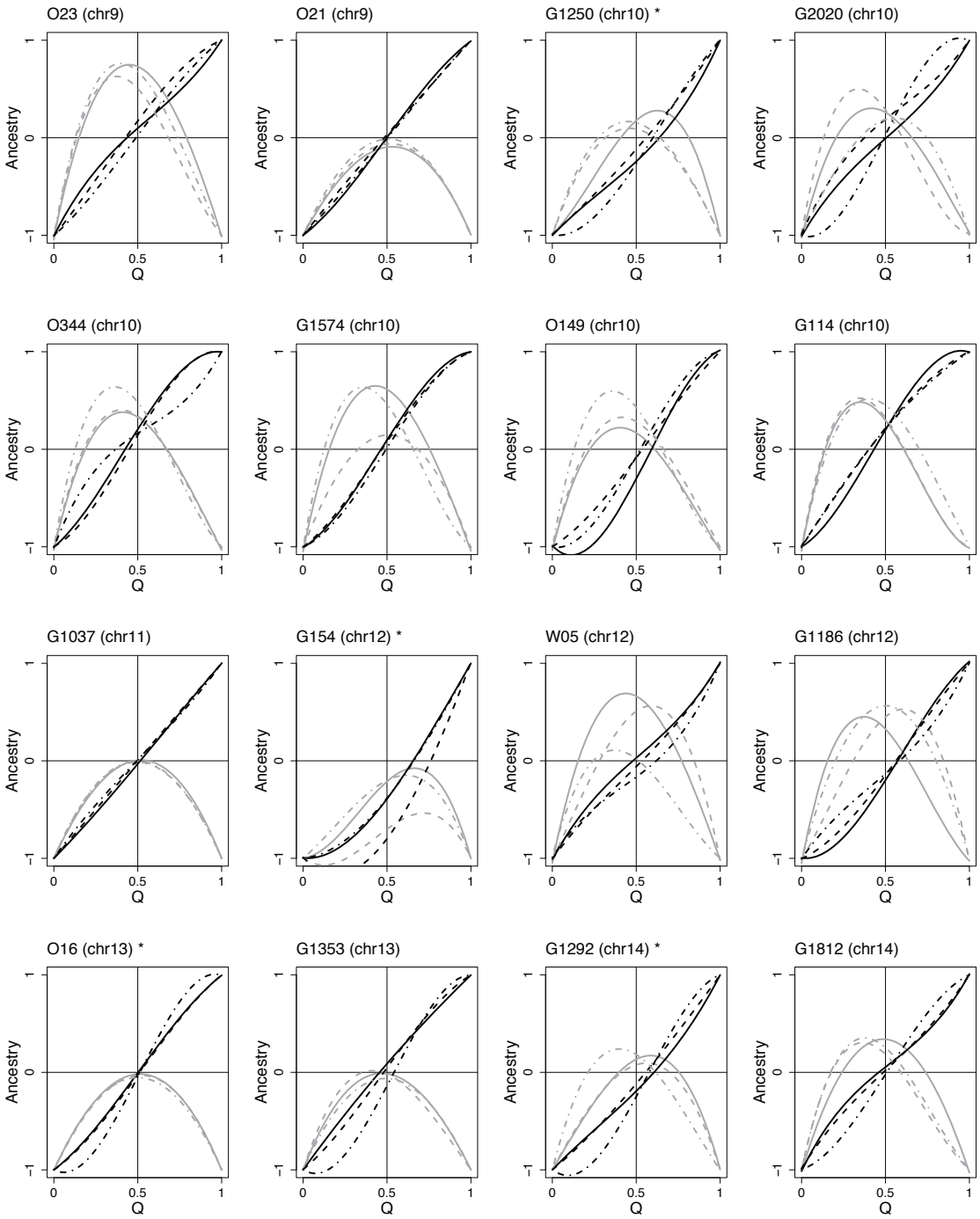
		Association (directional selection, deficiency of interspecific heterozygotes)					Association (balancing selection, excess of interspecific heterozygotes)				
Locality	Test	*p*-value	Odds ratio	95% CI		φ	*p*-value	Odds ratio	95% CI		φ
Italy	*BayeScan*	1.00	0.00	0.00	71.70	-0.05	0.50	1.64	0.32	9.16	0.09
	lositan	0.27	4.72	0.08	78.16	0.17	0.06	0.00	0.00	1.28	-0.26
	EW	0.53	1.67	0.03	20.07	0.05	1.00	0.86	0.25	2.85	-0.03
Austria	*BayeScan*	1.00	0.00	0.00	9.15	-0.11	0.34	0.00	0.00	2.93	-0.16
	lositan	0.57	1.46	0.03	17.25	0.04	1.00	0.00	0.00	6.58	-0.12
	EW	0.20	3.07	0.54	16.00	0.20	1.00	0.00	0.00	6.58	-0.12
Hungary	*BayeScan*	1.00	0.00	0.00	93.37	-0.04	0.71	0.51	0.05	3.01	-0.10
	lositan	1.00	0.00	0.00	93.37	-0.04	0.16	0.00	0.00	2.04	-0.21
	EW	1.00	0.00	0.00	48.17	-0.06	1.00	1.11	0.21	5.05	0.02

*) Fisher's exact tests and Phi coefficients were computed from binary coded, locus-specific variables (outlier/non-outlier), separately for the two specified directions of deviation.


