

# Patterns of genetic diversity and differentiation in resistance gene clusters of two hybridizing European *Populus* species

Celine Caseys<sup>1,2</sup>  · Kai N. Stölting<sup>1</sup> · Thelma Barbará<sup>1</sup> · Santiago C. González-Martínez<sup>3</sup> · Christian Lexer<sup>1,4</sup>

Received: 23 March 2015 / Revised: 10 June 2015 / Accepted: 24 June 2015 / Published online: 24 July 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Resistance genes (R-genes) are essential for long-lived organisms such as forest trees, which are exposed to diverse herbivores and pathogens. In short-lived model species, R-genes have been shown to be involved in species isolation. Here, we studied more than 400 trees from two natural hybrid zones of the European *Populus* species *Populus alba* and *Populus tremula* for microsatellite markers located in three R-gene clusters, including one cluster situated in the incipient sex chromosome region. The results indicate that all three R-gene clusters present extensive linkage disequilibrium (LD). Outlier tests suggest balancing selection as a potential driver of R-gene diversity, but divergent selection was also detected. Nucleotide-binding site leucine-rich repeat (NBS-LRR) type R-gene clusters exhibit low species differentiation and appear to be affected by shared selection pressures between species, recurrent interspecific gene flow, or

both. Our results are consistent with a role for R-gene clusters in the evolution of chromosome XIX, recently put forward as an incipient sex chromosome in *Populus*.

**Keywords** Resistance gene cluster · NBS-LRR · Plant defense · Hybridization · Sex chromosome · Selection

## Introduction

The understanding of mechanisms involved in the evolution of reproductive isolation and speciation is of great current interest in the field of population genomics (The Marie Curie SPECIATION Network 2011; Abbott et al. 2013). Studying the genomics of taxa with “porous genomes” provides an opportunity to obtain insights into the genetics of adaptation, reproductive isolation, and speciation (Wu 2001; Lexer and Widmer 2008; Feder et al. 2012). During the process of “speciation with gene flow,” reproductive isolation evolves gradually from specific isolated genome regions of different sizes (genomic islands or continents of speciation) to complete genomic isolation and post-speciation divergence (Feder et al. 2012). Before genomic isolation is complete, genetic exchange can occur uni- or bi-directionally and lead to introgression of genome segments due to neutral or selective forces (Arnold 2006; Petit and Excoffier 2009). Reproductive isolation was shown to result from a great number of different pre- and postzygotic mechanisms in animals and plants, including sexual, genetic, and ecological barriers (Smadja and Butlin 2011). In some animal species, sex determination regions host species isolation genes (reviewed by Qvarnström and Bailey 2009; Ellegren 2009). In plants, resistance genes have been put forward for their potential role in the origin and maintenance of reproductive barriers through hybrid necrosis (Bomblies and Weigel 2007; Bomblies 2009).

Communicated by P. Ingvarsson

This article is part of the Topical Collection on *Disease Resistance*

**Electronic supplementary material** The online version of this article (doi:10.1007/s11295-015-0904-8) contains supplementary material, which is available to authorized users.

✉ Celine Caseys  
celine.caseys@gmail.com

- <sup>1</sup> Unit of Ecology and Evolution, Department of Biology, University of Fribourg, Chemin du Musée 10, 1700 Fribourg, Switzerland
- <sup>2</sup> Department of Botany and Biodiversity Research Centre, University of British Columbia, 6270 University Blvd, Vancouver, BC V6T 1Z4, Canada
- <sup>3</sup> Department of Forest Ecology and Genetics, INIA-Forest Research Centre, 28040 Madrid, Spain
- <sup>4</sup> Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, 1030 Vienna, Austria

The two European *Populus* species, *Populus alba* and *Populus tremula*, have been described as species with incomplete reproductive barriers that hybridize despite relatively high levels of genomic divergence (Lexer et al. 2010; Lindtke et al. 2012; Stölting et al. 2013), largely concordant with available models of “porous genomes” and heterogenous genomic divergence (Wu 2001; Feder et al. 2012). Postzygotic selection has recently been put forward as a mechanism potentially maintaining the species barrier, but the exact selective factors acting on hybrid seedlings are currently unknown (Lindtke et al. 2014).

Despite high overall genomic divergence in these species, it was shown that the proximal end of chromosome XIX presents a block of low interspecific divergence (Macaya-Sanz et al. 2011; Stölting et al. 2013). This genome region was proposed by Yin et al. (2008) as an incipient sex chromosome, i.e., a chromosome morphologically similar to autosomes but in the process of developing characteristics of sex chromosomes. This chromosomal region is not only particular regarding its genetic diversity but also regarding its genomic architecture. Despite high levels of synteny among *Populus* genomes, the location of the gender determination locus appears to be variable in position along chromosome XIX in different sections of the *Populus* genus (Gaudet et al. 2008; Yin et al. 2008; Pakull et al. 2011; Kersten et al. 2014). In *P. alba* and *P. tremula*, sex determination appears to have a centromeric localization rather than the peritelomeric localization described for *Populus trichocarpa* (Tuskan et al. 2012). These different positions may point to the presence of groups of sex-determining and/or sexually antagonistic mutations along the chromosome, effectively representing a gender “super locus” due to recombination suppression (which has been documented in both subgenera; Yin et al. 2008; Macaya-Sanz et al. 2011; Gerales et al. 2015).

This block of low divergence on chromosome XIX is likely not only involved in sex determination but also in defense mechanisms: (a) the region contains the largest super cluster of NBS-LRR genes found in the *Populus* genome with 83 potentially encoding R-genes (Kohler et al. 2008; Bresson et al. 2011); (b) This super cluster co-locates with *Melampsora* rust resistance loci MER, R1, and RUS in *Populus deltoides* × *P. trichocarpa* (Yin et al. 2004; Jorge et al. 2005; Bresson et al. 2011); (c) the presence of 20 BEAF and DREF (BED finger) DNA-binding domains within R-genes of this super cluster could speed up downstream signaling by direct regulation of transcription. Some of these BED-NBS-LRR genes were also found in the grapevine and rice genomes, but the extent of these chimeric proteins in poplar seems unique, and this gene family presents an exceptional diversity on chromosome XIX (Germain and Séguin 2010).

The co-localization of sex determination and R-genes on chromosome XIX poses the question whether these two features have co-evolved (Tuskan et al. 2012). To understand the

potential co-evolutionary role of R-genes in this special region of the genome, we analyzed three R-gene clusters, including the super cluster on the incipient sex chromosome and two clusters on autosomal chromosomes. More specifically, we chose a small cluster of six NBS-LRR genes on chromosome III and the MXC3 region on chromosome IV (Yin et al. 2004) conferring resistance to *Melampsora* × *columbiana* rust hybrid species. For the MXC3 region, suppression of recombination has been shown in North American *Populus* species, similar to recent findings for the poplar sex determination region on chromosome XIX (Yin et al. 2008). No NBS-LRR genes were found in the MXC3 region, but five putative resistance genes were identified, including two thaumatin-like pathogenesis-related proteins and three receptor-like kinases (Yin et al. 2004; Germain and Séguin 2010). These pathogenesis-related (PR) proteins participate in plant-inducible immunity via their accumulation in response to biotic or abiotic stress. Indeed, transcriptome-based studies confirmed the role of PR proteins in resistance to *Melampsora* fungi (Petre et al. 2011).

Natural hybrid zones between species of the same ploidy level provide excellent opportunities to study genetic isolation, introgression, and the evolution of ecological divergence. Hybrid zones are often found along ecotones and thus are expected to present greater abundances of herbivores and pathogens presenting new challenges and opportunities for plant resistance (Traw and Bergelson 2010). In this study, we use two natural hybrid zones and adjacent populations of the parental species, situated in the Ticino river valley in Italy and Tisza river valley in Hungary, to compare patterns of genetic diversity and differentiation along three genome regions involved in plant resistance. We make use of microsatellite markers located within these R-gene clusters to address the following questions related to the interplay of R-gene evolution, sexual differentiation, and species differentiation in these two European *Populus* species: (1) What is the level of intra- and interspecific genetic diversity in resistance gene clusters, and how much do these genome regions contribute to species differentiation? (2) How great or small is the potential of these R-gene clusters for interspecific introgression in European *Populus* species? (3) How distinct is the super R-gene cluster on chromosome XIX, the *Populus* putative incipient sex chromosome, from other R-gene clusters with regard to genetic diversity, species differentiation, and linkage disequilibrium (LD)?

