MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Exploring the genetic component of ecologically important traits in hybridizing *Populus* species "

verfasst von / submitted by

Erik Szamosvári, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2019 / Vienna, 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 833
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Master's degree programme Ecology and Ecosystems
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Christian Lexer

Table of Contents

Acknowledgements

Zusammenfassung

Natürliche Hybridpopulationen haben in den letzten Jahrzehnten und insbesondere seit Beginn der Genomik-Ära großes Interesse für Untersuchungen der Evolutionsbiologie und der Öko-Evolutionsdynamik gefunden. In diesem Projekt haben wir die Auswirkungen genetischer Abstammung auf juvenile Fitness-Merkmale und funktionell wichtige Pflanzenmerkmale in den zwei ökologisch unterschiedlichen Baumarten *Populus alba*, *P. tremula* und ihren diploiden *P. x canescens*-Hybriden in Anbauversuchen erforscht. Wir fanden heraus, dass Rückkreuzungen mit *P. tremula* und andere rekombinante Hybridgenotypen in den ersten drei Lebensjahren in zwei Testgärten stärker von jugendlicher Mortalität betroffen waren als F1-Hybride, Rückkreuzungen mit *P. alba* und beide Elternarten. Sowohl die Bereitschaft zur klonalen Vermehrung (Rate überlebender Stecklinge der Klone), als auch das Wachstum der Jungpflanzen waren stark von der genetischen Abstammung beeinflusst. Der klonale Erfolg sowie auch die relativen Wachstumsraten waren bei Jungpflanzen mit einem höheren Abstammungsanteil von *P. alba* größer als bei anderen Hybridtypen und Rückkreuzungstypen zwischen *P. alba* und *P. tremula*. Unsere Ergebnisse aus Testgärten helfen die genomische Zusammensetzung natürlicher Hybridpopulationen dieser beiden weit verbreiteten Waldbaumarten zu erklären, die schon in früheren Studien präsentiert wurden. Darüber hinaus haben unsere Messungen der Blattreflexion (zur Bestimmung der Trichomen-Dichte am Blatt) eine signifikante Variation dieses funktional wichtigen Pflanzenmerkmals entlang der genetischen Durchmischungsachse von *P. tremula* bis *P. alba* zutage gebracht; mit einem Fehlen von Trichomen an einem Ende und hoher Trichomen-Dichte am anderen. Diese Ergebnisse wurden von Botanikern und Pappelzüchtern seit langem vorhergesagt, und unsere Daten und Befunde schaffen die Voraussetzung für die genetische Kartierung dieses funktional wichtigen Pflanzenmerkmals in natürlichen Pappelhybriden in den Testgärten. Zusammenfassend hat sich diese Arbeit mit den Auswirkungen genetischer Abstammung (neben Umwelt- und Entwicklungsaspekten) auf mehrere juvenile Fitness-Merkmale und funktional wichtige Pflanzenmerkmale in zwei weit verbreiteten, hybridisierenden Waldbäumen beschäftigt. Diese Ergebnisse bilden die Grundlage für genetische

Untersuchungen (z. B. genomweites Assoziations-Scanning) von fitnessbezogenen Merkmalsunterschieden bei diesen ökologisch wichtigen, hybridisierenden Pappelarten.

Abstract

Natural populations of hybrids have attracted great interest for studies of evolutionary biology and eco-evolutionary dynamics in recent decades, and particularly so since the onset of the genomics era. In this project, we tested the effects of genomic ancestry on juvenile fitness proxies and functionally important plant traits in the ecologically divergent tree species *Populus alba*, *P. tremula* and their diploid *P. x canescens* hybrid offspring in a common garden setting. We found that backcrosses towards *P. tremula* and other recombinant hybrid genotypes suffered more from juvenile mortality in two common gardens in their first three years of life, compared to F_1 hybrids, backcrosses to *P. alba*, and both parental species. Both the propensity to reproduce clonally (measured via clonal cutting survival) and sapling growth were strongly affected by genetic ancestry: clonal success was greater in saplings with a higher proportion of *P. alba* ancestry, and the same was the case for relative growth rates when comparing backcrosses towards this species (BCA) with other types of hybrids along the axis of genome admixture. These results from experimental common gardens help explain the genomic composition of natural hybrid populations of these wide-spread forest tree species seen in previous studies. In addition, our measurements of leaf reflectance (a surrogate of leaf trichome density) recovered significant variation for this functionally important plant trait along the admixture axis spanning from *P. tremula* to *P. alba*, with an absence of trichomes in the former and dense trichome cover in the latter. These results were long predicted by botanists and poplar breeders, and our data and findings set the state for the admixture mapping of this functionally important plant trait in common garden grown poplar hybrids. In summary, this thesis has addressed the effects of genetic ancestry (along with environmental and developmental aspects) on several juvenile fitness proxies and functionally important plant traits in two wide-spread, hybridizing forest trees. The results lay the foundation for genomic investigations (e.g. genome wide association scanning) of fitness-related trait differences in these ecologically important, hybridizing poplar species.

Introduction

Hybrid zones have gained great interest in ecology and evolutionary biology in recent decades and are considered as 'natural laboratories' for studying the mechanism of population divergence, adaptation, reproductive isolation, and speciation (Hewitt 1988, Harrison 1990). Natural hybrid zones form when two genetically distinct taxa with incomplete reproductive isolation meet, reproduce, and these mating events results in offspring with mixed ancestry (Harrison 1993). As a result of hybridization events, the alleles (=genetic variants) of divergent populations or species interact with one another, and these interactions may have neutral, beneficial or deleterious effects (Abbott et al. 2013). The possible consequences of hybridization in the context of speciation are manifold. Hybridization can affect barriers to gene flow in a positive or negative way. Also, recombination and gene flow between taxa may either impede species differentiation or may even fuel speciation (Abbott et al. 2013).

Introgression (=continued hybridization) results in recombination of chromosomes or genes from divergent taxa, which can lead to the emergence of novel phenotypes or species. Hybrid zones occur where the habitats of divergent lineages overlap, thus providing opportunity for genetic contact. The potential results of such reproduction events vary, e.g. depending on the parental species or the environment (Abbott 2017). In some cases, a broad spectrum of recombinant hybrids is present in hybrid zones, while in others only F_1 hybrids and parental genotypes are common. In other cases, hybrid progeny are able to colonize new habitats not accessible to the parental taxa, thus leading to the formation of new 'hybrid species' even without any change in ploidy (e.g. *Helianthus* spp.; Lexer et al. 2003) (Harrison 1993).

Novel techniques in genetics and genomics provide excellent opportunities to study the mechanisms underlying the origin and maintenance of reproductive isolation and trait differences, and thus of biodiversity (Buerkle and Lexer 2008, Lindtke et al. 2013). For example, the use of genome wide association studies in hybridizing (=genetically admixed) lineages, also known as 'admixture mapping', can help evolutionary biologists and ecological geneticists to explore relationships between phenotypic and genotypic differences among

divergent populations or species, i.e. the genetic architecture of important trait differences (Baack and Rieseberg 2007; Buerkle and Lexer 2008). Natural populations of hybridizing taxa can form 'genomic mosaics' with a broad range of genotypes and ancestry segments along the genome (Lindtke et al. 2014, Caseys et al. 2015). This great variety of genotypes can enrich the diversity of phenotypes in populations and communities (Whitham et al. 2006; Lexer et al. 2009) and make possible the investigation of unanswered questions in ecology and evolutionary biology.

In this study we used two 'model forest tree' species with porous genomes (Lexer et al. 2007), *Populus alba* and *P. tremula*, and their hybrids (a.k.a. *P. x canescens*) to examine the connections between genetic ancestry and fitness-related and/or functionally important traits. The two parental taxa differ in ecology, morphology and neutral DNA-based markers (Lexer et al. 2007, Lexer et al. 2010) and frequently form 'mosaic' hybrid zones in nature. Seeds from open-pollinated trees were collected in a natural hybrid zone of the Ticino river valley at the Swiss-Italian border (Lindtke at al. 2014). Seeds were germinated and common garden trials were established in Fribourg, Switzerland, and Salerno, Italy, by Christian Lexer and his team and collaborators. The biological materials for the present study were transferred from Fribourg to Vienna, Austria, in 2016 in the form of root cuttings and shoot cuttings ("stecklings"). In 2017, a common garden trial composed of 200 saplings was established in Vienna from these materials, and phenotypic measurements were conducted over the following two years. Genetic ancestry parameters previously estimated by Lindtke et al. (2014) and Bresadola et al. (2018) based on two different genotyping-by-sequencing (GBS) approaches were available to examine associations between genetic ancestry and fitness-related plant traits.

The main focus of this thesis was on analyzing the effects of genetic ancestry on a range of juvenile fitness proxies, including seedling survivorship, the propensity to reproduce asexually (=survival of clonal cuttings), and seedling growth (height and diameter increment). In addition, we were interested in examining leaf trichome density (hairiness of the abaxial leaf surface) using a digital morphometric approach (Lexer et al. 2009), because of reported links between this functionally important plant trait and both herbivory

response and transpiration (Plett et al. 2010). Specifically, this study asked how the genetic ancestry of common garden-grown poplar trees might affect their (1) survivorship, (2) propensity to clone, (3) growth patterns, and (4) variation in a functionally important leaf trait, trichome density.