O137 (chr1)
G124 (chr1)
O30_1 (chr1)
ASP376 (chr1)
P2852 (chr1)
ASP302 (chr1)
G1719 (chr1)
G1274 (chr1) *
G1158 (chr2)
G1376 (chr2)
G1133 (chr3)
O30_2 (chr3)
G1887 (chr3) *
G1629 (chr3)
O203 (chr3)
G1869 (chr3)
Ancestry
Q
−1
0
1
0.5


G1688 (chr3)
O127 (chr4)
O220 (chr4)
G1809 (chr4)
G1255 (chr5)
G1192 (chr5)
G1838 (chr5)
G20 (chr5)
W15 (chr5)
G139 (chr6)
G1831 (chr6)
G1074 (chr6)
O26 (chr6)
O167 (chr6)
ASP933 (chr6)
O190 (chr6)
Ancestry
Q
−1
0
1
0.5


W12 (chr6)
G2034 (chr6)
O369 (chr6)
O60 (chr6)
G1065 (chr6)
ASP322 (chr6)
G1260 (chr7)
W17 (chr7)
G1416 (chr7)
G1295 (chr7) *
O312 (chr7)
G2062 (chr8)
O374 (chr8)
O202 (chr8)
O268 (chr8)
G1949 (chr9)
Ancestry
Q
−1
0
1
0.5


O23 (chr9)
O21 (chr9)
G1250 (chr10) *
G2020 (chr10)
O344 (chr10)
G1574 (chr10)
O149 (chr10)
G114 (chr10)
G1037 (chr11)
G154 (chr12) *
W05 (chr12)
G1186 (chr12)
O16 (chr13) *
G1353 (chr13)
G1292 (chr14) *
G1812 (chr14)
Ancestry
Q
−1
0
1
0.5