## Materials and methods

### Plant materials

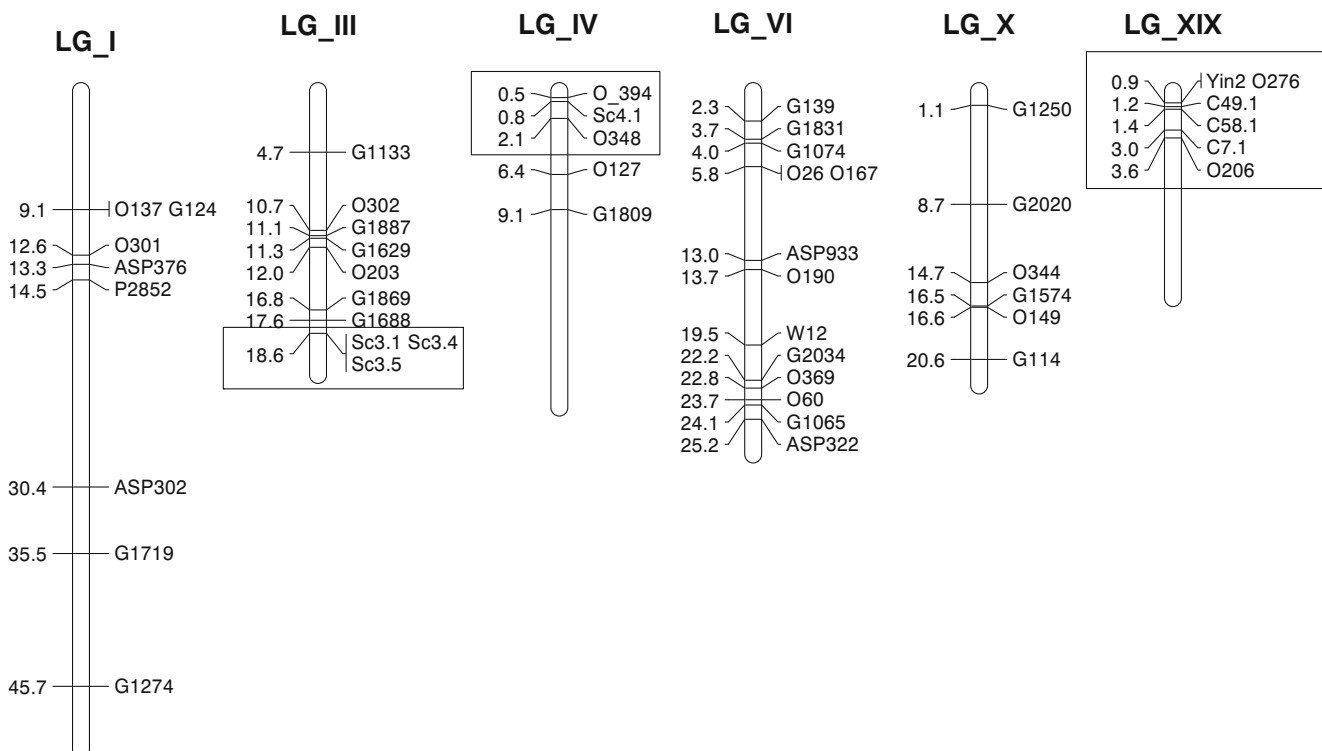
We analyzed a total of 405 individuals sampled in two natural hybrid zones of *P. alba* (white poplar) and *P. tremula*

(European aspen). These individuals were previously characterized as parental species or hybrids according to their hybrid index measured as the admixture proportion ( $Q$ ) assessed with 77 microsatellite markers (Lexer et al. 2010; Lindtke et al. 2012). Individuals were classified as *P. alba* when  $Q > 0.95$ , as hybrids when  $0.95 > Q > 0.05$ , and as *P. tremula* when  $Q < 0.05$ . In the Ticino river (Italy) and Tisza river (Hungary) hybrid zones, respectively, 42 and 96 *P. alba*, 68 and 56 *P. tremula*, and 109 and 34 hybrids were identified.

### Molecular genetic markers

Three R-gene clusters were characterized with a total of 12 microsatellite markers: three markers spanning 64 kilobases (kb) on chromosome III, three distributed across 1.3 megabases (Mb) on chromosome IV, and six spanning 2.76 Mb on chromosome XIX (Fig. 1). The six markers on the proximal end of chromosome XIX were included because of their positions within the largest resistance gene cluster present in the *Populus* genome, as identified in previous studies (Macaya-Sanz et al. 2011), and because of its potential involvement in sex determination (Yin et al. 2008). To amplify the target genomic regions on chromosomes III and IV, 16 primer pairs were designed based on *P. trichocarpa* genome (version 2), and microsatellite loci were tested for amplification and allelic diversity in the focal European

*Populus* species. Three of the microsatellite markers developed from chromosome III and one on chromosome IV in addition to two markers from Yin et al. (2004) were polymorphic and were thus retained for all subsequent analyses (Table S1). Polymerase chain reactions (PCR) for DNA amplification followed Lexer et al. (2005). Forward primers for markers on chromosome XIX included fluorescent dyes (FAM, VIC, NED, and PET from Applied Biosystems), whereas markers on chromosomes III and IV were labeled using a M13-tail protocol (Schuelke 2000). Amplified fragment sizes were analyzed on an Applied Biosystems 3130 Genetic Analyzer with LIZ (Applied BioSystems) as internal size standard. Genetic markers positioned outside of R-gene clusters on chromosomes III and IV and markers from chromosomes I, VI, and X (Fig. 1), available from a previous study (Lindtke et al. 2012), were also used for comparative purposes. These additional chromosomes were selected as a reference for genetic diversity outside of R-gene clusters. Marker positions (Fig. 1) were assessed by BLAST searches against the *P. trichocarpa* genome assembly (built *P. trichocarpa\_156*) available on phytozome website. Chromosome XIX marker positions are difficult to assess, since both sequence content and marker positions tend to change between different *P. trichocarpa* genome assemblies (Geraldes et al. 2015). However, most of the chromosome XIX markers used in the present study were shown to be genetically linked on linkage



**Fig. 1** *Populus* chromosomes showing the location of microsatellite marker loci included in this study. Map of physical distances in Mb according to the *P. trichocarpa* genome sequence (V2) are indicated on

the left of each chromosome. Boxes indicate the focal markers of the present study, selected specifically for their proximity to resistance gene clusters

group XIX in a *P. alba* × *P. tremula* genetic map (Macaya-Sanz et al. 2011), within the usual limits imposed by marker polymorphism in a single mapping pedigree. Thus, wherever possible, our discussion of genetic diversity and differentiation on chromosome XIX is based on genetic map distances rather than physical distances from various *P. trichocarpa* genome assemblies, as also recommended by Gerald et al. (2015). Physical distances in the present study were estimated from marker positions in *P. trichocarpa* genome and converted to genetic distances based on an estimated 1 centimorgan (cM) per 200 kb. This rough approximation of relationships between physical and genetic distances, estimated from the *P. trichocarpa* genome sequence (Tuskan et al. 2006), represents a simplification, since recombination rates are known to vary greatly across genomes (Bresson et al. 2011; Slavov et al. 2012).

### Genetic diversity

The overall genetic structure and diversity of European *Populus* hybrid zones have been previously studied (Lexer et al. 2005, 2010; Lindtke et al. 2012). In this study, descriptive population genetic parameters focused on gene diversity ( $H_e$ ), allelic richness corrected for sample size, and genetic differentiation ( $F_{ST}$ ) between parental species in each hybrid zone are provided only for the target and reference genomic regions (see above). Population genetic parameters were estimated using FSTAT software (Goudet 1995). Means and standard deviations (SD) for each of these genetic descriptors were calculated over 45 microsatellite markers. Means and SD were used to compare genetic diversity and differentiation of markers within resistance gene clusters with markers outside of these genome regions.