Materials and methods
Study system
Two tree species from the diploid genus *Populus* of family *Salicaceae*, *Populus alba* (white poplar), *Populus tremula* (European aspen), and their hybrids (*P. x canescens*) comprised the study system of our research. *Populus alba* is a deciduous pioneer tree of Eurasian lowland floodplains and other habitats with severe abiotic stress regimes (flooding, drought, salinity), dominant in many riparian woodlands. It is widely distributed in Eurasia, native from northern Africa, the Iberian-peninsula along the Mediterranean, Central and Western Europe to Central Asia and Western Siberia (Dickmann and Kuzovkina 2008). The other parental species is *P. tremula*, known as a deciduous upland pioneer of Eurasia with an even wider distribution. The range of the species extends from cooler, boreal regions of Europe and Asia, to Spain and North Africa, from Ireland to Kamchatka and the central Island of Japan (Dickmann and Kuzovkina 2008). *Populus tremula* is the second most widely distributed tree in the world after *Pinus sylvestris* (EUFORGEN. 2009, Caudullo and de Rigo 2016b). *Populus alba* and *P. tremula* are phenotypically and genetically divergent species. They not only differ in habitat and morphology (Lexer et al. 2009, Lindtke et al. 2013), but also show variation in phytochemical traits (Casey et al. 2012, 2015). Both parental species, considered as 'model forest trees', play important roles in ecology such as creating habitats, stepping stones and in ecosystem services including watershed protection and soil stabilization (Dickmann and Kuzovkina 2008, Caudullo and de Rigo 2016a, 2016b). In spite of distinct habitat types, their populations are overlapping geographically. In fact, in many cases the two species occur in sym- or parapatry (e.g. along the river systems of Europe), creating mosaic hybrid zones (Van Loo et al. 2008, Lexer et al. 2005, 2010). Hybrids (*P. x canescens*) can be found in these zones as a result of mating via wind pollination, thus indicating that reproductive isolation between these genetically divergent taxa is incomplete (Lexer et al. 2010, Lindtke et al. 2012, 2014). Fertile and viable first (F_1) and subsequent

generation hybrids (F_n) meet and mate in these contact zones not only with hybrids, but also with the parental taxa, thus giving rise to backcrosses and other types of recombinant hybrid genotypes (Lexer et al. 2010, Lindtke at al. 2014; Christe et al. 2016). The presence of ecologically divergent taxa with "porous" genomes and the availability of reference genomes (e.g. *P. trichocarpa*; Tuskan at al., 2006) provide an opportunity with huge potential to study speciation, adaptation and the genetic basis of ecologically important traits (e.g.: Buerkle & Lexer 2008, Lexer et al. 2009, Carvalho et al. 2010, Lindtke et al. 2013).

Common garden in Vienna

This study was based on a common garden (CG) trial established at the Botanical Garden of University of Vienna. The biological materials used in the CG stem from a natural hybrid zone in the Ticino river valley at the Swiss-Italian border. Propagules from catkins of maternal trees were collected (Lindtke at al. 2014) and seeds were germinated in Fribourg, Switzerland. Trees, propagated from root (n=185) and shoot (n=15) cuttings, were used to establish the CG at the Vienna location. A randomized block design was used for setting up the CG of 200 seedlings in the spring of 2017. To render the sampling homogenous, 5 planting blocks were created, consisting of identical numbers of genotypes in each block: 2 *P. tremula*, 34 *P. x canescens* and 4 *P. alba*. In each block, 40 trees were placed randomly with approximately 1 meter distance between the trees and 2 meters between the blocks (see Figure 1). To maintain homogeneous environmental conditions, an automated watering system was installed in early 2017. Apart from substrate containing humus, seedlings were grown without any interference from our side. Additional phenotypic data for a related common garden trial established from the same seedling families near Salerno, Italy (not published) and Fribourg, Switzerland (Caseys et al. 2015) were also analysed in the present study.
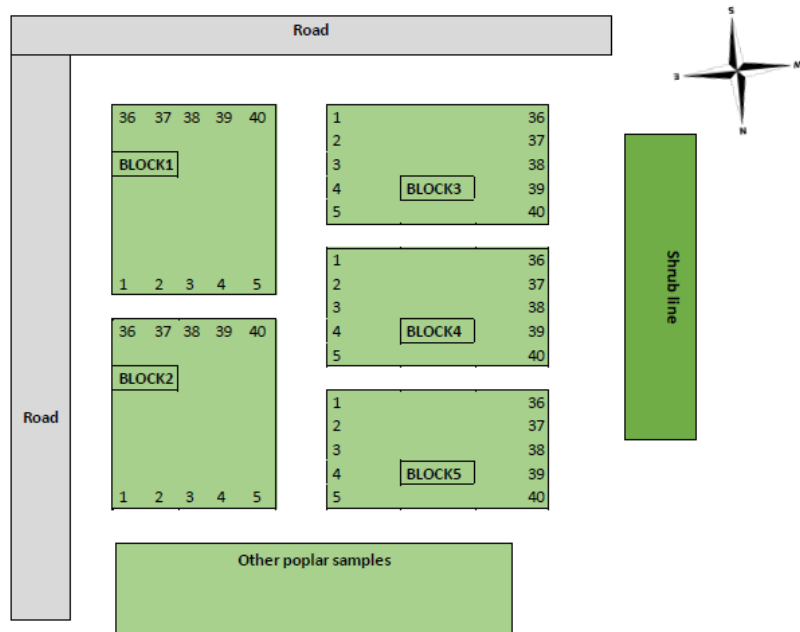
Figure 1. The block positions of the common garden In Vienna

Genetic parameters

Admixture analysis of hybrids in natural contact zones represents an opportunity to examine the genetic basis of phenotypic variation in F1 and recombinant hybrids among divergent parental species (Buerkle & Lexer 2008, Lindtke et al. 2013). In our study, genetic ancestry parameters for individual trees were provided and calculated by Bresadola et al. (2018). These parameters were based on restriction site associated DNA (RAD) sequencing data and Bayesian estimates of genetic ancestry following Christe et al. (2016). We additionally used genetic information evaluated from genotyping by sequencing (GBS) data by Lindtke et al. (2014). These datasets include estimate of genome-wide admixture proportions. The values of the hybrid or admixture index (q) can be interpreted in the following way: if $q \leq 0.1$, the individual belongs to *P. tremula*, if $q \geq 0.9$, the individual belongs to *P. alba*, if $0.1 < q < 0.9$, the individual is classified as a hybrid, *P. x canescens* (Lindtke et al. 2012). A second genetic ancestry parameter, called inter-source ancestry index (Q_{12}), was also estimated from the same data, following Lindtke at al. (2012) and Lindtke et al. (2014). Q_{12} is the proportion of the loci that are of mixed ancestry between the parental species, in analogy to heterozygosity in classical population genetics. The inter-source ancestry, which also ranges between 0.00 and 1.00, indicates if the individual is more likely to belong to the first generation of hybrids (F$_1$) or if it is more likely to be of a recombinant hybrid generation

($F_{2,3,....n}$). If the $1 \geq Q_{12} \geq 0.75$, the individual can tentatively be classified as an F_1. If the Q_{12} value is lower, the individual is expected to belong to a subsequent, recombinant generation (e.g. parental-like backcrosses or F_n), due to the decrease of interspecific heterozygosity with each generation of recombination (Buerkle & Lexer 2008). Figure 2 shows the genetic ancestry parameters of the trees used in the experiment. Many individuals are genetically closer to *P. alba,* which reflects genotype proportions of these plants in natural hybrid zones, presumably due to the low abundance of *P. tremula* in riparian forest habitats (Lindtke et al. 2014, Lexer et al. 2010, Christe et al. 2016).
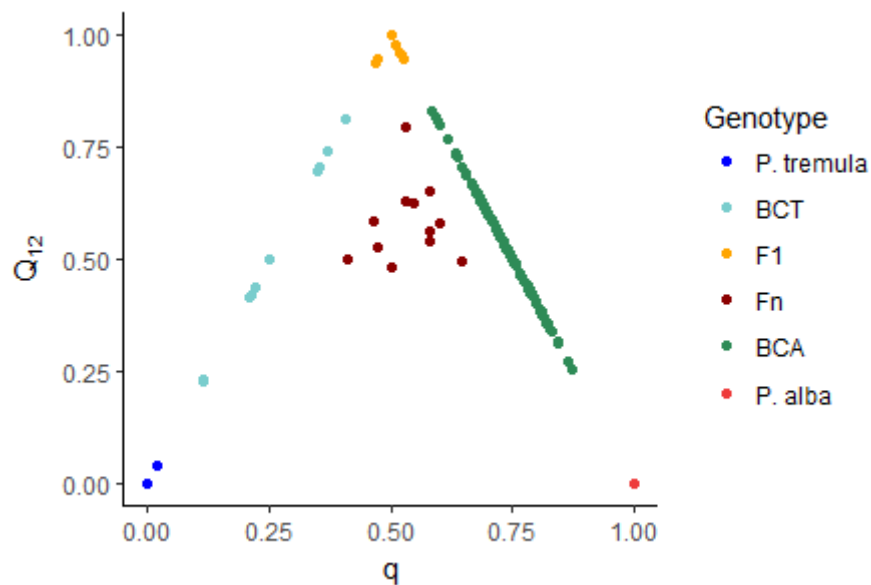

Figure 2 The plot of genetic parameters with q and Q12 as explained in text. Genetic ancestry parameters were estimated by Bresadola et al. (2018), following Lindtke et al. (2014) and Christe et al. (2016). (Abbreviations: BCT: backcrosses towards *Populus tremula*, Fn: individuals belong to subsequent generations, BCA: backcrosses towards *P. alba*)

Seedling survivorship

The wind-pollinated parental species reproduce both sexually and asexually. Seed production is abundant, but seedling mortality is high for these pioneer species, because necessary conditions for germination are often lacking (e.g.: moist soil is required) (Dickmann and Kuzovkina 2008). This phenomenon is known in other *Populus* species as well (Schweitzer et al. 2002). Our aim was to examine the relationship between genetic ancestry (q, Q_{12}) and seedling survivorship of parental taxa and their hybrids, as a juvenile fitness proxy, in the studied common garden trials. Previous experiments from Salerno, Italy and Fribourg, Switzerland contributed to the data available for statistical analysis. For analysing

relationships between survivorship and ancestry, we used previously available ancestry estimates based on genotyping-by-sequencing (GBS) (Lindtke et al 2014; Christe et al. 2016), rather than RAD-seq genotype data. The use of GBS data allowed us to directly compare our results on survivorship at different time point in two CG locations to results obtained for a single CG location in a previous study (Christe et al. 2016).