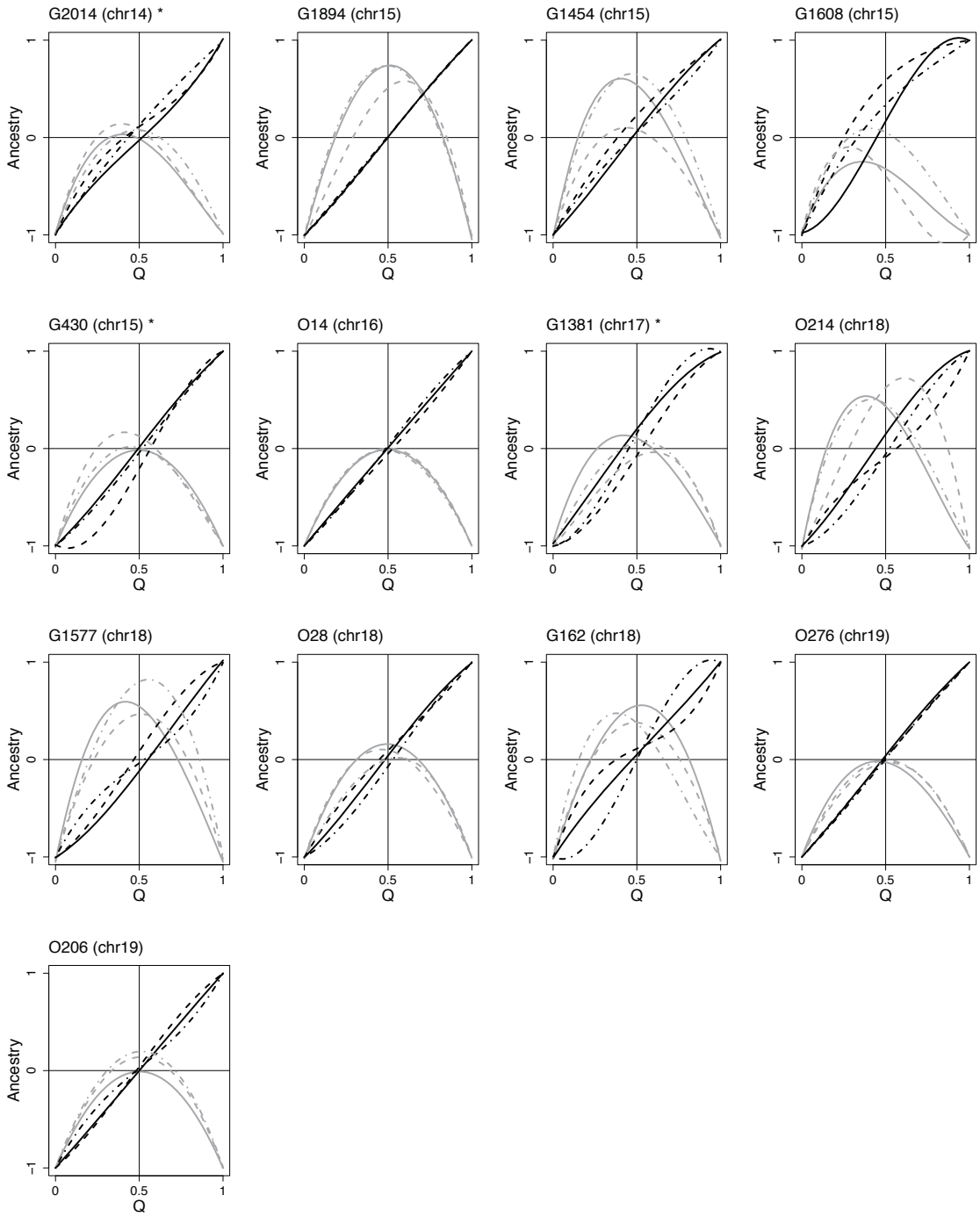


Figure S1: Cline estimates for all 77 markers used in this study and all three localities. Cubic regressions of *Q* on ancestry estimates calculated from *structure* output (evidence for specific homozygosity, black; evidence for interspecific heterozygosity, gray) are shown (see main text for details). Positive ancestry values indicate a surplus of *P. alba* or heterozygote genotypes, respectively. Solid lines, Italy; dashed lines, Austria; dash dotted lines, Hungary. Marker names including chromosome number are shown; dominant loci are tagged with an asterisk.