### Genetic differentiation outlier analysis

We tested all 45 markers for signatures of natural selection within and between the two species using two different outlier detection approaches with the software BayeScan 2.1 (Foll and Gaggiotti 2008) and LOSITAN (Antao et al. 2008), which implements the  $F_{ST}$  outlier detection method implemented in Fdist2 (Beaumont and Nichols 1996). The BayeScan outlier test is based on Bayesian multinomial-Dirichlet inference to assess subpopulation  $F_{ST}$  coefficients. BayeScan runs consisted of 50,000 burn-in iterations followed by 20 runs of 5,000 iterations with a thinning interval of 10. The Fdist2 approach identifies outliers by comparing genetic data to coalescent simulations of  $F_{ST}$  conditioned on heterozygosity using an infinite island model of population structure. Outlier analyses were performed for both the infinite allele mutation (IAM) model and the stepwise mutation model (SMM). Four runs of one million simulations for each mutation model were performed to assess repeatability.

### Introgression analysis

Patterns of introgression were investigated with the R package INTROGRESS (Gompert and Buerkle 2010). This R package calculates admixture proportions and genomic clines based on maximum-likelihood methods. The genomic cline approach was applied to detect excess of homozygosity or heterozygosity for each marker compared to its genomic background along the admixture gradient present between *P. alba* and *P. tremula*. Allele frequency differentials ( $\delta$ ) measure the levels of differentiation between species and are highly correlated with  $F_{ST}$ . Ancestry plots (Fig. S1) carry qualitative information regarding the exchange of alleles among the parental species through hybrid individuals.

### Haplotype reconstruction

Haplotypes were inferred from markers in the MXC3 region on chromosome IV to obtain a better understanding of haplotype sharing between species and facilitate comparison to introgression analysis (above). The MXC3 region was selected for this computationally rather intensive analysis, because of its low allelic richness (Table S3). In diploid species, haplotypes can be obtained by different approaches, e.g., using family data with known linkage phase, cloning of individual DNA strands from diploid individuals sampled in natural populations, or inference of haplotypes with computational algorithms. Haplotype reconstruction was conducted with the Bayesian method implemented in the PHASE 2.1.1 software (Stephens et al. 2001). Following a method proposed by Koopman et al. (2006), haplotype reconstruction consisted of 10 runs of 500 iterations and thinning interval of 5 and 100 burn-in iterations. All ten runs started from different points in the input file according to the -X10 option of the software, and results from the run with the highest overall likelihood were used for further analyses.

### Linkage disequilibrium analysis

LD (non-random association of alleles between loci) was estimated between pairs of loci. Since the studied microsatellite markers are more often multiallelic than biallelic (Tables 1, 2, 3) and are of unknown linkage phase, LD was calculated as the correlation-based composite LD coefficient  $R^2$  following Zaykin et al. (2008). This composite LD measure is superior to previous methods because of its robustness under departures from Hardy-Weinberg equilibrium and suitability for analyzing multiallelic markers. In addition, to facilitate comparisons with previous work on chromosome XIX (Macaya-Sanz et al. 2011) in the same two *Populus* species,  $\chi^2$ -based multiallelic LD  $R^2$  coefficients (Table S5) were computed with Linkdos (Garnier-Gere and Dillmann 1992) web interface (<http://genepop.curtin.edu.au/linkC.html>).

**Table 1** LD coefficient  $R^2$  measured following Zaykin et al. (2008) among markers on chromosome III of *P. alba* and *P. tremula* in the Tisza (above the diagonal) and Ticino (below the diagonal) river hybrid zones

Loci	G1133	O30_2	G1629	O203	G1869	G1688	SC3.1	SC3.4	SC3.5
<i>P. alba</i>									
G1133	–	0.114	<b>0.194</b>	0.127	0.121	0.125	0.166	0.135	0.120
O30_2	0.234	–	0.108	0.093	0.101	0.112	<b>0.138</b>	0.109	<b>0.112</b>
G1629	0.173	0.187	–	0.102	0.120	0.103	0.132	0.123	0.115
O203	0.232	0.197	0.166	–	0.105	<b>0.143</b>	0.106	0.104	0.099
G1869	0.205	<b>0.203</b>	<b>0.21</b>	0.154	–	0.117	0.128	<b>0.120</b>	0.103
G1688	0.218	0.196	<b>0.253*</b>	0.110	0.140	–	0.109	0.102	0.105
SC3.1	0.260	<b>0.272</b>	<b>0.279</b>	0.208	<b>0.254</b>	<b>0.301</b>	–	0.135	0.126
SC3.4	0.196	0.182	<b>0.204*</b>	0.188	0.173	0.169	<b>0.272*</b>	–	<b>0.121*</b>
SC3.5	<b>0.238</b>	0.189	0.182	<b>0.225</b>	0.159	0.181	<b>0.251</b>	<b>0.194*</b>	–
<i>P. tremula</i>									
G1133	–	0.123	0.148	0.149	0.136	0.101	<b>0.188</b>	0.159	<b>0.179</b>
O30_2	0.111	–	0.123	0.122	0.144	0.138	0.147	0.151	0.134
G1629	0.141	0.127	–	0.111	0.155	0.129	0.119	0.146	0.137
O203	0.140	0.138	0.114	–	0.135	<b>0.177</b>	0.133	0.117	0.127
G1869	<b>0.145</b>	0.111	<b>0.141</b>	<b>0.176</b>	–	<b>0.164</b>	0.164	0.149	<b>0.162</b>
G1688	0.096	0.115	0.089	0.150	0.143	–	0.094	0.136	0.132
SC3.1	0.134	0.124	0.139	0.120	0.130	<b>0.189</b>	–	<b>0.178</b>	<b>0.163</b>
SC3.4	<b>0.142</b>	0.114	0.119	0.117	<b>0.133</b>	0.129	<b>0.148</b>	–	<b>0.154</b>
SC3.5	<b>0.139</b>	0.094	0.130	0.091	0.116	0.102	0.148	<b>0.156</b>	–

Bold values indicate significance at  $p < 0.05$ , and  $R^2$  values indicated with an asterisk are significant after Bonferroni correction for multiple testing. For the full results including  $p$ -values, see Fig. S4

## Results

### Genetic diversity and species differentiation

Estimates of genetic diversity and differentiation for mapped microsatellite marker loci (Fig. 1) indicated great variation

among chromosomes within species and also between *P. alba* and *P. tremula* and, to a lesser extent, across geographical regions (Fig. 2). Allelic richness and gene diversity within the chromosome III R-gene cluster were high, especially in *P. alba* (Fig. 2b, c). Indeed, allelic richness in both species and localities deviated more than one standard deviation from the overall

**Table 2** LD coefficient  $R^2$  measured following Zaykin et al. (2008) among markers on chromosome IV of *P. alba* and *P. tremula* in the Tisza (above the diagonal) and Ticino (below the diagonal) river hybrid zones

Loci	O394	SC4.1	O348	O127	G1809
<i>P. alba</i>					
O394	–	<b>0.201</b>	<b>0.328*</b>	0.023	0.029
SC4.1	0.000	–	<b>0.252</b>	0.159	0.059
O348	<b>0.389*</b>	0.000	–	<b>0.184</b>	0.068
O127	0.072	0.000	0.147	–	0.021
G1809	0.077	0.000	0.112	0.05	–
<i>P. tremula</i>					
O394	–	0.175	0.095	<b>0.190</b>	0.151
SC4.1	0.087	–	0.108	0.120	0.153
O348	0.123	0.042	–	0.062	0.108
O127	0.130	<b>0.207</b>	0.147	–	<b>0.166</b>
G1809	0.098	0.147	<b>0.182</b>	0.144	–