Clonability

Asexual or clonal reproduction is common in the genus *Populus*. This phenomenon is well-known in *P. alba*, *P. tremula* and in their hybrids, but their propensity to clone differs dramatically. Asexual reproduction contributes to the maintenance of populations and genotypes in time and space in these frequently hybridizing species (Schwitzer at el. 2002, Van Loo et al. 2008). In spite of high number of viable seeds, *P. tremula* reproduces asexually via root suckers as well (Latva-Karjanmaa et al. 2006). *Populus alba* is also capable to spread vegetatively (Koop 1987, Dickmann and Kuzovkina 2008). Hybrids (*P. x canescens*) can even produce more ramets and conquer larger areas in natural hybrid zones than parental taxa (Van Loo et al. 2008). The ability or potential to form clones, referred to as clonability here, is an asexual fitness proxy, which is an important means of propagation in nature (Van Loo et al. 2008). In this study we tested genetic effects on clonability of shoot cuttings to obtain information on the propensity of individual genotypes to propagate asexually in a common garden trial. According to Dickmann and Kuzovkina (2008), *P. tremula* can be rooted from softwood shoot cuttings under suitable conditions. Cuttings of *P. alba* can be propagated even from hardwood, but success varies among genotypes. Several shoot cuttings from the same trees with known genetic parameters were planted in early spring of 2016 in the Botanical Garden of University of Vienna. Photographic documentation was used for scoring the strength of starting material (diameter covariate). Cuttings were grouped in 5 categories according to the thickness of shoots. This covariate was included in a dataset with total number of cuttings and the number of cuttings still alive at beginning of July in 2016. From this information, so called "cutting success" was calculated for each tree in the following way: number of cuttings alive / total number of cuttings. After removing genotypes with less than 5 cuttings planted, altogether 231 statistical units were used during analysis of clonability.

Growth

The two parental species of *Populus* are considered as fast-growing trees (Caudullo and de Rigo 2016a, 2016b). *Populus tremula* is a short-lived, medium-sized tree, which shows increased and quick growth in its early lifespan, but with canopy competition this rate slows down later on (Dickmann and Kuzovkina 2008, Caudullo and de Rigo 2016b). *Populus alba* is considered as a medium to large tree with moderately long lifetime, and this species also has similarly rapid growth rates. (Dickmann and Kuzovkina 2008, Caudullo and de Rigo 2016a). Growth during early life stage in natural habitats can be influenced by environmental stress (e.g.: flooding, fire, drought) and competition for light and other resources (Stettler et al. 1996). Despite these common forces, there are several others with great regulatory impact. In our study, we explored growth rate and height variation of young saplings from different genotypes in a CG trial and potential effects of genetic ancestry parameters (estimated from RAD-seq) on this juvenile fitness component. Only trees germinated from root cuttings were included in the analysis, as suitable sample sizes were available for these. During the second year of the CG trial, from the beginning of the vegetation period until autumn of 2018, measurements (n=6) of diameter and of the height of each tree were taken in 5-week intervals. Diameter was measured with a digital caliper between the north and south side of the trunks at 8 cm from the ground in the pots, while height was measured as distance from the ground to the top of the leading shoot. Relative growth rate was calculated for each tree following a common equation (Fisher 1921) (Figure 3). Relative growth rate (RGR) can facilitate more meaningful comparisons than absolute growth rate (Hunt 1990).

$$RGR = \frac{(\log_e(measurement_2) - \log_e(measurement_1)}{time_2 - time_1}$$

Figure 3. Equation of relative growth rate.

Leaf reflectance

Admixture analysis provides an opportunity to examine phenotypic variation in taxa with "porous genomes" in recently diverged species (Buerkle and Lexer 2008, Lindtke et al. 2013). *Populus alba* and *P. tremula* meet and mate in natural contact zones, where fertile hybrids also breed with each other and their parental taxa. Their offspring form genomic mosaics. Populations contain different generations of recombinants and show high variability of

admixture proportions between the parental species (Lindtke et al. 2014, Caseys et al. 2015). This genomic mosaic also suggests the presence of phenotypic mosaics of hybrids due to the genetic recombination between the parental gene pools and the introgression of traits and alleles via backcrosses (Lexer et al. 2009). Nevertheless, since F_1 and genetically intermediate F_n hybrids are also present in hybrid zones of these species, phenotypic traits are often intermediate (Lindtke et al. 2013). The two species differ in leaf morphology, including not only leaf shape, leaf area and petiole length, but also hairiness (trichome density) of the leaf undersides (Bartha 2004, Lexer et al. 2009). *Populus alba* leaves exhibit a dense indumentum of hairs on their leaf undersides, while the abaxial leaf surface of *P. tremula* is glabrous (Lexer et al. 2009). Morphological diversity and previously estimated genomic parameters provide the chance to explore the relationships between heritable phenotypic trait differences and genetic ancestry. We used digital image-based morphometrics to study functional trait variation in these hybridizing model forest tree taxa. Our study is focussed on leaf reflectance, which is known to be correlated with trichome density (Plett et al. 2010). Trichomes on plants are important modified epidermal outgrowths of several different shapes and functions (Plett et al. 2010). Samples (n=574) were collected from 198 trees during mid-summer of 2017 in the CG at the Botanical Garden of University of Vienna. From each tree, three fully expanded leaves of the leading shoot without any trace of herbivory were carefully taken. Samples were immediately placed in a standard herbarium press after labelling. The collected materials were dried for 14 days in a herbarium drying cabinet. Temperature was kept at 32 Celsius with continuous air circulation to ensure rapid drying. All samples were digitized and saved in Tiff image format using a Canon 9000F Mark II type scanner. With the help of a Kodak CAT 1527654 colour standard, black and white points were set manually in each case to keep standard threshold exposure values, following Lexer et al. (2009). The lower (abaxial) surface of leaves were scanned in random order to avoid human or technical bias. We always used the same settings: automatic sharpness, in colour mode, 300 dpi resolution and black background. 8 bit images of scanned specimens were analysed in Image J open source software (https://imagej.nih.gov/ij/). We measured grey values of a set area (circle-shape) between primary and secondary veins next to the main vein. This grey value of lower leaf surface provides information on leaf hairiness (Lexer et al.

2009) and thus trichome density (Plett et al. 2010). Testing the effects of genetic ancestry on leaf reflectance / trichome density can provide a first glimpse of the degree of genetic determination of this functional trait. This can also be of relevance for the fitness of hybrids, because trichome density plays an important role in pest resistance, reducing transpiration, and photosynthesis, and it may indirectly impact on growth rates (Plett et al. 2010). During the analysis, mean grey values of leaves originating from the same individual were merged, and average values were used in down-stream statistical analyses.

Statistical analysis

The open source software environment R (R Development Core Team, 2008) was used for statistical computing and plotting. The significance of p-values was set at a minimum α=0.05 and all assumptions were tested prior to running and interpreting statistical models.

Examining the effects of genetic ancestry on seedling survivorship required binary logistic regression, because of the binary nature of our survivorship response variable (seedling dead or alive at a particular time point). Our dataset contained 255 observations for year 3 after planting. As predictor variables, we used both linear and quadratic values of the hybrid or admixture index (q) and of inter-source ancestry index (Q_{12}). Generalized linear mixed effects models (GLMM) in the *lme4* package (Bates et al. 2014) were used to analyse third year survivorship. The location of common gardens (Salerno or Fribourg) was introduced into models as a fixed factor variable and maternal family (classification according to mother trees, n=15) was introduced as a random factor. We created different models from the simplest to the most complex ones. We applied this method also to all other models described below. For GLMM we ran analysis of variance (ANOVA) and chose the best fitting model according to the Akaike Information Criterion (AIC). Significance of coefficients with 95% intervals were determined for fixed effects of the most accurate model to interpret which predictor variable contributed most to predicting the response.

To interpret relationships between clonability and genetic parameters we again ran GLMM using the *lme4* package (Bates et al. 2014). This was necessary because the response variable was calculated from binary data (clonal cuttings alive or dead), categorized as a percentage, and germination year of propagation source was applied as random factor in the model. First

analysis of covariance (ANCOVA) was used to test which fixed predictor (over the genetic indices) we could insert in the model as a covariate or as a random factor, choosing between diameter classes of shoot cuttings and germination year of propagation source. In fact, if these variables vary between groups (here: genotype) then they could remove information from explanatory variables (here: q and Q_{12}). Our dataset consisted of 230 sample units and the final models included admixture index (q), inter-source ancestry index (Q_{12}) and scored diameter classes as fixed predictor variables. Germination year of propagation source was considered as a random factor, and cutting success as response variable. Subsequently, model comparisons and calculation of significance of model coefficients were carried out as described above.

To analyse how genetic ancestry can affect growth in our common garden trial, we used linear mixed effect models (LMEM) within the *nlme* R-package (Pinheiro et al. 2018). Our dataset comprised measurements for 181 trees, which were used to test four kind of approaches with three different response variables in our models. Spearman's rank correlation confirmed the close relationship between growth rate calculated from diameter and height, respectively, and all subsequent analyses were carried out using height only. In the first approach, model comparison included height measurements of each month of all trees as a response variable. We applied the following fixed predictor variables: both genetic parameters, germination year of trees used as propagation source and linear, quadratic and cubic values of measurement time points (April was used as baseline, as the first measurement, n=6), to see how time affects height differences. Location in the CG (block number) was added to the models as a random factor. In our second approach, we replaced the response variable with RGR of each month, with height values from April as reference point during calculating RGR. Here, we applied the same fixed predictor variables and random factor as previously. The third approach was to use overall RGR from the growing season of 2018 as response variable. In these models we included the same random factor and fixed predictor variables as before, with one exception: we removed the values of the measurement time points (effect of time), because we investigated only one given period. In our last approach, we also analysed a second dataset of overall RGR where all trees

categorised as pure *P. alba* and *P. tremula* based on the genetic data were removed. This allowed us to explore the impact of genetic ancestry and recombination in hybrid genotypes only, excluding the reference trees of the pure parental species. During model comparisons, AIC values were extracted from ANOVA to obtain information about the best fitting models and significance of coefficients.