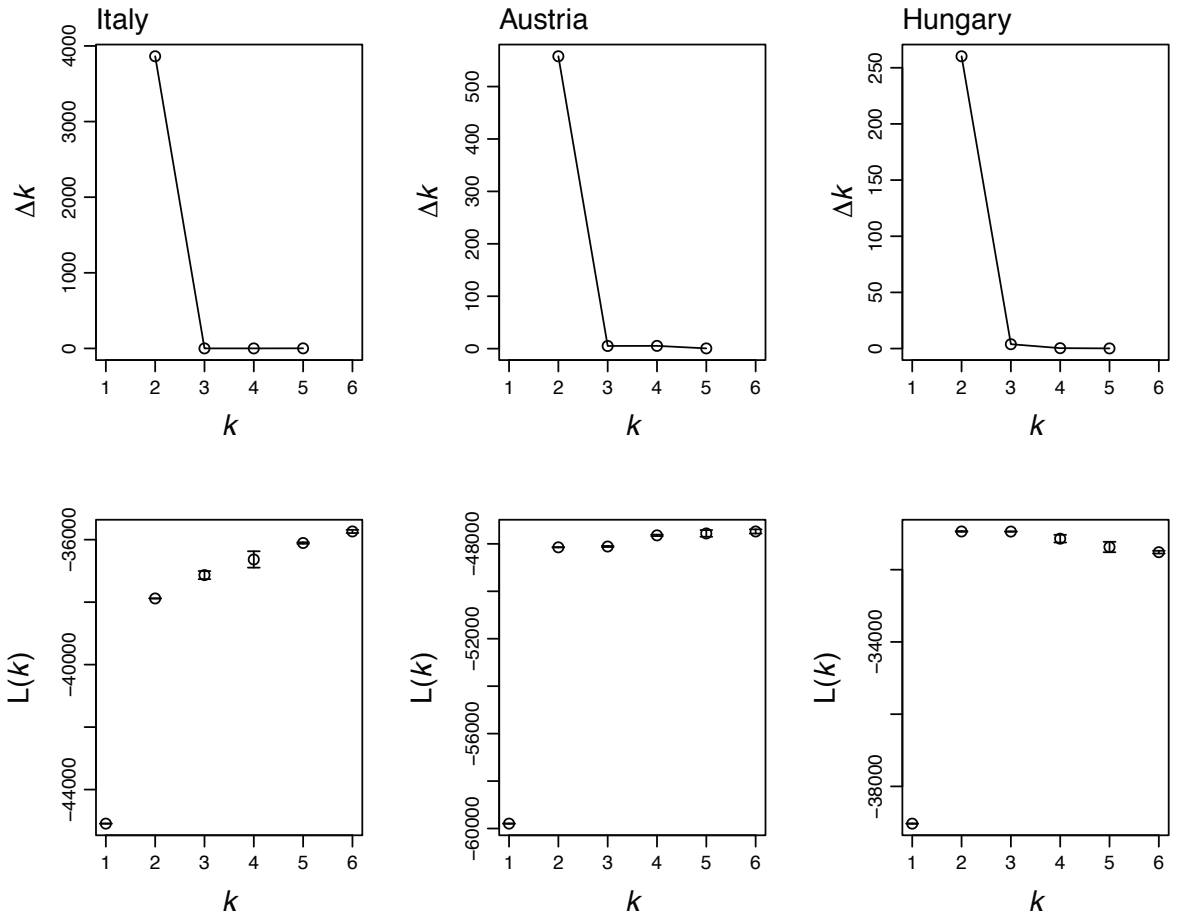


Figure S2: Most likely number of genetic clusters *k* at the three localities. Δ*k* (top) was calculated after the method of Evanno *et al.* (2005) from L(*k*). L(*k*) (bottom) was obtained from three replicate MCMC chains of *k* from one to six (circles are mean L(*k*), error bars are standard deviations).

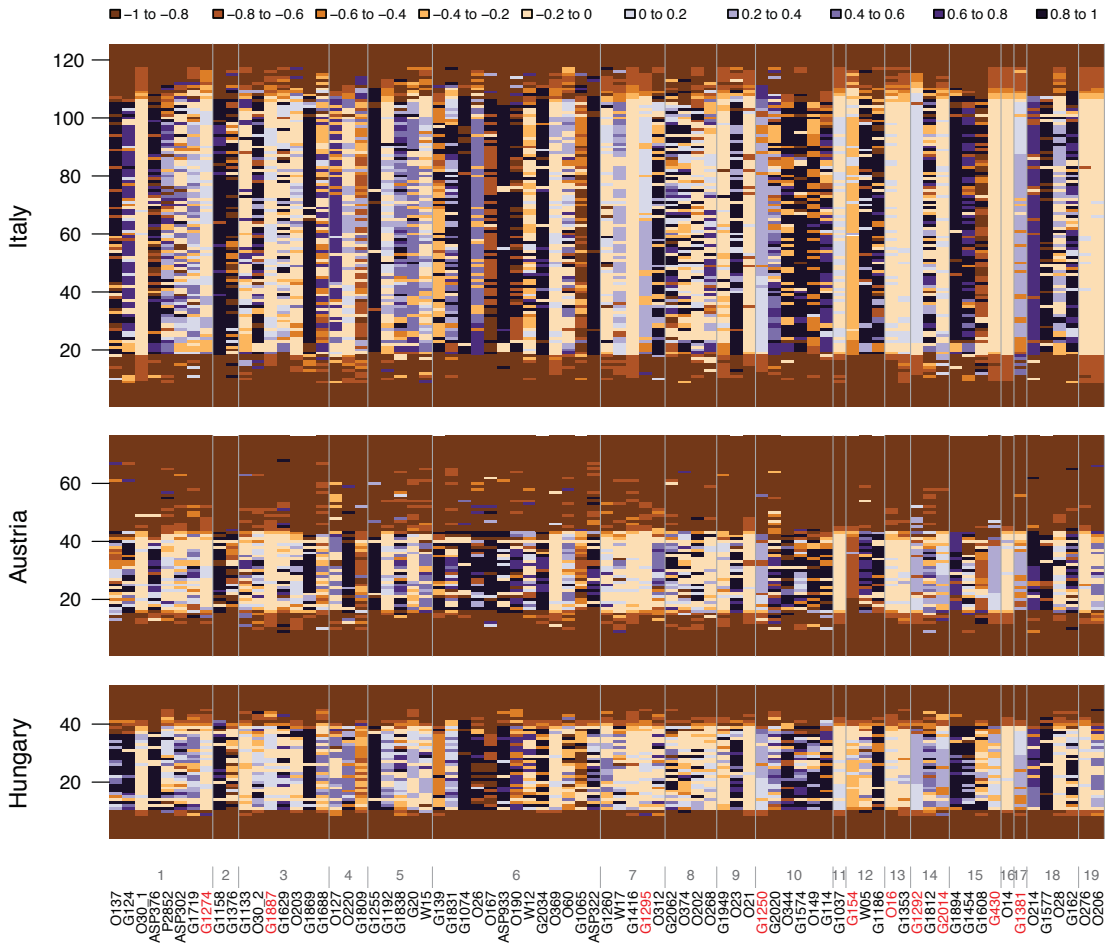


Figure S3: Locus specific ancestry (LSA) described by evidence for interspecific heterozygosity at the three localities for all potentially admixed plus eight of each of the parental-like individuals (y-axis), ordered by *Q* value as in Fig. 2, main article (at each locality: top, *P. alba* like; middle, admixed; bottom, *P. tremula* like), analyzed with 77 markers across 19 chromosomes in map position (x-axis). Dark brown colors correspond to strong evidence for being a specific homozygote (value close to -1), dark blue colors correspond to strong evidence for being an interspecific heterozygote (value close to +1) (see main article for details). Pale colors result from uncertainty of belonging to either one or the other class. Output from the *structure* linkage model was used to calculate evidence for interspecific heterozygosity. Light gray vertical lines mark boundaries between the 19 chromosomes. Marker names in red indicate dominant markers.