Bold values indicate significance at  $p < 0.05$ , and  $R^2$  values indicated with an asterisk are significant after Bonferroni correction for multiple testing. For the full results including  $p$ -values, see Fig. S4



**Table 3** LD coefficient  $R^2$  measured following Zaykin et al. (2008) among markers on chromosome XIX of *P. alba* and *P. tremula* in the Tisza (above the diagonal) and Ticino (below the diagonal) river hybrid zones

Loci	Yin2	O276	C49.1	C58.1	C7.1	O206
<i>P. alba</i>						
Yin2	–	<b>0.117</b>	0.109	<b>0.135*</b>	<b>0.136*</b>	0.09
O276	0.093	–	0.099	0.095	0.116	0.087
C49.1	<b>0.213</b>	0.175	–	0.108	0.089	0.112
C58.1	0.187	0.113	<b>0.244</b>	–	<b>0.178*</b>	0.114
C7.1	0.194	0.142	<b>0.249</b>	0.194	–	0.120
O206	<b>0.210*</b>	0.093	0.180	0.176	0.186	–
<i>P. tremula</i>						
Yin2	–	0.163	0.130	0.139	0.149	0.128
O276	0.123	–	0.128	<b>0.204</b>	<b>0.176</b>	0.103
C49.1	0.126	0.117	–	0.146	0.156	0.145
C58.1	0.134	0.126	0.130	–	0.153	0.117
C7.1	0.147	0.147	0.102	0.149	–	0.120
O206	0.127	0.137	0.148	0.129	0.112	–

Bold values indicate significance at  $p < 0.05$ , and  $R^2$  values indicated with an asterisk are significant after Bonferroni correction for multiple testing. For the full results including  $p$ -values, see Fig. S4

mean (Fig. 2b). In contrast, genetic diversity on LG IV was low in *P. alba* and *P. tremula* in both localities, with many estimates in *P. alba* dropping more than one standard deviation from the mean. LG XIX presented no obvious deviation in genetic diversity compared to reference linkage groups (Fig. 2b, c).

A strong decrease of  $F_{ST}$  on the proximal end of chromosome XIX was found, confirming previous results by Macaya-Sanz et al. (2011) and Stölting et al. (2013). For this specific genome region,  $F_{ST}$  values ranged from 0.000 to 0.216 and from 0.010 to 0.549 (Table S3) in the Ticino and Tisza hybrid zones, respectively, and were generally lower than the genome-wide mean ( $F_{ST}$  values of 0.354 and 0.331, respectively; Fig. 2a). A similar decrease of  $F_{ST}$  was observed in the R-gene cluster on chromosome III, with a range of  $F_{ST}$  from 0.106 to 0.156 and from 0.013 to 0.133 (Table S3). Consequently,  $F_{ST}$  for many markers fell outside of the standard deviation of all markers analyzed in the present study (Fig. 2a). In comparison,  $F_{ST}$  on chromosome IV was comparable to our reference chromosomes (Fig. 2a).

### Genetic differentiation outliers and footprints of selection in R-gene clusters

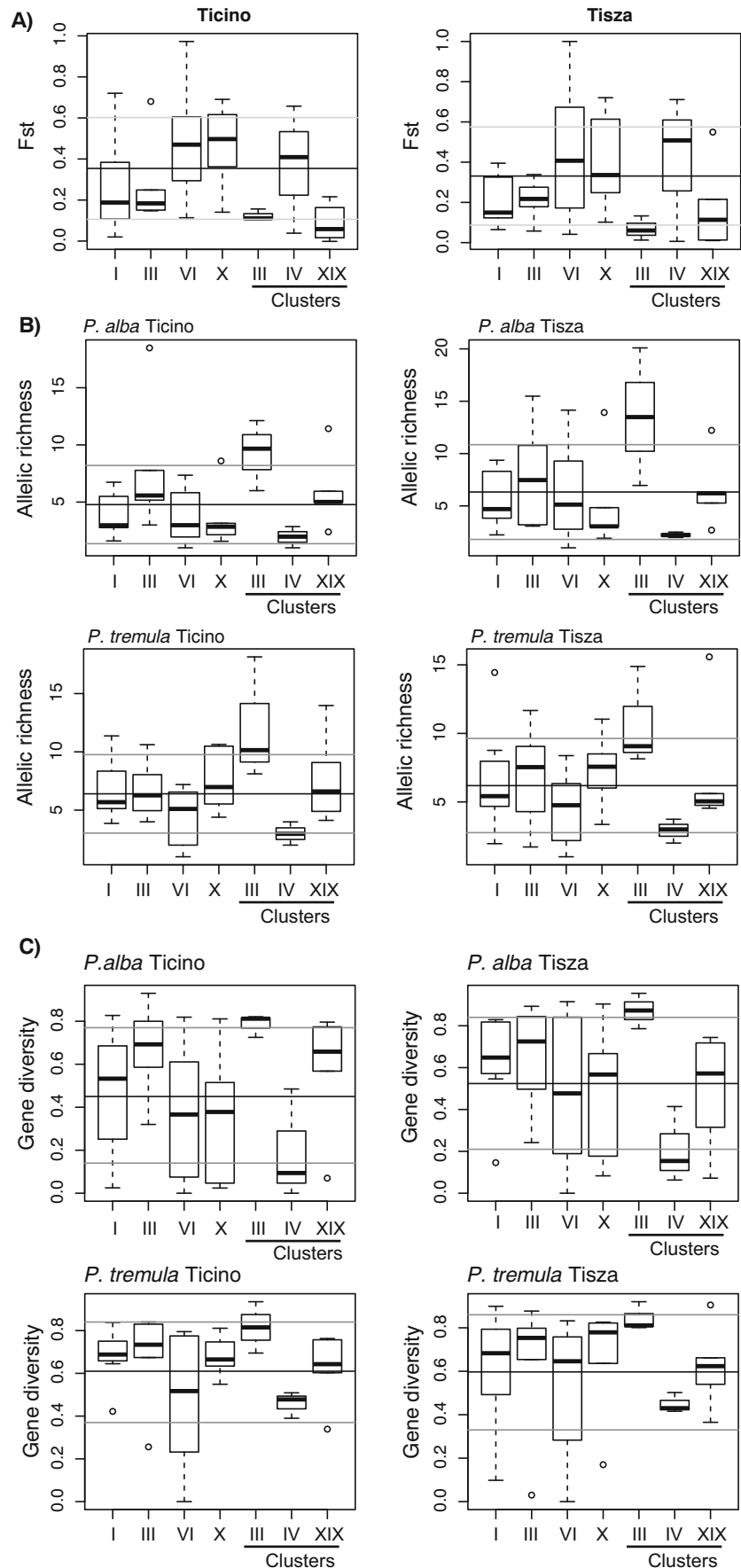
Outlier analyses revealed that R-gene clusters are targets of selection at the intra- and interspecific level (Table 4). At the species level, the results for *P. alba* and *P. tremula* were highly similar. On chromosome III, marker SC3.4 presented signs of balancing selection in both species, whereas SC3.5 presented balancing selection in *P. alba* only. On chromosome XIX, Yin2 was significant for balancing selection, whereas C7.1 and O206 showed traces of positive selection in both species and with most tests applied. Thus, balancing selection appears

to contribute to patterns of genetic diversity between species in the NBS-LRR cluster on chromosome III and also for markers of the super cluster on chromosome XIX.