To explore the effect of genetics on leaf reflectance as a surrogate of trichome density, we used linear regression models (LM). As the hierarchical organization of trees within blocks did not affect our response variable, there was no need to include it as a random factor. The means of leaf reflectance (=grey values) across replicate leaves for each tree were used as response variables. After merging mean grey values of leaves from each tree, our dataset consisted of 198 statistical units. In our linear regression models, inter-source ancestry index (Q_{12}) and both the linear and quadratic values of admixture index (q) were set as predictor variables. To obtain the most accurate model, ANOVA was run and differences in Residual Sums of Squares (RSS) between models were evaluated via *F*-statistics. The significance of the coefficients and 95% intervals for the chosen predictors were also estimated.

Results
Seedling survivorship
To detect impacts of genetic indices on survivorship of young poplar saplings we used complementary ancestry estimates from a GBS study by Lindtke et al. (2014), and from a recent RAD-seq study by Christe et al. (2016). Our extended dataset included survivorship observations three year after planting in two different CG locations. In the third year, there were 236 live and 19 dead (8.05%) trees in the two CGs. We used family membership of trees as a random factor while building GLMMs for 3rd year survivorship. The best fitting model according to our ANOVA of model comparison (Table S1) contained linear and quadratic values of admixture index and linear values of inter-source ancestry (ANOVA of GLMM: ΔAIC= 21.95, *p*-value= <0.0001.)

Table 1. Fixed coefficients of the best fitted GLMM for the third year survivorship

	Estimate	Std..Error	z.value	Pr(>|z|)
Intercept	1.799	0.432	4.165	<0.0001
q	-31.334	6.372	-4.917	<0.0001
q^2	32.630	6.350	5.138	<0.0001
Q_{12}	16.777	3.825	4.386	<0.0001

The evaluation of fixed coefficients of the best fitting model showed that all predictors bring significant changes in survivorship of saplings (Table 1). When location of common garden was inserted as fixed predictor, it did not decrease the AIC value (Table S1). Therefore, there were no significant differences in proportion of survivorship between the CGs of Salerno and Fribourg. The positive effect of Q_{12} indicates higher chances of survivorship with increased Q_{12} values, which mean F_1s were more likely to survive than subsequent recombinants (Figure 4). Also the quadratic effect of q was positive, which reflects that individuals with low to intermediate admixture index, i.e. backcrosses to *P. tremula*, have a lower chance to survive (Figure 4). These tendencies are also illustrated in Figure 5. Our results suggest that recombinant hybrids with an affinity to *P. tremula* suffered most from mortality.
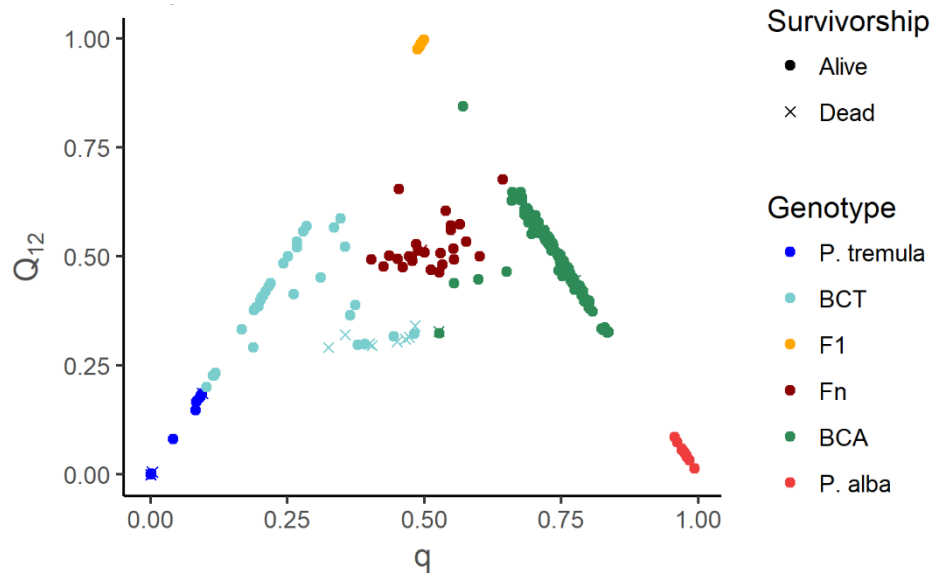

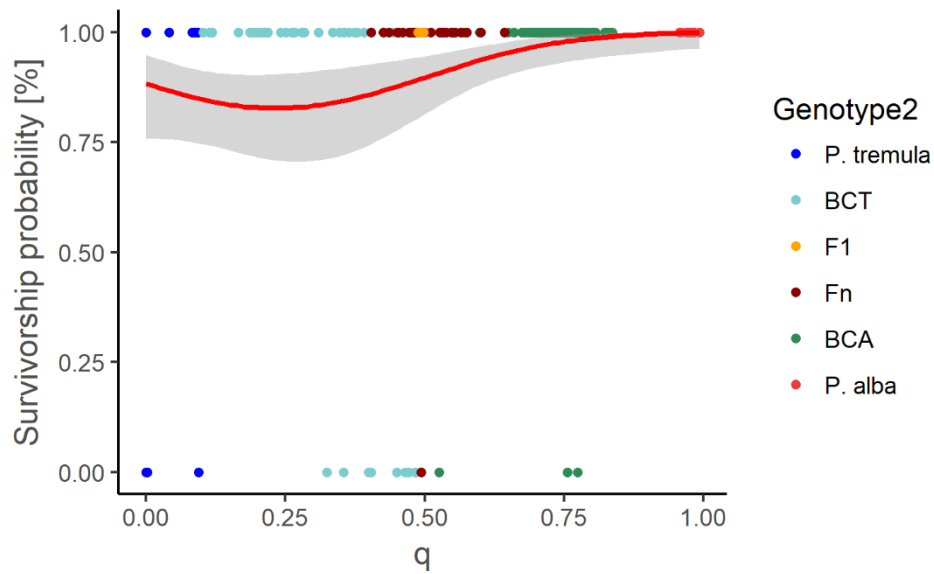*Figure 4. Triangle plot of survivorship in the third year according to q (x-axis) and Q12 (y-axis) values*

Figure 5. Polynomial logistic curve based on admixture index for best fitting models of GLMM for 3rd year survivorship in two CG locations.

Clonability

The statistical analyses showed that admixture index and germination year of propagation source affect the propensity for asexual reproduction via shoot cuttings in a common garden trial. To evaluate clonability, the cutting success for each statistical unit (n=231) was estimated (mentioned above) and applied as response variable in our GLMMs. First, we ran ANCOVA, whose outcome was to consider germination year of propagation source (GY) as a random factor according to the interaction coefficients with q and Q_{12} in GLM (Pr (<|z|) <0.0001 of q:GY; Pr (<|z|) = 0.0031 of GY:Q_{12}). To explain the cutting success of poplar samples, we used model comparisons of GLMM. According to the significant change of AIC values, we found our best fitting model when introducing admixture index as a fixed predictor with GY as a random factor (ANOVA, ΔAIC= 88.36, *p*-value= <0.0001, Table S6). Admixture index had a positive effect on clonability (Table 2), in agreement with the expectation that reproduction of *P. tremula* via shoot cutting is less successful compared to *P. alba*. Among hybrids, individuals genetically more similar to *P. alba* exhibited greater cutting success.

Table 2. Fixed coefficient of the best fitted model regarding cutting success with addition p-value

	Estimate	Std..Error	z.value	p.value
Intercept	-1.90166	0.232032	-8.19569	<0.0001
q	1.836514	0.205571	8.933744	<0.0001

Figure 6 and 7 suggest that *q* and germination year of propagation source are both applicable to predict the probability of cutting success. As *q* value is approaching its maximum (*q*=1.0), also the probability of clonability (survivorship of germinated shoot cuttings) is increasing. Backcrosses towards *P. alba* are more likely to be able to reproduce asexually via shoot cuttings successfully. Variation in germination year of propagation source was clearly present in the data. However, individuals of *P. tremula*, which had lower cutting success than those of *P. alba*, were unevenly distributed among sampling years (Fig. 6). Therefore, their predicted probability was generalized on the entire scale of *q* (Fig. 7). This generalization showed that the lowest probability was associated with the minimum value of admixture index (*q*=0.00) and with older propagation source. These results supported our hypothesis that clonability is dependent on genetic ancestry and affected by the age of the material used for clonal propagation.
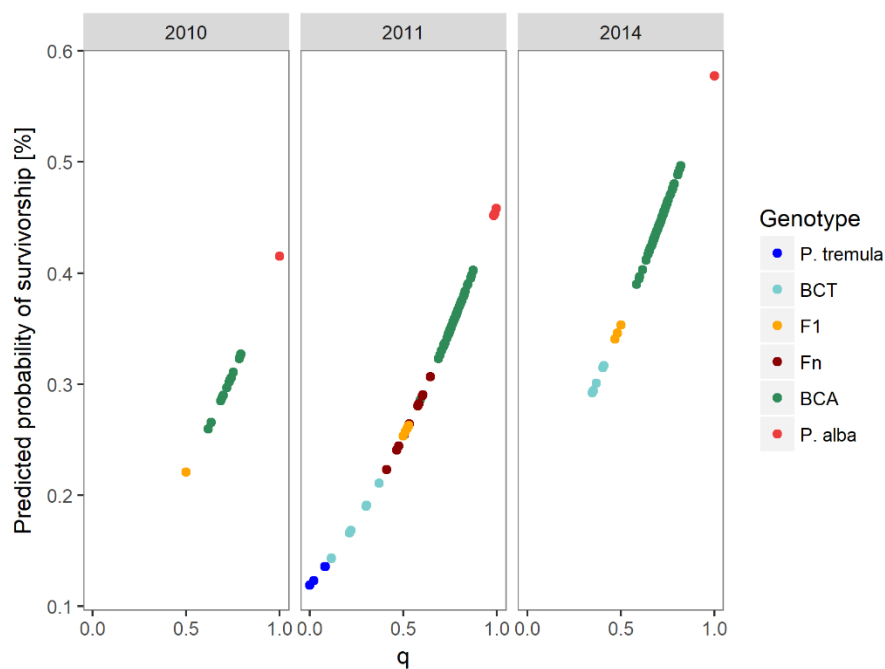