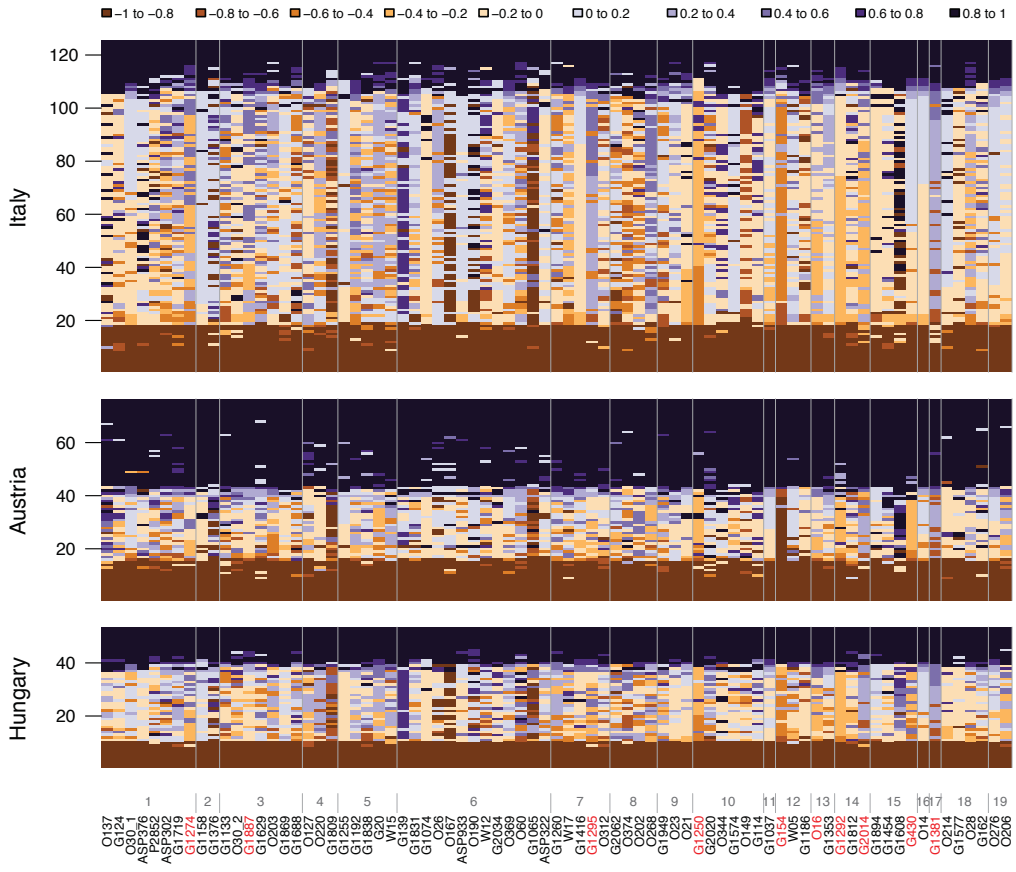


Figure S4: Locus specific ancestry (LSA) described by evidence for specific homozygosity at the three localities for all potentially admixed plus eight of each of the parental-like individuals (y-axis), ordered by *Q* value as in Fig. 2, main article (at each locality: top, *P. alba* like; middle, admixed; bottom, *P. tremula* like), analyzed with 77 markers across 19 chromosomes in map position (x-axis). Dark brown colors correspond to strong evidence for being homozygous for *P. tremula* alleles (value close to -1), dark blue colors correspond to strong evidence for being homozygous for *P. alba* alleles (value close to +1) (see main article for details). Pale colors result from uncertainty of belonging to either one or the other class. Output from the *structure* linkage model was used to calculate evidence for specific homozygosity. Light gray vertical lines mark boundaries between the 19 chromosomes. Marker names in red indicate dominant markers.