### Introgression and haplotype sharing of targeted markers

We applied a genomic cline approach to detect excess of ancestry of parental alleles in hybrids and introgression in hybrid zones. Allele frequency differentials between species (Table S2) were low for chromosome XIX and the R-gene cluster on chromosome III in the Tisza river hybrid zone (see also  $F_{st}$  plots in Fig. 2a), indicating low interspecific differentiation in these genome regions, and thus low power of the genomic cline method to detect departures from neutrality. Significance of genomic clines for each marker ( $p$ -values) is indicated in Table S2. The genomic cline approach revealed little deviation in the excess of heterozygosity or homozygosity in the focal genome regions, compared to reference genome regions (Table S2). As shown by reference chromosomes (I, VI, and X), occasional departures from expected genomic clines at single markers are possible in all regions of the genome (Fig. S1). Our interest was in detecting repeated introgression of markers within targeted R-gene clusters in different hybrid zone localities. R-gene clusters did not present extensive patterns of introgression. Significant patterns of introgression were rare for R-gene clusters (C49.1 on chromosome XIX and Sc3.1 on chromosome III; see Fig. S1 and Table S2) and not repeated among hybrid zones. The MXC3 locus on chromosome IV presented no ancestry patterns consistent with introgression, but haplotype reconstruction revealed haplotype sharing between species. In

**Fig. 2** Plots of **a**  $F_{ST}$ , **b** allelic richness, and **c** gene diversity in populations of *Populus alba* and *P. tremula* adjacent to two natural hybrid zones, Ticino river (Italy) and Tisza river (Hungary). On the horizontal axis are control markers: chromosomes I (7 markers), III (6 markers outside of the R-gene cluster present on this chromosome), VI (13 markers), and X (5 markers), serving as reference genome regions for comparisons, followed by markers within R-gene clusters on chromosome III (3 markers), IV (3 markers), and XIX (6 markers). Horizontal black lines indicate the means for all studied markers (45 markers), and gray lines represent standard deviations around these means



**Table 4** Three genetic differentiation ( $F_{ST}$ ) outlier tests were carried out in *P. alba* and *P. tremula* for markers located within R-gene clusters on chromosomes (Chr.) III, IV, and XIX

Loci	Chr.	BayeScan			Fdist2	
		Alpha	PO	<i>q</i> value	IAM	SMM
<i>P. alba</i>						
2SC3.1	III	−0.022	0.064	0.731	0.491	0.532
SC3.4	III	<b>−1.215</b>	<b>1.000</b>	<b>0.000</b>	0.389	0.730
SC3.5	III	<b>−1.497</b>	<b>0.986</b>	<b>0.004</b>	0.049	0.132
O394	IV	0.065	0.113	0.583	<b>0.022</b>	<b>0.000</b>
SC4.1	IV	0.040	0.090	0.663	0.750	0.625
O348	IV	−0.117	0.129	0.513	0.282	0.325
Yin2	XIX	<b>−1.659</b>	<b>0.998</b>	<b>0.001</b>	0.049	0.076
O276	XIX	0.106	0.130	0.465	0.895	0.713
C49.1	XIX	−0.034	0.072	0.720	0.218	0.274
C58.1	XIX	−0.293	0.244	0.277	<b>0.020</b>	<b>0.026</b>
C7.1	XIX	<b>0.975</b>	<b>0.898</b>	<b>0.024</b>	<b>0.998</b>	<b>0.997</b>
O206	XIX	<b>1.143</b>	<b>0.808</b>	<b>0.052</b>	<b>0.982</b>	<b>0.999</b>
<i>P. tremula</i>						
SC3.1	III	−0.005	0.048	0.691	0.492	0.485
SC3.4	III	<b>−1.430</b>	<b>1.000</b>	<b>0.000</b>	<b>0.005</b>	0.139
SC3.5	III	0.002	0.034	0.703	0.555	0.643
O394	IV	−0.080	0.107	0.525	0.321	0.325
SC4.1	IV	−0.011	0.084	0.617	0.542	0.593
O348	IV	−0.090	0.116	0.509	0.167	0.073
Yin2	XIX	<b>−1.106</b>	<b>0.944</b>	<b>0.017</b>	0.247	0.353
O276	XIX	0.152	0.151	0.435	0.351	0.338
C49.1	XIX	−0.018	0.061	0.672	0.116	0.093
C58.1	XIX	−0.070	0.102	0.552	0.098	0.110
C7.1	XIX	<b>1.482</b>	<b>1.000</b>	<b>0.000</b>	<b>1.000</b>	<b>1.000</b>
O206	XIX	<b>1.552</b>	<b>0.999</b>	<b>0.000</b>	0.958	0.942
Inter-species						
SC3.1	III	<b>−0.978</b>	<b>0.848</b>	<b>0.022</b>	0.018	0.106
SC3.4	III	<b>−2.549</b>	<b>1.000</b>	<b>0.000</b>	<b>0.000</b>	0.161
SC3.5	III	<b>−0.915</b>	<b>0.999</b>	<b>0.000</b>	0.355	0.673
O394	IV	0.057	0.098	0.394	0.978	0.979
SC4.1	IV	0.025	0.061	0.492	0.945	0.904
O348	IV	−0.003	0.069	0.474	0.267	0.305
Yin2	XIX	<b>−2.164</b>	<b>1.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.014</b>
O276	XIX	−0.063	0.099	0.370	0.168	0.224
C49.1	XIX	−0.011	0.044	0.585	0.481	0.592
C58.1	XIX	0.006	0.038	0.633	0.484	0.594
C7.1	XIX	<b>−0.864</b>	<b>0.783</b>	<b>0.038</b>	0.079	0.145
O206	XIX	−0.017	0.049	0.574	0.569	0.513

BayeScan alpha values predict balancing ( $\alpha < 0$ ) or divergent ( $\alpha > 0$ ) selection, with high numerical values indicating strong selection. Posterior odds (PO) values indicate the probability of the model including selection, and *q* values correspond to probabilities corrected for false discovery rates (FDR). Fdist2 runs were carried out both for the infinite allele model (IAM) and the stepwise mutation model (SMM). *P* values of these models predict balancing selection for probabilities in the lower tail ( $p < 0.05$ ) and divergent selection in the upper tail ( $p > 0.95$ ). Bold values indicate significance at 10 % FDR for all outlier tests

*P. alba*, *P. tremula*, and hybrids, 5 and 7, 10 and 11, and 11 and 8 haplotypes were detected in the Ticino and Tisza river hybrid zones, respectively. Two haplotypes (haplotypes 7 and 8) were shared among all species and localities (Fig. 3).