Figure 6. Predicted probabilities of clonal propagation success (predicted probability of clonal cutting survivorship) plotted against admixture index, divided by germination year of propagation source.
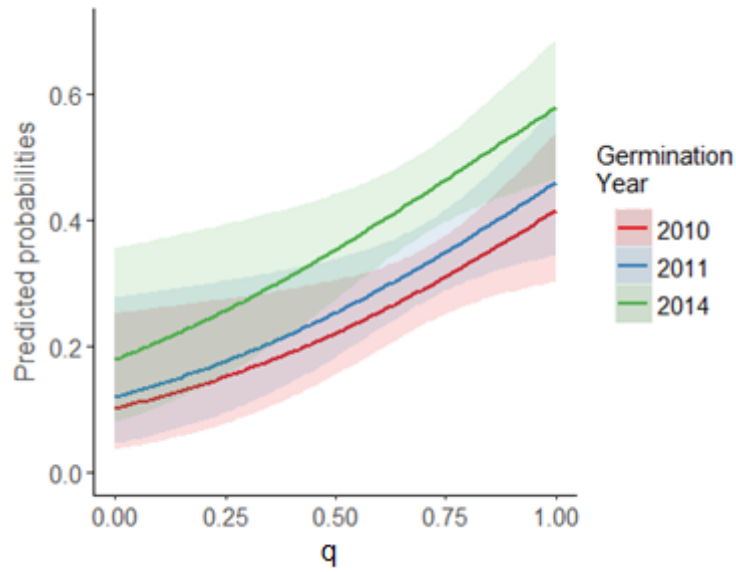
Figure 7. Predicted probability curves for each germination year of cutting success along the scale of q with 95% confidence intervals

Growth

To obtain the best explanation for growth differences, during our first approach we investigated the height of trees as a cumulative index. Height differences across genotypes were detected during the vegetation period of 2018. The best fitting model according to ANOVA of LMEM contained: germination year of propagation sources (GY), admixture index (q), linear, quadratic and cubic values of measurement time points as fixed predictor variables and block position in CG as random factor (ANOVA of LMEM, ΔAIC= 9.1, *p*-value=0.0009, Table S2). The positive effect of admixture index (Table 3) indicates increased height of individuals genetically closer to *P.alba* (at Figure 8 all height measurements were taken during the investigated period). Germination year of propagation source showed a negative effect on height (Table 3). We considered the effect of measurement time points, equal to the effect of time during the vegetation period. Figure 9 demonstrates a slow growth in the beginning and at the end of the vegetation period. The locations of the blocks in the CG affected the growth and individuals in different blocks varied in height (Figure 10).

Table 3. Significance of fixed coefficients of the best fitting model using height measurements of each months as a response variable

	Value	Std.Error	DF	t.value	p.value
Intercept	5206.755	1511.047	1076	3.445794	0.0006
Month	-24.4884	12.00477	1076	-2.03989	0.0416
I(Month^2)	12.178	3.841469	1076	3.170142	0.0016
I(Month^3)	-1.16784	0.363008	1076	-3.21713	0.0013
q	31.5313	5.561592	1076	5.669473	<0.0001
GY	-2.50248	0.75127	1076	-3.33099	0.0009
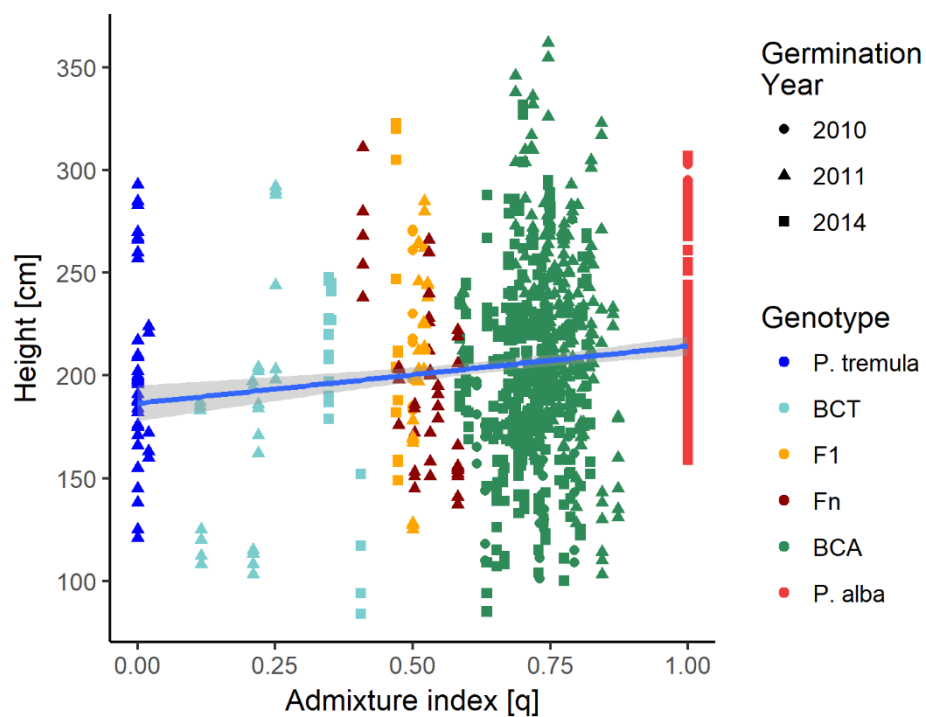
Figure 8. Linear regression line with 95% confidence interval for relationship between all height measurements during the vegetation period and admixture index for the best fitting model
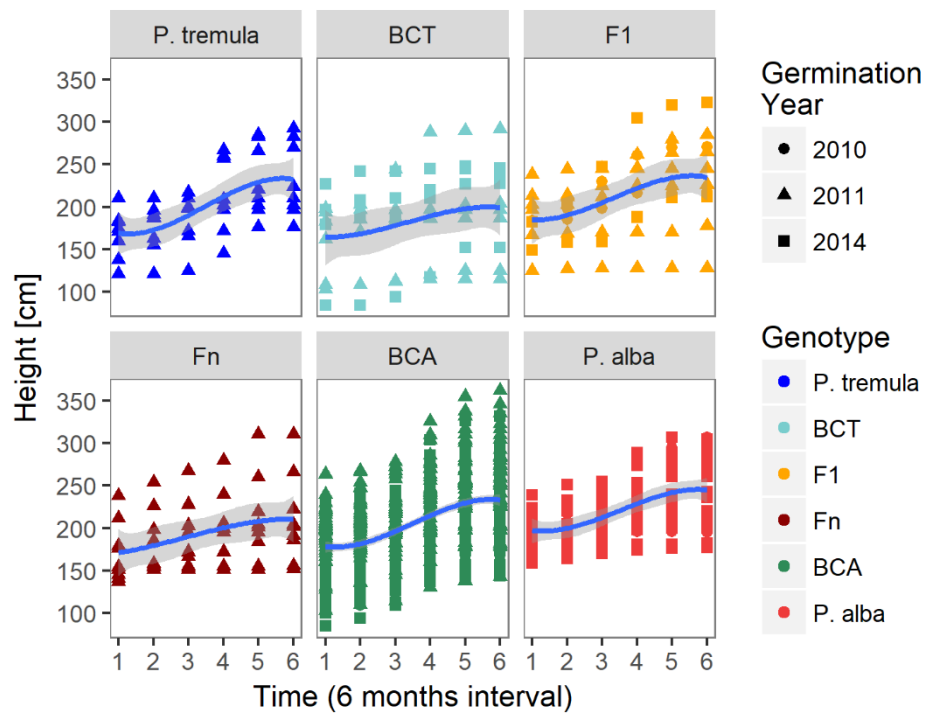
Figure 9. Polynomial regression line for the same model divided by the genotypes, representing changes in height during the 6 measurement time points
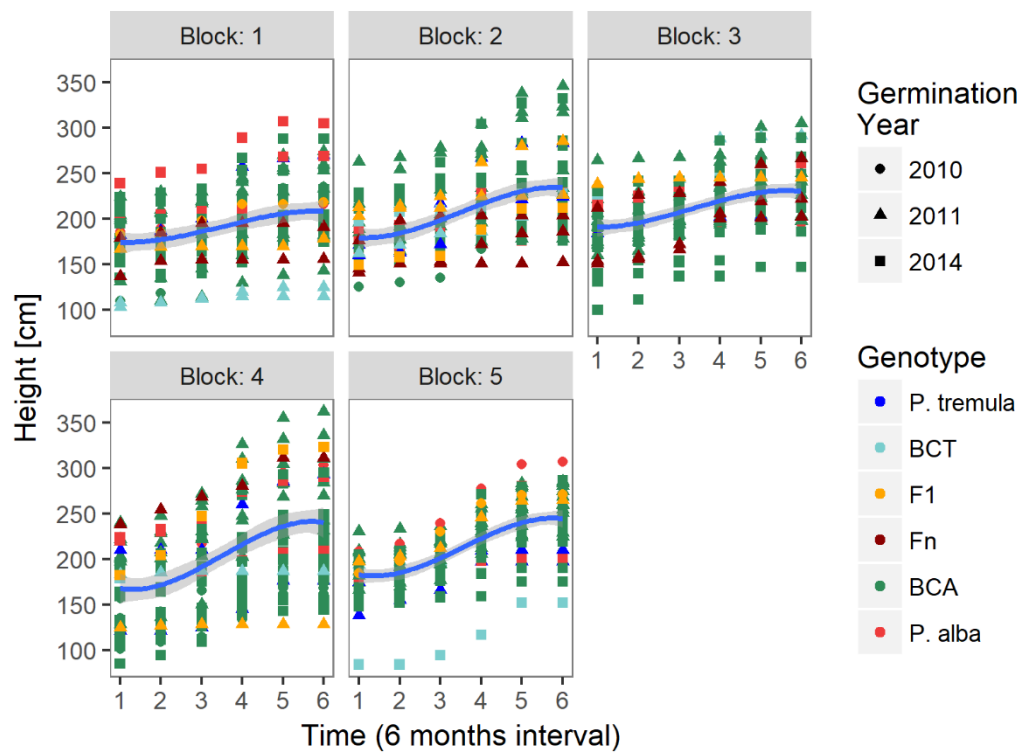
Figure 10. Scatterplot with 95% confidence interval of the best fitting model for height changes in each month, showing the effects of blocks on height.