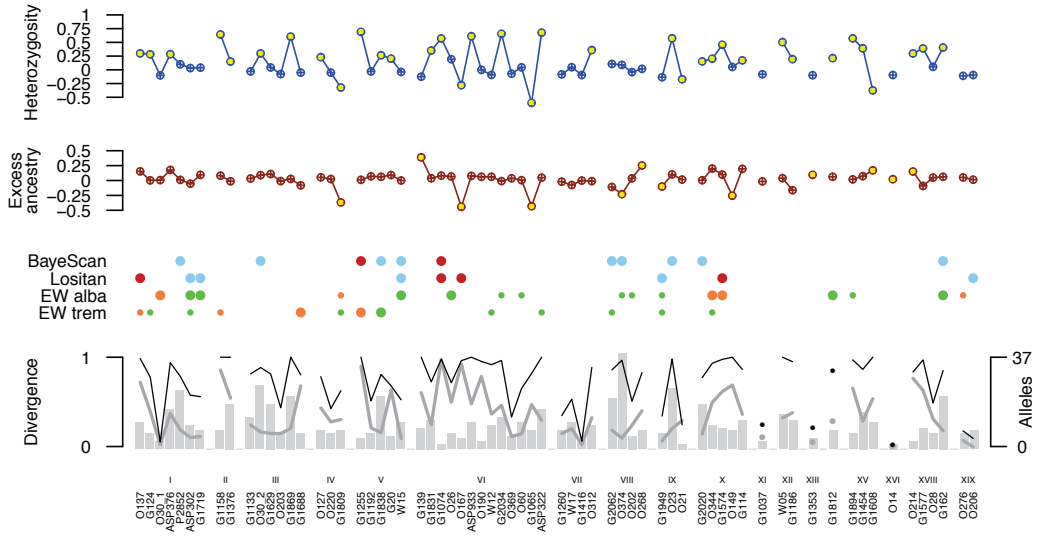


Figure S5: Summary for 67 codominant loci for the Italy hybrid zone, including interspecific heterozygosity and excess ancestry in admixed individuals, and divergence- and diversity-based genome scans involving the two parental species *P. alba* and *P. tremula*. From top to bottom: Blue, interspecific heterozygosity; brown, excess ancestry; yellow points indicate values exceeding simulated 95% CIs (see Figs. 4 and 5, main article). Divergence based genome scans (*BayeScan* and *lositan*): dark red, loci identified as under directional selection; light blue, loci identified as under balancing selection. Diversity based genome scans (Ewens-Watterson test for *P. alba*, EW alba; for *P. tremula*, EW trem): orange, loci identified as under directional selection; green, loci identified as under balancing selection. For the Ewens-Watterson tests, loci identified after controlling for the false discovery rate (FDR) at FDR = 0.1 are indicated with large points. See Text S1 for more details on outlier scans. Black line, allele frequency differentials (δ); bold gray line, F_{ST} values; gray bars, numbers of alleles at each locus (see Fig. 3, main article).

Text S1: Genome scans for parental populations

METHODS

Genome scans in parental populations. In addition to the analysis of LSA in admixed individuals, genome scans were performed by using measurements for genetic divergence (F_{ST}) to capture the dynamics of gene exchange between parental populations in the more distant past (Whitlock 1992), and to detect loci that might be subject to selection within each parental population by investigating genetic diversity. All genome scans were done separately for each of the three study sites.

F_{ST} outlier detection. Loci showing unusually low or high levels of genetic divergence, compared to the remainder of the genome, provide evidence that their evolutionary history is extreme relative to the majority of the genome. Outlier loci might be subject to balancing or directional selection, or have been shaped by other exceptional evolutionary processes. For loci with more than two alleles, caution is needed when using the statistic F_{ST} to quantify the level of divergence, as this measurement is highly dependent on the level of genetic variation (Hedrick 2005). The method to detect divergence outliers proposed by Beaumont & Nichols (1996) plots F_{ST} against heterozygosity and compares this distribution with that expected under a symmetrical Island Model (Wright 1951). To detect outlying values of F_{ST} between *P. alba* and *P. tremula*, we used the *fdist* approach based on the method of Beaumont & Nichols (1996) by applying *lositan*, which is software constructed around *fdist2* (Antao *et al.* 2008). We carried out three replicate analyses in *lositan* for each of the three hybrid zones using 95,000 simulations under the infinite alleles model, applying the recommended settings (1) to compute the initial mean F_{ST} after removing potentially selected loci in a first run, and (2) to force the average simulated F_{ST} to approximate the average value found in the real data set. Tests were performed given the null-hypothesis that simulated F_{ST} < sampled F_{ST} at $\alpha = 0.05$, where $p < \alpha$ and $p > 1 - \alpha$ can be used as a hint for balancing and directional selection, respectively. Only loci showing significant values in all three runs were regarded as outliers.

The outlier detection approach of Beaumont & Nichols (1996), *fdist2*, can be misleading because real populations will often deviate from a symmetrical Island Model (Beaumont & Balding 2004; Foll & Gaggiotti 2008; Excoffier *et al.* 2009b; Nielsen *et al.* 2009). Thus, we additionally applied a hierarchical-Bayesian method to identify outlier loci, implemented in the software *BayeScan* (Foll & Gaggiotti 2008). This approach, which is based on the method of Beaumont & Balding (2004), decomposes estimates of locus-population specific estimates of F_{ST} into population-specific (β) and locus-specific (α) components. A positive value of α suggests that the locus might be under directional selection, whereas a negative value can indicate balancing selection. We chose a threshold of log10 (Bayes Factor) > 2, corresponding to a *p*-value > 0.99, as 'decisive' evidence (Jeffreys 1961) that the preferred model includes α. We ran *BayeScan* using 10 pilot runs of 5,000 iterations for parameter tuning, followed by a burn-in of 100,000 and a final number of 1,000,000 iterations with a thinning interval of 50, as these settings showed a good convergence between different chains.