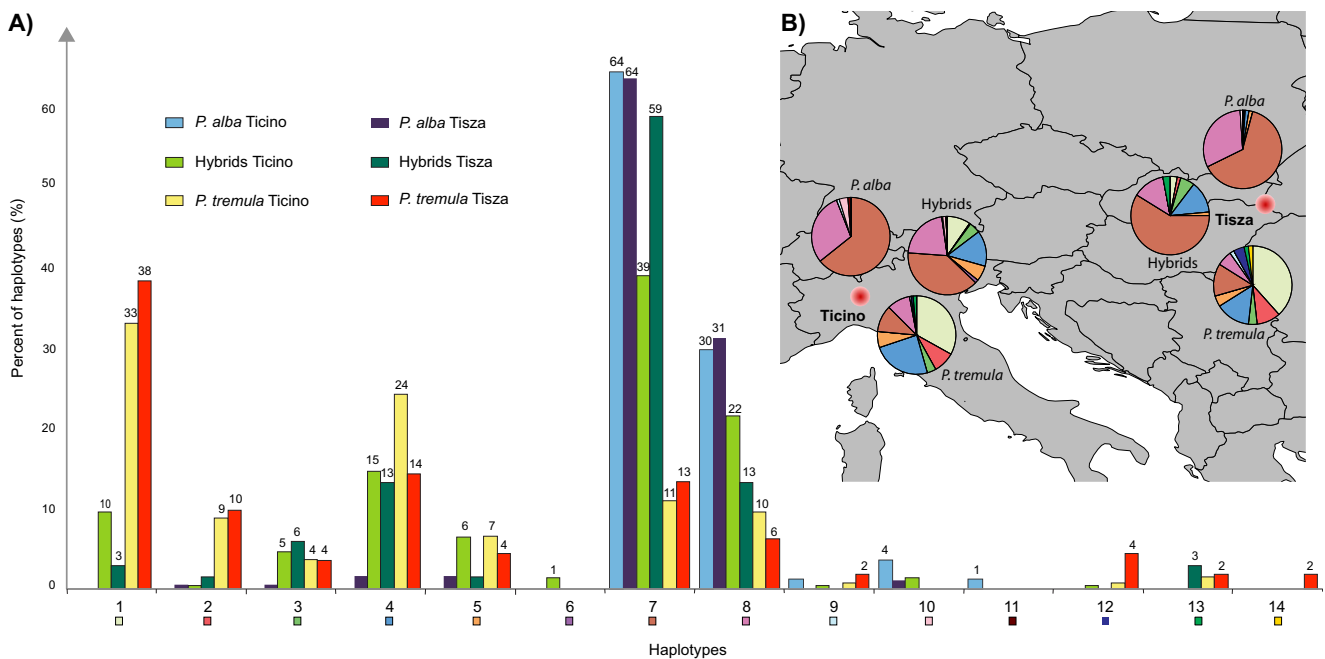
### Extent of linkage disequilibrium

Far-ranging LD was detected in all three R-gene clusters under study. LD was stronger in *P. alba* than in *P. tremula* (Tables 1–3 and S4) in accordance with previous results for effective population sizes ( $N_e$ ) in these species (lower in *P. alba*; Lexer et al. 2005). On chromosome XIX, LD extended up to 2.5 cM (~500 kb) in both species and localities (Tables 3 and S4). On chromosome III, LD was strong and extended over 0.32 cM (~64 kb) in the NBS-LRR gene cluster (Tables 1 and S4). However, strong LD was also detected in markers outside of this cluster, revealing that LD can extend up to 14 Mb in this chromosome. These measures of LD could be explained by a lack of synteny, e.g., chromosomal rearrangements between *P. trichocarpa* and European *Populus* species for this chromosome. On chromosome IV, LD was present between markers within the MXC3 locus in *P. alba* (Tables 2 and S4) extending to 1.5 Mb (~7.5 cM). In *P. tremula*, patterns of LD were generally more diffuse, consistent with its largely panmictic metapopulation structure and larger effective population size ( $N_e$ ) observed in previous studies (Lexer et al. 2005; Macaya-Sanz et al. 2011).

### Discussion

European *Populus* species have previously been shown to present high levels of genomic divergence (Lexer et al. 2010; Lindtke et al. 2012; Stölting et al. 2013). The impact of R-gene clusters on the genomics of population divergence of *Populus* species with porous genomes is poorly known, even though R-genes were put forward as gene families that, through diversification by tandem duplication followed by genomic reorganization and gene conversion, have contributed greatly to the evolution of biological responses in poplars (Tuskan et al. 2006). By contributing to responses of plants to their natural enemies, R-genes represent an important target for selective forces and co-evolution between plants and pathogens (Bergelson 2001; Rausher 2001; Keith and Olds 2013). R-genes were also put forward as potential contributors to barriers to gene flow in plants by being responsible for hybrid necrosis and autoimmune responses in hybrids (Bomblies 2009). However, genes associated with resistance to pathogens and herbivores can also be exchanged between species and can contribute to adaptation in the recipient species (Whitney et al. 2006).





**Fig. 3** Frequencies of 14 haplotypes reconstructed from three microsatellites markers located within the MXC3 region on LG IV, expressed as percentages of total numbers of haplotypes in each species and locality. **a** Haplotype frequencies with *P. alba* represented in blue, hybrids in green, *P. tremula* in orange, Ticino river hybrid zone in light

colors, and the Tisza river hybrid zone in dark color. Numbers above bars indicate rounded percentages. **b** Map of central Europe with localization of the Ticino and Tisza river hybrid zones (red dots). Pies for each species represent the percentage of haplotypes in each locality. Colors specific to each haplotype are provided below the haplotype numbers

### Evolution of R-gene clusters in European *Populus* species

The studied *Populus* species and hybrid zones are known to have distinct demographic histories due to different population structure during the last glaciation, followed by re-colonization of Europe from different glacial refugia (Fussi et al. 2010). *P. alba* is thought to have survived the last glaciation in disconnected refugia in Southern Europe, whereas *P. tremula*, less sensitive to cold, survived in various refugial areas closer to the ice shields (Lexer et al. 2010; Fussi et al. 2010; Stölting et al. 2015). Thus, the studied hybrid zones should be derived from different source populations and should thus represent fairly independent “evolutionary replicates.”

Despite the known differentiation between the Italian and Hungarian populations, this study revealed high similarity in the population dynamics of R-gene loci. Both hybrid zones presented low interspecific differentiation in the two studied NBS-LRR clusters on chromosomes III and XIX (Fig. 2). In addition, R-gene clusters on chromosome III presented outstanding genetic diversity as shown by both allelic richness and gene diversity (Fig. 2). The low interspecific differentiation observed in NBS-LRR clusters can be explained by two mechanisms, as already suggested by both microsatellite and RAD-sequencing data for chromosome XIX (Macaya-Sanz et al. 2011; Stölting et al. 2013): (1) shared selection pressure may maintain ancestral polymorphism, and (2) these gene pools are affected by recurrent introgression between these divergent species.

Outlier tests carried out in the present study suggest that NBS-LRR clusters are targets of balancing selection (Table 4) in both *P. alba* and *P. tremula*, consistent with a major role of balancing selection in the evolution of genetic variability of R-genes (reviewed by Keith and Olds 2013). The maintenance of high levels of genetic variation by balancing selection is known as the “trench warfare” model of R-gene evolution (Stahl and Bishop 2000; De Meaux and Mitchell-Olds 2003; Brown and Tellier 2011). This model explains how genetic diversity is maintained over the long term by plants and pathogens that adapt to one another via virulence and resistance factors. These mechanisms for the maintenance of polymorphism may be particularly favored in tree species with long generation time to assure recognition of quickly evolving pathogen virulence factors. On the contrary, the evolutionary “arms race” model of evolution predicts the loss of genetic diversity by the spread of advantageous genotypes in populations. Divergent selection, compatible with the arms race model, was detected in the super cluster on chromosome XIX, in which both balancing and divergent selection appear to co-occur (Table 4).

In contrast to the two studied NBS-LRR clusters, the MXC3 locus on chromosome IV presented no deviations in interspecific differentiation or genetic diversity. In addition, this region exhibited no clear patterns of selection in outlier analyses (Table 4). Thaumatin-like pathogenesis-related genes identified in the MXC3 region (Yin et al. 2004; Germain and Séguin 2010) are known to be responsive to biotic and abiotic

stress. These generalist defense genes, although identified as forming part of a rust resistance locus in North American poplars, present no trace of selection in single-locus outlier tests of the studied populations of European poplars.

### Allele sharing and LD in R-gene clusters

As an alternative to shared selection pressure maintaining ancestral polymorphism, recurrent introgression between the species could explain the low interspecific differentiation in NBS-LRR clusters on chromosomes III and XIX. Hybridization is well documented in these species (Lexer et al. 2010; Lindtke et al. 2012), and recurrent gene flow was suggested as a likely explanation for allele sharing between them (Stölting et al. 2013). However, low allele frequency differentials between species (Table S2) in each hybrid zone indicated an absence of differentiation and thus low information content for fitting genomic clines (Lindtke et al. 2012). Therefore, no strong evidence for introgression was found in the present study (Fig. S1). This suggests that more powerful analysis methods for detecting introgression (more markers, more powerful statistical methods) may yield different results than the method employed here. Future studies should address these issues using larger numbers of biallelic single nucleotide polymorphisms (SNPs), e.g., from genotyping-by-sequencing (as in Stölting et al. 2013) or low-coverage genome resequencing, and more powerful methods of ancestry analysis (e.g., Gompert and Buerkle 2012).

The MXC3 locus on chromosome IV presented no ancestry patterns indicative of introgression (Table S2), whereas haplotype reconstruction showed extensive haplotype sharing among parental species and hybrids (Fig. 3). The high levels of within-species haplotype sharing among localities (Italy and Hungary) are also remarkable, considering the known demographic histories of the two species (above).