For subsequent analyses with RGR, we chose data for height measurements over those for diameter. Spearman's rank correlations (Figure 11) showed a highly significant positive correlation between RGR calculated from height and from diameter values (Spearman's rank test, p-value < 2.2e-16, ρ= 0.5898).
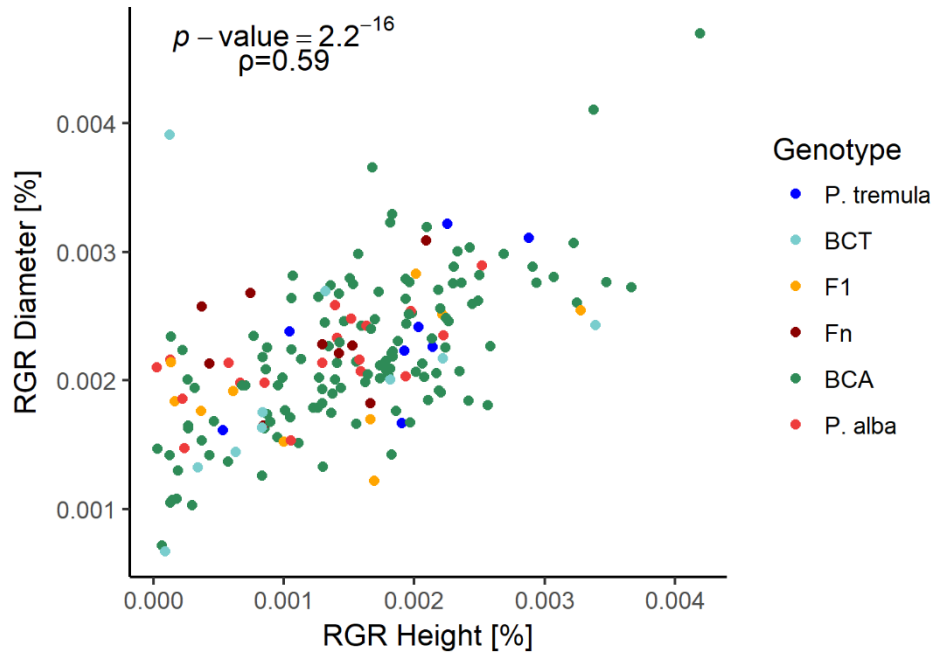
Figure 11. Plot of Spearman's rank test shows a positive correlation between RGR calculated from height data and RGR from diameter

In our second approach we changed our response variable to RGR calculated for each month with April as reference and applied the same model structure as for height (above). After comparison of LMEMs, our best fitting model (ANOVA of LMEM, ΔAIC= 15.96, *p*-value <0.0001, Table S3) only included linear, quadratic and cubic effects of time and block position in CG as random factor. Neither genetic parameters nor GY caused a significant change when introduced in the models.

In our third approach, RGR was calculated for the vegetation period of 2018 to see an overall picture of growth during this time. The best fitting model only included intercept and block number in CG as random factor (ANOVA, ΔAIC= 26.07, *p*-value<0.0001, Table S4). None of known genetic parameters brought significant changes in model comparisons. Graphical inspection (Figure S1) suggested merit in the removal of individuals of the pure parental species from the dataset to bring out patterns among admixed individuals (=hybrids) more

clearly. Examination of hybrids revealed an additional trend not seen in the earlier analyses. In particular, many of the BCAs exhibited increased RGR, compared to *P. alba*. In the fourth and last approach we ran another LMEM comparison with the same model structure as previously. The best fitting model included admixture index as fixed predictor variable and block position as a random factor (ANOVA, ΔAIC= 5.35, *p*-value= 0.0067, Table S5). A positive effect of closer relatedness to *P. alba* on RGR became clearly visible from the analysis of hybrid individuals (Figure 12; Table 4). Blocks in the CG differed in RGR (Figure 13), with Block 4 showing greatly increased RGR for BCAs, while in other blocks differences were more moderate.

Table 4. Significance of fixed coefficients of the top-rank model using only overall RGR of hybrids as a response variable

	Value	Std.Error	DF	t.value	p.value
Intercept	0.000734	0.00034	148	2.159	0.0324
q	0.001156	0.000424	148	2.727	0.0071
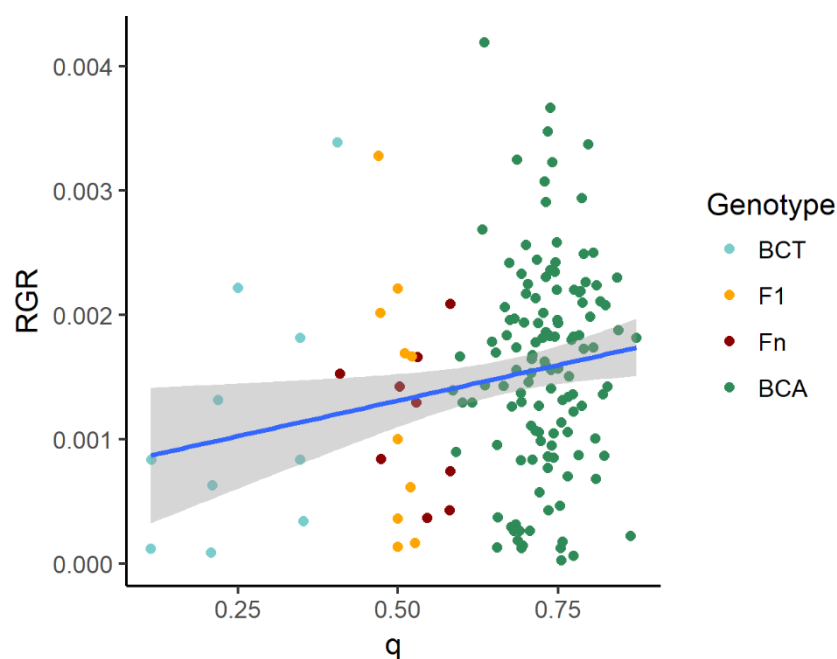
Figure 12. Scatterplot with 95% confidence interval of the best model for the overall RGR of hybrids
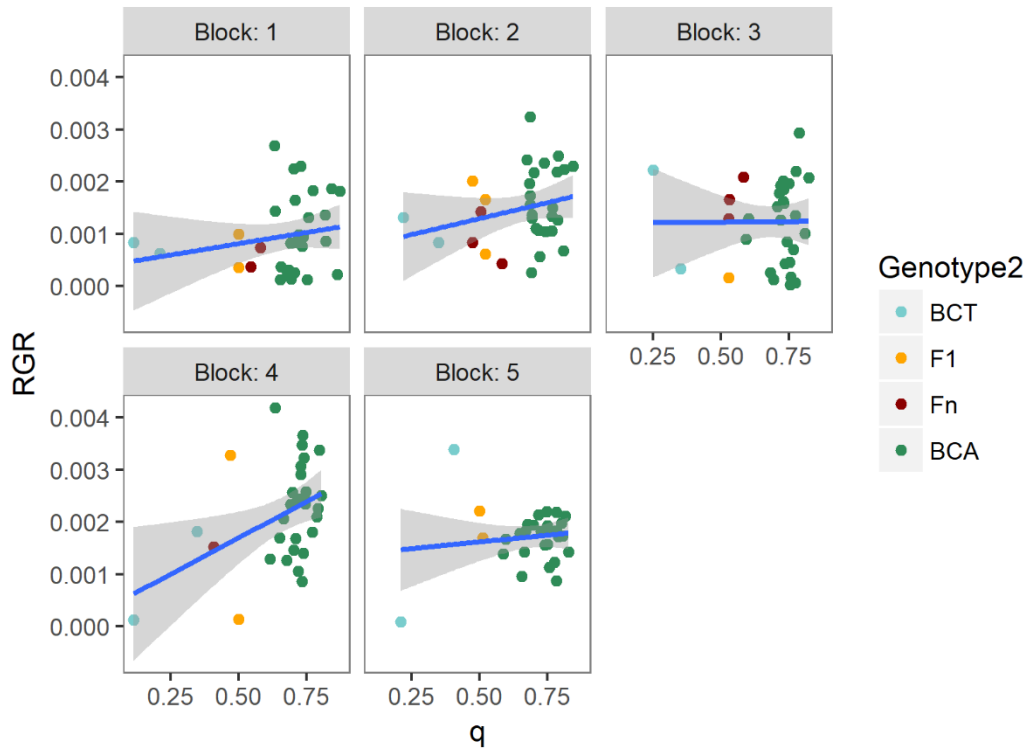
Figure 13. Regression line of the top rank model for overall RGR of hybrids showing effects of blocks in the CG

Leaf reflectance

Our analysis of the hairiness of abaxial leaf surface revealed how genetic ancestry indices influence this functionally important phenotypic trait. Measurements of mean, modal and median grey values for individual trees were extremely correlated (Pearson correlation coefficient, $r > 0.99$ in all three pairs of variables), thus mean grey values were used for all subsequent analyses. ANOVA was carried out to evaluate the need for mixed models. This resulted in no significant decrease of AIC values, therefore the arrangement of trees in blocks did not influence our response variable (ANOVA, ΔAIC= -1.79, *p*-value= 0.6465). Linear models were built and compared with ANOVA using *F*-statistics (Table 5) to find the best fitting model. Introduction of admixture index (*q*) brought a decrease of residual sum of squares (RSS), but adding the quadratic value of *q* resulted in an even lower RSS with a significant *p*-value ($p<0.05$, Table S7). Inclusion of inter-source ancestry (Q_{12}) did not result in significant changes during model comparison. To explain the variability of abaxial leaf hairiness we used the polynomial form of admixture index ($y \sim x + x^2$) as predictor variable in our best fitting model.