Genetic diversity based outlier detection. A drawback of using microsatellite loci for genome scans is their heterogeneity in mutation rate (Schlotterer 2000; Ellegren 2004), which might erroneously lead to the conclusion that a locus is exceptional and potentially under selection. Thus, in addition to the F_{ST} based scans, we also made use of a genetic diversity based method (Ewens-Watterson test) that is independent of locus-specific mutation rates.

The Ewens-Watterson test of selective neutrality is done separately for each locus and population and is based on the formula derived by Ewens (1972) to calculate the expected number of different alleles at a locus at mutation-drift equilibrium. To test if the expected neutral distribution deviates from the observed allele frequency distribution, expected Hardy-Weinberg homozygosity of the allele frequencies under neutrality (F_{exp}) is compared to the observed Hardy-Weinberg homozygosity ($F_{obs} = \sum p_i^2$, where p_i is the frequency of allele i) (Watterson 1978). To achieve this, F_{exp} is computed for a large number of random samples exhibiting the same number of alleles and individuals as the original population. The probability of observing random samples with F_{exp} identical or smaller than F_{obs} is recorded and can be used as an indication for balancing selection (excess of heterozygotes, $F_{obs} < F_{exp}$, $p < \alpha$) or directional selection (excess of homozygotes, $F_{obs} > F_{exp}$, $p > 1 - \alpha$). Ewens-Watterson tests, using 50,000 simulated samples each at $\alpha = 0.05$, were carried out with *arlequin* 3.5 (Excoffier & Lischer 2010). To control for multiple testing, the false discovery rate (FDR) was controlled using the algorithm described in Benjamini & Hochberg (2000) at FDR = 0.1.

RESULTS

Genomic divergence (F_{ST}) based outlier detection in the parental populations: Highly variable levels of allelic diversity across loci (Table 3 and Fig. 3, main article; Table S2) presented a challenge in identifying departures from neutrality using genome scans based on interspecific F_{ST}. By using *BayeScan* and *lositan* to detect loci potentially under selection, we found that loci with large numbers of alleles (pointing to high mutation rates) tended to be detected as potentially under balancing selection, even when almost no alleles were shared between species (Table S4; Fig. 3, main article; compare relative numbers of alleles with estimates of F_{ST} and δ). This was more pronounced for the *BayeScan* method, which is known to be sensitive to differences in mutation rates (Foll & Gaggiotti 2008). Keeping in mind the well-known limitations of F_{ST} based genome scans for marker loci with variable mutation rates (Butlin 2010; Feder & Nosil 2010; Michel *et al.* 2010), a striking feature of our interspecific F_{ST} scans is the great concordance of genomic patterns across localities (Table S4), which in turn is consistent with highly similar patterns of ancestries across localities in hybrids (main article).

Genetic diversity based outlier detection in the parental populations: In the Ewens-Watterson test, which is independent of locus specific mutation rates, a large number of loci was identified as being non-neutral, most of them in Italy (38.8% of loci with any signature), and more in *P. alba* than in *P. tremula* (43.3% and 29.9% of loci with any signature, respectively). Of the non-neutral loci, more were consistent with balancing (55.8%) than directional selection (44.2%). Only few loci showed the same signature of selection in more than one locality within a species (Table S4). Because of multiple testing, three to four loci in each of the Ewens-Watterson test series are

expected to randomly deviate from neutrality in both of the tested directions. However, as patterns shared between populations are unlikely to result from chance, statistically significant outliers with and without FDR are reported.

Association between divergence and diversity based outliers in parentals and interspecific heterozygosity in admixed individuals: No association between loci showing a deficiency or excess of interspecific heterozygosity (outside simulated 95% CIs, Figs. 4 and 5, main article) and loci identified as under directional or balancing selection by *BayeScan*, *lositan*, or Ewens-Watterson tests could be detected (Table S4 and S5, Fig. S5).