All three R-gene clusters presented extensive LD in *P. alba* and, to a lesser extent, in *P. tremula*, in accordance with their respective effective population sizes (Lexer et al. 2005). Genome resequencing in *P. trichocarpa* showed that  $R^2$  LD coefficients should drop below 0.2 within 3–6 kb (Slavov et al. 2012), whereas higher values (0.389) were found in *P. alba* (Ticino locality, see Tables 1–3 and S4) in the MXC3 region for distances of up to 1.5 Mb. Our results on LD on chromosome XIX (Tables 3 and S4) confirm available data from different populations and a different set of markers (Macaya-Sanz et al. 2011). The extent of LD in R-gene clusters may be explained by selection, epistasis among defense components, or participation of different genes of the cluster to heterodimer proteins or even larger protein complexes (reviewed by Bomblies 2009). Such interactions were found in different plant species and were suggested to contribute to hybrid necrosis, a phenomenon awaiting further study in controlled or open-pollinated progeny of these Eurasian poplar species.

### R-genes and sex determination

Chromosome XIX is a special genome region in *Populus*. Recent studies of this chromosome have revealed different aspects of its complex history: different genomic architectures of sex determination depending on species (Geraldes et al. 2015; Kersten et al. 2014) but also a chromosome that is younger than the rest of the *Populus* genome with turnover events in the recent past (Geraldes et al. 2015; Hou et al. 2015). We discuss the potential roles of both R-genes and sex determination in shaping the genomic architecture of this chromosome.

Although genetic maps show that sex determination in European *Populus* species is located further towards the centromere (Paolucci et al. 2010; Tuskan et al. 2012; Kersten et al. 2014) and ongoing genomic work in *P. trichocarpa* continues to discover novel aspects of the sex determination system (Hou et al. 2015; Geraldes et al. 2015), this chromosome block was shown to present segregation distortion and reduced recombination in *P. alba* and *P. tremula*, as expected for an incipient sex chromosome (Macaya-Sanz et al. 2011). The presence of a large NBS-LRR gene cluster may explain the evolution of this particular chromosome. Suppression or reduction of recombination is not a phenomenon unique to sex chromosomes but is widespread in the genome of most organisms (Charlesworth 2013). Resistance genes have been shown to play a role in recombination, with some R-gene clusters showing complete suppression of recombination and others acting as hotspots of recombination (Bresson et al. 2011). In particular, suppression of recombination was demonstrated not only in the MXC3 locus region in *Populus* (Stirling et al. 2001; Yin et al. 2004) but also in R-gene loci in barley, tomato, and *Arabidopsis* (Yin et al. 2008).

To understand the role of R-gene clusters on chromosome XIX, we compared similarities between the two NBS-LRR clusters and their putative influence on the sex determination locus positioned on the same chromosome. Compared to the R-gene cluster on chromosome III, the super cluster on chromosome XIX presents levels of genetic variability that are comparable to other chromosomes (Fig. 2). This difference might be explained by balancing selection on chromosome III, whereas both balancing and divergent selection were detected on chromosome XIX. Indications for divergent selection in BayeScan analyses of both species (Table 4) accumulate especially in the chromosome part closer to the sex determination region of European species (Gaudet et al. 2008; Yin et al. 2008; Paolucci et al. 2010; Pakull et al. 2011; Tuskan et al. 2012). These patterns support the hypothesis that the evolution of NBS-LRR genes on chromosome XIX is, directly or indirectly, linked to the evolution of sex determination.

In addition to the R-gene super cluster, chromosome XIX presents a remarkable density of micro RNA's targeting and

potentially regulating the expression of NBS-LRR genes (Klevebring et al. 2009). The hypothesis that resistance and sex determination genes co-evolved in this chromosome region, thus permitting the emergence of a nascent sex chromosome, was developed further by Tuskan et al. (2012). These authors proposed that this genome region emerged under selective pressure for resistance to a female-specific floral pathogen. Suppressed recombination associated with resistance genes apparently allowed this genome region to be inherited as a largely cohesive ancestry block. Dioecy in *Populus* may result from a translocation of sex determination genes close to the R-gene super cluster during the more recent history of this chromosome. Our results thus support a role of resistance genes in suppressed recombination and selective pressures observed in this genome region. Indeed, a recent genome-wide association study of gender in *P. trichocarpa* and *Populus balsamifera* indicates a potential link between gender and defense response genes on chromosome XIX (Gerald et al. 2015). However, it remains to be tested whether the presence of micro RNAs and expression of NBS-LRR genes are gender specific, which would further support the hypothesis that resistance genes and sex determination have co-evolved.

## Conclusion and outlook

Our results are consistent with a role of R-genes in the evolution of chromosome XIX of *Populus*. However, contrary to expectations for both sex determination genes and R-genes from other species (Bomblies and Weigel 2007; Qvarnström and Bailey 2009; Ellegren 2009), chromosome XIX of European poplar species exhibits unusually low levels of species differentiation (Macaya-Sanz et al. 2011; Stölting et al. 2013; this study). More extensive work based on high-throughput genotyping-by-sequencing and whole genome resequencing is currently underway to clarify the precise roles of recurrent gene flow and shared selection pressures in generating the conspicuous patterns of allele sharing and LD observed in this enigmatic genome region of poplar. The relative roles of recurrent gene flow and ancestral polymorphism may be resolved by exploring genomic patterns of diversity in a biogeographically explicit context (Muir and Schlötterer 2005). Different types of shared selection pressures may be identified by genome-wide screens for balancing selection and based on characteristic fine-scale signatures of genomic divergence along chromosomes (Roesti et al. 2014). Thus, we anticipate that these important open questions can be resolved successfully by analyzing genome-scale data within a biogeographic context, making use of the excellent genomic resources and extensive natural populations available in *Populus* spp.

**Acknowledgments** This study was supported by grants 31003A\_127059 and 31003A\_149306 of the Swiss National Science Foundation (SNF) to CL.

**Data archiving statement** Microsatellite genotype data can be found on dryad doi:10.5061/dryad.2593t.

## References

- Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJ, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R et al (2013) Hybridization and speciation. *J Evol Biol* 26:229–246
- Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinf* 9:323
- Arnold ML (2006) Evolution through genetic exchange. Oxford University Press, Oxford
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Philos Trans R Soc Lond B Biol Sci* 263:1619–1626
- Bergelson J (2001) Evolutionary dynamics of plant R-genes. *Science* 292:2281–2285
- Bomblies K (2009) Too much of a good thing? Hybrid necrosis as a by-product of plant immune system diversification. *Botany* 87:1013–1022
- Bomblies K, Weigel D (2007) Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nat Rev Genet* 8: 382–393
- Bresson A, Jorge V, Dowkiw A, Guerin V, Bourgaït I, Tuskan GA, Schmutz J, Chalhoub B, Bastien C, Faivre Rampant P (2011) Qualitative and quantitative resistances to leaf rust finely mapped within two nucleotide-binding site leucine-rich repeat (NBS-LRR)-rich genomic regions of chromosome 19 in poplar. *New Phytol* 192:151–163
- Brown JKM, Tellier A (2011) Plant-parasite coevolution: bridging the gap between genetics and ecology. *Annu Rev Phytopathol* 49: 345–367
- Charlesworth D (2013) Plant sex chromosome evolution. *J Exp Bot* 64: 405–420
- De Meaux J, Mitchell-Olds T (2003) Evolution of plant resistance at the molecular level: ecological context of species interactions. *Heredity* 91:345–352
- Ellegren H (2009) The different levels of genetic diversity in sex chromosomes and autosomes. *Trends Genet* 25:278–284
- Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends Genet* 28:342–350
- 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
- Fussi B, Lexer C, Heinze B (2010) Phylogeography of *Populus alba* (L.) and *Populus tremula* (L.) in Central Europe: secondary contact and hybridisation during recolonisation from disconnected refugia. *Tree Genet Genomes* 6:439–450
- Garnier-Gere P, Dillmann C (1992) A computer program for testing pairwise linkage disequilibria in subdivided populations. *J Hered* 83:239
- Gaudet M, Jorge V, Paolucci I, Beritognolo I, Mugnozza GS, Sabatti M (2008) Genetic linkage maps of *Populus nigra* L. including AFLPs, SSRs, SNPs, and sex trait. *Tree Genet Genomes* 4:25–36
- Gerald A, Hefer CA, Capron A, Kolosova N, Martinez-Nuñez F, Soolanayakanahally RY, Stanton B, Guy RD, Mansfield SD, Douglas CJ, Cronk QCB (2015) Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). *Mol Ecol*. doi: 10.1111/mec.13126
- Germain H, Séguin A (2010) Innate immunity: has poplar made its BED? *New Phytol* 189:678–687