Table 5. ANOVA and F test between our linear models

Predictors	Df	SS	RSS	F	p.value
Intercept	NA	NA	142654.31	NA	NA
q	1	92793.099	49861.21	373.4907	<0.0001
$q+q^2$	1	1587.918	48273.29	6.391344	0.0123
$q+q^2+Q_{12}$	1	74.347	48198.95	0.299243	0.5850

According to the significance of coefficients for fixed predictors of the best fitting model (Table 6), we rejected our null hypothesis. These results also allowed us to conclude that both linear and quadratic values of *q* had significant effects on our response variable. These positive effects of *q* indicate that backcrosses towards *P. alba* (BCA) exhibit increased trichome density. The polynomial form brought greater significance to the model (Fig. 14), which may indicate epistasis or episodes of selection (discussed below).

Table 6. Significance of fixed coefficients of the top-rank linear model

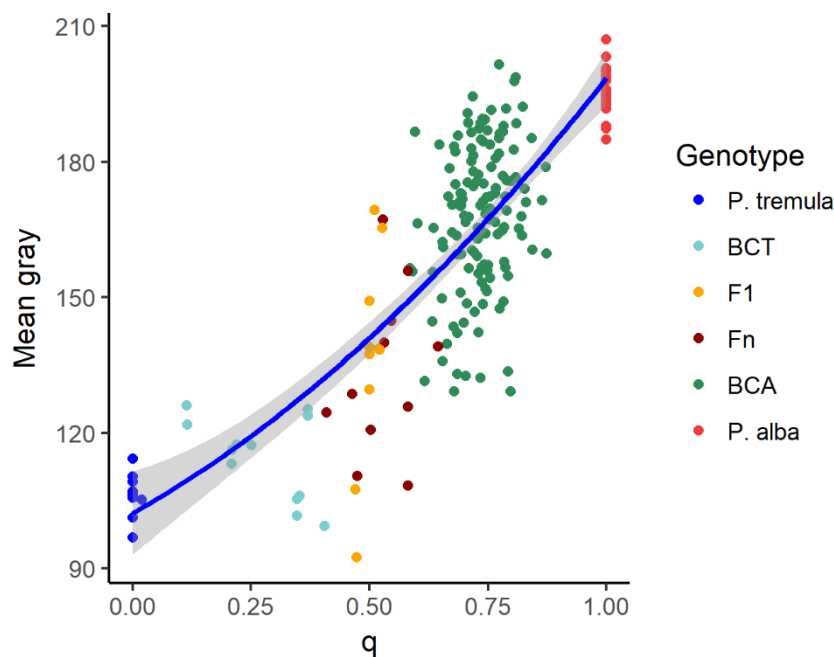| Predictors	Estimate	Std..Error	t.value	Pr(>|t|)
Intercept	102.2989	4.64277	22.03401	<0.0001
q	58.6453	16.21729	3.61622	0.000381
q^2	37.4115	14.77157	2.53267	0.012108

Figure 14. Polynomial regression plot with 95% confidence interval, showing the relationship between admixture index (q) and mean grey value of leaf lower surface

Discussion

Seedling survivorship

Populus species can reproduce asexually *and* sexually, but even despite high numbers of seeds, seedling mortality can be high and juvenile survivorship differs among species (Scarascia-Mugnozza et al. 1997, Schweitzer et al. 2002, Dickmann and Kuzovkina. 2008, Lindtke et al. 2014). In this thesis, survivorship and mortality of young *P. alba*, *P. tremula* and hybrid saplings was recorded three years after planting in two different common gardens (CGs) (Salerno, Italy and Fribourg, Switzerland). Genetic ancestry indices evaluated from previously presented GBS data (Lindtke et al. 2014) were used to examine relationships between ancestry and fitness-related traits at early life stages in these model forest tree species and their hybrids. Our analyses of GLMMs for two different CG localities supported the findings of Christe et al. (2016) on a single CG locality. Reduced inter-source ancestry (Q_{12}) and lower, intermediate admixture index (q) were associated with a higher chance of seedling mortality. Thus, backcrosses towards *P. tremula* and recombinant hybrid genotypes showed reduced survivorship, while individuals that were genetically closer to *P. alba* and first generation (F1) hybrids had higher survival rates. This points to the presence of heterosis in F1´s, as hypothesized by Christe et al. (2016). In both studies, the same fixed predictor variables were highly significant, including both linear and quadratic aspects of q and linear aspects of Q_{12}. This suggests that not only the proportion of parental genomes (q) but also diploid ancestry combinations (Q_{12}) determine this juvenile fitness proxy (Lindtke at al. 2014, Christe et al. 2016). Introducing the locality of the two CGs in our model as fixed predictors was a novel step, but it did not change the results significantly (Table S1). The survivorship differences were therefore independent of the locality of common gardens, i.e. they appear to have a strong intrinsic component. Previous studies identified backcrosses to *P. alba* (BCAs) (Lexer et al. 2005, Van Loo et al. 2008, Lexer et al 2010, Lindtke et al. 2012) and F1s (Santos del Blanco et al. 2012, Lindtke et al. 2014, Christe et al. 2016) as the dominant classes of hybrid genotypes in natural populations, and the present results suggest that these patterns may partially be due to asymmetric patterns of post-zygotic selection. Despite the high genetic diversity of hybrid seedlings found in natural populations, adult trees show much more moderate diversity due to selection (Christe et al. 2016). The strong

effect of post-zygotic selection during juvenile life stages is expected to play an important role in the composition of natural hybrid populations of these species (Lindtke et al. 2014, Christe et al. 2016). Our results on similar plant mortality in two very different common garden environments (Fribourg and Salerno) suggest that selection during juvenile stages of *Populus* hybrids has an important intrinsic component, most likely related to plant development and performance, as expected for the genetic mechanisms underlying heterosis (Bar-Zvi et al. 2017).

Clonability

Our findings on the propensity to reproduce asexually from shoot cuttings supported available knowledge from the literature. *Populus alba* and *P. tremula* have mixed reproduction, both asexual and sexual, but their propensity to clone via shoot cuttings differs (Dickmann and Kuzovkina. 2008). In our experiment *P. tremula* and its backcrosses had poor performance, while hybrids genetically more related to *P. alba* and the parental species themselves had higher success in clonability. This pattern is strikingly similar to the composition of hybrid populations of these species in nature (Lexer et al. 2005, Van Loo et al. 2008), where BCA hybrids are more frequent than BCTs. Our results also revealed that not only the admixture index, but also the germination year (=age) of the propagation source has an effect on asexual reproduction. The predicted probability of shoot cutting survivorship, and thus clonability in our common garden experiment, pointed to lower mortality of cuttings derived from younger propagation sources (Fig. 7). In natural habitats, *Populus* species are more likely to apply a related form of asexual reproduction, namely ramet production via root suckers (Dickmann and Kuzovkina. 2008). This strategy also differs among species, and clone size and population composition can vary with time and also depending on the locality (Barsoum et al. 2004, Latva-Karjanmaa et al. 2006, Santos del Blanco et al. 2012). In northern Utah, USA, Schweitzer et al. (2002) observed that hybrids of *P. fremontii* x *P. angustifolia* can produce more ramets than their parental species. In the study of Van Loo et al. (2008) no significant difference was detected in the number of ramets between *P. alba* and *P. x canescens*, but *P. x canescens* clones seemed to be able to cover larger areas in the riparian forests of the Danube in Central Europe. Even if sexual reproduction of these species is more frequent than previously considered (Latva-Karjanmaa

et al. 2006, Van Loo et al. 2008), the importance of asexual reproduction cannot be overstated. This strategy can provide fitness benefits for these pioneer species, especially in harsh and extreme environments, for example at the margin of species´ranges (Eckert 2002) or in mosaic hybrid zones, particularly in riparian floodplain forests where conditions fluctuate and floods can be followed by periods of heat and drought. Another important aspect is that diploid hybrids and their parental species of *Populus* can maintain their adapted and beneficial genotypes (Schweitzer at al. 2002, Van Loo et al. 2008). These properties are not only desired in silviculture (Scarascia-Mugnozza et al. 1997), but the ability to propagate vegetatively is also a crucial and important survival strategy in natural forest habitats with relevance for ecosystem services.