DISCUSSION

Genomic patterns and outlier detection in parental populations: The high level of genetic differentiation between the parental species (mean across localities: F_{ST} = 0.34; δ = 0.71) reflects their long divergence time of at least several million years (Stettler *et al.* 1996) and indicates the presence of strong RI. In particular, the high δ at most, but not all, markers is consistent with divergent species that still share a portion of their genomes via gene flow (Wu 2001; Fig. 3, main article). Accordingly, F_{ST} outlier detection based on deviations from putatively neutrally evolving regions (which might themselves be weakly selected) will only be suitable to detect the most extreme regions (Michel *et al.* 2010) that are not necessarily associated with RI (Via 2009; Feder & Nosil 2010). In addition, many evolutionary processes other than selection are likely to vary across the genome and to affect outlier detection (Buerkle *et al.* 2011), and much of the differentiation directly involved in RI might be obscured by divergence hitchhiking and genetic drift occurring after initial divergence (Via & West 2008; Via 2009; Feder & Nosil 2010). These general issues related to the detection of F_{ST} outlier loci, combined with the high mean F_{ST} between the two species, make divergence scans an inappropriate tool to detect genetic regions involved in RI. Likewise, inferences on balancing selection are confounded by the known sensitivity of F_{ST} to differences in mutation rates (Hedrick 2005; Foll & Gaggiotti 2008), which are strongly suggested by the weak relationship between F_{ST} and simple allele frequency differentials δ for loci with high numbers of alleles in our study (Fig. 3, main article).

Regardless of these specific issues related to F_{ST} outlier detection, the high similarity of patterns of genetic distance between *P. alba* and *P. tremula* across localities (Fig. 3, main article; Table S2 and S4) is in line with consistently acting, intrinsic mechanisms of isolation or high levels of intra-specific gene flow. The reduced congruence among replicate localities for loci identified by a genetic diversity-based neutrality test (Ewens-Watterson test; Table S4), on the other hand, might reflect locally varying ecological selection, which has already been shown to be operating in *P. tremula* (De Carvalho *et al.* 2010), or differences in the spatial or demographic history of the local populations (Excoffier *et al.* 2009a).

Association between divergence and diversity based outliers in parentals and interspecific heterozygosity in admixed individuals: The missing association between F_{ST} outliers for parental populations and loci with increased or decreased interspecific heterozygosity in hybrids results probably from the inapplicability of divergence based outlier scans for highly divergent species like the examined pair of *Populus* and

general issues related to the detection of F_{ST} outlier loci, further complicated by the use of microsatellite markers (see above). The missing association between diversity based outliers for the parental populations and increased or decreased interspecific heterozygosity in hybrids might reflect that most of the within species patterns of selection are not associated with loci involved in RI between species or vice versa, or that selective sweeps involved in RI are too ancient to be detected.

REFERENCES

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a F-st-outlier method. *BMC Bioinformatics* **9**.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology* **13**, 969-980.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London Series B-Biological Sciences* **263**, 1619-1626.

Benjamini Y, Hochberg Y (2000) On the adaptive control of the false discovery fate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics* **25**, 60-83.

Buerkle CA, Gompert Z, Parchman TL (2011) The n = 1 constraint in population genomics. *Molecular Ecology* **20**, 1575-1581.

Butlin RK (2010) Population genomics and speciation. *Genetica* **138**, 409-418.

De Carvalho D, Ingvarsson PK, Joseph J, *et al.* (2010) Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Molecular Ecology* **19**, 1638-1650.

Ellegren H (2004) Microsatellites: Simple sequences with complex evolution. *Nature Reviews Genetics* **5**, 435-445.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Ewens WJ (1972) Sampling theory of selectively neutral alleles. *Theoretical Population Biology* **3**, 87-112.

Excoffier L, Foll M, Petit RJ (2009a) Genetic consequences of range expansions. *Annual Review of Ecology Evolution and Systematics* **40**, 481-501.

Excoffier L, Hofer T, Foll M (2009b) Detecting loci under selection in a hierarchically structured population. *Heredity* **103**, 285-298.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.

Feder JL, Nosil P (2010) The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution* **64**, 1729-1747.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977-993.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* **59**, 1633-1638.

Jeffreys H (1961) *Theory of probability*, 3rd edn. Clarendon Press, Oxford.

Michel AP, Sim S, Powell TH, *et al.* (2010) Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences of Sciences of the United States of America* **107**, 9724-9729.
Nielsen EE, Hemmer-Hansen J, Poulsen NA, *et al.* (2009) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology* **9**.
Schlotterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**, 365-371.
Smulders MJM, Van Der Schoot J, Arens P, Vosman B (2001) Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). *Molecular Ecology Notes* **1**, 188-190.
Stettler RF, Bradshaw HD, Heilman PE, Hinckley TM (1996) Biology of *Populus* and its implications for management and conservation. *NCR Research Press*, Ottawa.
Tuskan GA, DiFazio S, Jansson S, *et al.* (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596-1604.
Tuskan GA, Gunter LE, Yang ZMK, *et al.* (2004) Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* **34**, 85-93.
van der Schoot J, Pospiskova M, Vosman B, Smulders MJM (2000) Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.). *Theoretical and Applied Genetics* **101**, 317-322.
Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 9939-9946.
Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology* **17**, 4334-4345.
Watterson GA (1978) Homozygosity test of neutrality. *Genetics* **88**, 405-417.
Whitlock MC (1992) Temporal fluctuations in demographic parameters and the genetic variance among populations. *Evolution* **46**, 608-615.
Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323-354.
Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**, 851-865.