- Gompert Z, Buerkle CA (2010) INTROGRESS: a software package for mapping components of isolation in hybrids. *Mol Ecol Resour* 10: 378–384
- Gompert Z, Buerkle CA (2012) bgc: software for Bayesian estimation of genomic clines. *Mol Ecol Resour* 12:1168–1176
- Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *J Hered* 86:485–486
- Hou J, Ye N, Zhang D, Chen Y, Fang L, Dai X, Yin T (2015) Different autosomes evolved into sex chromosomes in the sister genera of *Salix* and *Populus*. *Sci Rep* 5:9076
- Jorge V, Dowkiw A, Faivre-Rampant P, Bastien C (2005) Genetic architecture of qualitative and quantitative *Melampsora larici-populina* leaf rust resistance in hybrid poplar: genetic mapping and QTL detection. *New Phytol* 167:113–127
- Keith R, Olds TM (2013) Genetic variation for resistance to herbivores and plant pathogens: hypotheses, mechanisms and evolutionary implications. *Plant Pathol* 62:122–132
- Kersten B, Pakull B, Groppe K, Lueneburg J, Fladung M (2014) The sex-linked region in *Populus tremuloides* Turesson 141 corresponds to a pericentromeric region of about two million base pairs on *P. trichocarpa* chromosome 19. *Plant Biol* 16:411–418
- Klevebring D, Street NR, Fahlgren N, Kasschau KD, Carrington JC, Lundeberg J, Jansson S (2009) Genome-wide profiling of *Populus* small RNAs. *BMC Genomics* 10:620
- Kohler A, Rinaldi C, Duplessis S, Bauchner M, Geelen D, Duchaussoy F, Meyers B, Boerjan W, Martin F (2008) Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. *Plant Mol Biol* 66:619–636
- Koopman WJM, Li Y, Coart E, van de Weg WE, Vosman B, Roldan-Ruiz I, Smulders MJM (2006) Linked vs. unlinked markers: multilocus microsatellite haplotype-sharing as a tool to estimate gene flow and introgression. *Mol Ecol* 15:243–256
- Lexer C, Widmer A (2008) The genic view of plant speciation: recent progress and emerging questions. *Philos Trans R Soc B* 363:3023–3036
- Lexer C, Fay MF, Joseph J, Nica M, Heinze B (2005) Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): the role of ecology and life history in gene introgression. *Mol Ecol* 14:1045–1057
- Lexer C, Joseph J, van Loo M, Barbará T (2010) Genomic admixture analysis in European *Populus* spp. reveals unexpected patterns of reproductive isolation and mating. *Genetics* 186:699–712
- Lindtke D, Buerkle CA, Barbará T, Heinze B, Castiglione S, Bartha D, Lexer C (2012) Recombinant hybrids retain heterozygosity at many loci: new insights into the genomics of reproductive isolation in *Populus*. *Mol Ecol* 21:5042–5058
- Lindtke D, Gompert Z, Lexer C, Buerkle CA (2014) Unexpected ancestry of *Populus* seedlings from a hybrid zone implies a large role for postzygotic selection in the maintenance of species. *Mol Ecol* 23: 4316–4330
- Macaya-Sanz D, Suter L, Joseph J, Barbará T, Alba N, Gonzalez-Martinez SC, Widmer A, Lexer C (2011) Genetic analysis of post-mating reproductive barriers in hybridizing European *Populus* species. *Heredity* 107:478–486
- Muir G, Schlötterer C (2005) Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Mol Ecol* 14:549–561
- Pakull B, Groppe K, Mecucci F, Gaudet M, Sabatti M, Fladung M (2011) Genetic mapping of linkage group XIX and identification of sex-linked SSR markers in a *Populus tremula* × *Populus tremuloides* cross. *Can J Forest Res* 41:245–253
- Paolucci I, Gaudet M, Jorge V, Beritognolo I, Terzoli S, Kuzminsky E, Muleo R, Mugnozza GS, Sabatti M (2010) Genetic linkage maps of *Populus alba* L. and comparative mapping analysis of sex determination across *Populus* species. *Tree Genet Genomes* 4:25–36
- Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends Ecol Evol* 24:386–393
- Petre B, Major I, Rouhier N, Duplessis S (2011) Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar. *BMC Plant Biol* 11:33
- Qvarnström A, Bailey RI (2009) Speciation through evolution of sex-linked genes. *Heredity* 102:4–15
- Rausher MD (2001) Co-evolution and plant resistance to natural enemies. *Nature* 411:857–864
- Roesti M, Gavrillets S, Hendry AP, Salzburger W, Berner D (2014) The genomic signature of parallel adaptation from shared genetic variation. *Mol Ecol* 23:3944–3956
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18:233–234
- Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers-Melnick E, Lipphardt MF, Pennacchio CP, Hellsten U, Pennacchio LA, Gunter LE et al (2012) Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytol* 196:713–725
- Smadja CM, Butlin RK (2011) A framework for comparing processes of speciation in the presence of gene flow. *Mol Ecol* 20:5123–5140
- Stahl EA, Bishop JG (2000) Plant-pathogen arms races at the molecular level. *Curr Opin Plant Biol* 3:299–304
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989
- Stirling B, Newcombe G, Vrebalov J, Bosdet I, Bradshaw HD Jr (2001) Suppressed recombination around the MXC3 locus, a major gene for resistance to poplar leaf rust. *TAG* 103:1129–1137
- Stölting KN, Nipper R, Lindtke D, Caseys C, Waeber S, Castiglione S, Lexer C (2013) Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. *Mol Ecol* 22:842–855
- Stölting KN, Paris M, Meier C, Heinze B, Castiglione S, Bartha D, Lexer C (2015) Genome-wide patterns of differentiation and spatially varying selection between postglacial recolonization lineages of *Populus alba* (Salicaceae), a widespread forest tree. *New Phytol.* doi:10.1111/nph.13392
- The Marie Curie SPECIATION Network (2011) What do we need to know about speciation? *Trends Ecol Evol* 27:27–39
- Traw M, Bergelson J (2010) Plant immune system incompatibility and the distribution of enemies in natural hybrid zones. *Curr Opin Plant Biol* 13:466–471
- Tuskan GA, DiFazio SP, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam NH, Ralph SG, Rombauts S, Salamov A et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray). *Science* 313:1596–1604
- Tuskan GA, DiFazio S, Faivre Rampant P, Gaudet M, Harfouche A, Jorge V, Labbé JL, Ranjan P, Sabatti M, Slavov GT, Street N, Tschaplinski TJ, Yin TM (2012) The obscure events contributing to the evolution of an incipient sex chromosome in *Populus*: a retrospective working hypothesis. *Tree Genet Genomes* 8:559–571
- Whitney KD, Randell RA, Rieseberg LH (2006) Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *Am Nat* 167:794–807
- Wu C (2001) The genic view of the process of speciation. *J Evol Biol* 14: 851–865
- Yin T, DiFazio SP, Gunter L, Jawdy S, Boerjan W, Tuskan GA (2004) Genetic and physical mapping of *Melampsora* rust resistance genes in *Populus* and characterization of linkage disequilibrium and flanking genomic sequence. *New Phytol* 164:95–105
- Yin T, DiFazio SP, Gunter LE (2008) Genome structure and emerging evidence of an incipient sex. *Genome Res* 18:422–430
- Zaykin DV, Pudovkin A, Weir BS (2008) Correlation-based inference for linkage disequilibrium with multiple alleles. *Genetics* 180:533–545