Growth

For this juvenile fitness trait, we tested four different approaches to explore the relationships between genetic indices, height, and growth properties. We compared height measurements for each month, which resulted in significant effects of time, age of propagation source (germination year, GY) and admixture index (q). The effect of time indicated increased growth between May and September, as expected (Fig. 8). The tendency was lower in the beginning and at the end of the vegetation period. From the genetic indices only q brought significant changes to our model. Individuals which were more closely related to *P. alba* were more likely to grow taller and F_1s performed better than BCTs or Fns (Fig. 9). When we analysed relative growth rate (RGR) for each month and overall RGR including the parental species, inter- or intraspecific differences caused by genetic ancestry were not detected. After we approached our dataset in a more targeted way, focusing only on the overall RGR of hybrid genotypes, we uncovered differences in RGR according to the admixture index. In this analysis, the genetic relatedness exhibited the same benefits as in the height analyses. Growth differences were influenced also by the locations of planting blocks in the CG in both analyses; trees planted in Block 4 deviated from the rest of the samples in height increment and overall RGR of hybrids (Fig. 10 & 13). This observation may be explained by the presence of distinct microenvironments, despite the equal treatment and small distances between blocks. Scarascia-Mugnozza et al. (1997) observed growth differences derived by environment in other *Populus spp*., which indicated sensitivity to

differences in soil composition. In our CG, the same type of soil was used during planting, therefore soil type should not be the main influence. The same can be expected for watering regime, as we employed an automated watering system. It is more plausible that other biotic or abiotic factors played a role (e.g. shade/sun), but we did not investigate this question further in this study. This phenomenon could emphasize the importance of microenvironmental conditions, not only in a CG, but also in natural habitats. In the same study, Scarascia-Mugnozza et al. (1997) discussed hybrid superiority (*Populus trichocarpa x P. deltoids*) in stem volume and in total height of trees. Our findings do not support heterosis of *P. x canescens* in growth, but show a trend of poorer performance connected to relatedness to *P. tremula*. Features of growth can be correlated, among others, with bud break and bud set of trees, because these are associated with the length of the growing season (De Carvalho et al. 2010). Bud break and bud set are determined by the place of origin of the propagation source, and these phenological traits are known to vary along latitudinal gradients (Scarascia-Mugnozza et al. 1997, Luquez et al. 2008, De Carvalho et al. 2010). In our study, material from both species was collected from the same sampling site, and thus it is more plausible that heritable differences between these two ecologically divergent species and their hybrids played an important role in plant growth. We measured early growth as a juvenile fitness proxy, because lifetime fitness is difficult to observe in these long-lived tree species in a CG. Growth measured for a time period can indicate fitness and access to resources, and juvenile growth differences in critically important periods can affect the fate of plants (Coley et al. 1985). As both studied species and their hybrids are pioneer trees, increased growth during early life stages can be an essential benefit in inter- and intraspecific competition during the colonization of new habitats (Zuest and Agrawal, 2017).

Leaf reflectance

Leaf reflectance of the abaxial leaf surface, as a surrogate for trichome density, differed along the ancestry / admixture axis between these two divergent *Populus* species. Our results revealed that backcrosses towards *P. alba* and the parental species expressed a denser indumentum of hairs on the leaf underside. The importance of genetic determination of this trait is also emphasized by the non-significant random factor in our models, i.e. the

distinct locations of the blocks in the CG had no influence on this trait. This suggests that divergence in leaf hairiness was more likely due to genetic differences than environmental factors. The significance of the quadratic term in the best fitting model can be explained by several possible hypotheses. One hypothesis is that the non-linear fit can be interpreted by missing backcrosses to *P. tremula* (BCT). The absence of these individuals in our CG is the result of selection during the early life stages of these seedlings (Christe et al. 2016). A higher number of trichome-expressing phenotypes might have survived, compared to saplings lacking this feature. An alternative possibility is that the quadratic term could indicate epistasis (non-additive genetic interactions between alleles of different loci) involving genes of the two parental species transmitted to hybrids (Christe et al 2016; Bresadola et al 2018). Although backcrosses towards *P. alba* present a higher numbers of observations in our dataset, these samples are nevertheless equally distributed around both the linear and the polynomial regression line (Fig. 14). F_1 and especially F_n genotypes deviate more strongly from the linear trend between response and predictor variables. The polynomial regression represented a better fit especially for these individuals (Fig. 14), which is consistent with the epistasis hypothesis. Much research has been carried out on trichomes, because these specialized epidermal cells in plants alter functional traits and also have taxonomic importance. Increased trichome density can reduce transpiration and thus water loss (Perez-Estrada et al. 2000), which can prevent heat stress in riparian habitats (e.g. in Heißlend habitats in river floodplains) when the water level is low in mid-Summer. Negative correlations between trichome density and herbivory have been observed as well (Plett at al. 2010); defence against herbivory is an important ecological and evolutionary feature of natural populations of *Populus* species (Whitham et al. 2006). Backcrosses towards *P. alba* may gain greater resistance against herbivory due to higher trichome density, thus increasing survival and fitness. Trichomes may also alter photosynthesis via light absorption, but evidence for this is not unequivocal (Plett at el. 2010). The relationships between growth and trichome density and possible trade-offs between growth and defence are heavily debated topics (Zuest and Agrawal, 2017). Plant growth can be negatively influenced by the costly production of trichomes (Gruber et al. 2006). In other cases, individuals with higher trichome density showed increased growth tendency (Plett et al.

2010). Our findings suggest that the increased leaf hairiness might provide benefits to the trees. These altered, fitness-related features can determine the fate of hybrids, therefore the ecological importance of this functionally important plant trait is supported. The significant relationship between genetic ancestry and trichome density in these common garden grown hybrids indicates a great potential to unlock the genomic architecture of this functionally important plant trait using admixture mapping, i.e. genome wide association scanning in hybridizing populations (Buerkle & Lexer 2008; Bresadola et al. 2018).

References

Abbott, Richard & Albach, Dirk & Ansell, S & Arntzen, Jan & Baird, Stuart & Bierne, Nicolas & Boughman, Janette & Brelsford, Alan & Buerkle, C.A. & Buggs, Richard & Butlin, Roger & Dieckmann, U & Eroukhmanoff, Fabrice & Grill, Andrea & Cahan, S.H. & Hermansen, Jo & Hewitt, G & Hudson, Alan & Jiggins, Chris & Zinner, Dietmar. (2013). Hybridization and speciation. Journal of Evolutionary Biology. 26. 246. 10.1111/j.1420-9101.2012.02599.x.

Abbott, Richard. (2017). Plant speciation across environmental gradients and the occurrence and nature of hybrid zones: Hybrid zones and plant speciation. Journal of Systematics and Evolution. 55. 10.1111/jse.12267.

B. Pérez-Estrada, Leticia & Cano-Santana, Zenón & Oyama, Ken. (2000). Variation in leaf trichomes of Wigandia urens: Environmental factors and physiological consequences. Tree physiology. 20. 629-632. 10.1093/treephys/20.9.629.

Baack, Eric & H Rieseberg, Loren. (2007). Baack EJ, Rieseberg LH. A genomic view of introgression and hybrid speciation. Curr Opin Genet Dev 17: 513-518. Current opinion in genetics & development. 17. 513-8. 10.1016/j.gde.2007.09.001.

Barsoum, Nadia & Muller, Etienne & Skøt, Leif. (2004). Variations in levels of clonality among *Populus nigra L.* stands of different ages. Evolutionary Ecology. 18. 601-624. 10.1007/s10682-004-5146-4.

Bartha D. (2004). Short description and key to the poplar taxa (*Populus* L.) existing in Hungary (published in Hungarian). Flora Pannonica 2 (2):85-101

Bar-Zvi, D., Lupo, O., Levy, A. A., & Barkai, N. (2017). Hybrid vigor: The best of both parents, or a genomic clash? Current Opinion in Systems Biology, 6, 22– 27.

Bates D, Maechler M, Bolker B and Walker S (2014). *lme4:* Linear mixed-effects models using Eigen and S4. R package version 1.1-7, http://CRAN.R-project.org/package=lme4.

Bresadola, Luisa & Caseys, Céline & Castiglione, Stefano & Alex Buerkle, C & Wegmann, Daniel & Lexer, Christian. (2018). Admixture mapping in interspecific *Populus* hybrids identifies classes of genomic architectures for phytochemical, morphological and growth traits. BioarXiv. doi: https://doi.org/10.1101/497685

Buerkle, C. A., & Lexer, C. (2008). Admixture as the basis for genetic mapping. Trends In Ecology & Evolution, 23(12), 686-94

Caseys C, Glauser G, Stolting KN, Christe C, Albrectsen BR, Lexer C. Effects of interspecific recombination on functional traits in trees revealed by metabolomics and genotyping-by-resequencing. Plant Ecol Divers. 2012; 5:457–71.

Caseys C, Stritt C, Glauser G, Blanchard T, Lexer C (2015) Effects of hybridization and evolutionary constraints on secondary metabolites: the genetic architecture of henylpropanoids in European *Populus* species. PLoS ONE, 10, e0128200.

Caudullo G., de Rigo D., 2016a. *Populus alba* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), European Atlas of Forest Tree Species. Publ. Off. EU, Luxembourg, pp. e010368+

Caudullo, G., de Rigo, D., 2016b. *Populus tremula* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), European Atlas of Forest Tree Species. Publ. Off. EU, Luxembourg, pp. e01f148+

Christe C, Stölting KN, Bresadola L, Fussi B, Heinze B, Wegmann D, Lexer C. (2016) Selection against recombinant hybrids maintains reproductive isolation in hybridizing *Populus* species despite F1 fertility and recurrent gene flow. Molecular Ecology. PMID 26880192 DOI: 10.1111/mec.13587

De Carvalho D, Ingvarsson PK, Joseph J, Suter L, Sedivy C, Macaya-Sanz D, et al. Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula L*.), a widespread forest tree. Mol Ecol. 2010; 19:1638–50. doi: 10.1111/j.1365-294X.2010.04595.x PMID: 20345678

Dickmann, D. I. – Kuzovkina, J. (2008): Poplars and Willows of the World, with Emphasis on Silviculturally Important Species. FAO Forest Management Division Working Paper IPC/9-2; 129 p. FAO Forest Management Division, Rome, Italy, 3–64.

Eckert, Christopher G. "The Loss of Sex in Clonal Plants." Evolutionary Ecology 15, no. 4 (July 1, 2001): 501–20. https://doi.org/10.1023/A:1016005519651.

EUFROGEN: Description of *Populus alba, P. tremula* (www.eufrogen.org)

Fisher RA. 1921. Some remarks on the methods formulated in a recent article on `The quantitative analysis of plant growth'. Annals of Applied Biology 7: 367±372.

Gruber, MY & Wang, S & Harris, Stephanie & Holowachuk, J & C Bonham-Smith, P & Soroka, Juliana & Lloyd, A. (2006). "HAIRY CANOLA" – *Arabidopsis* GL3 induces a dense covering of trichomes on Brassica napus seedlings. Plant molecular biology. 60. 679-98. 10.1007/s11103-005-5472-0.

Harrison, R.G. (1990) Hybrid zones: Windows on evolutionary process. Oxford Surveys in Evolutionary Biology, 7, 59.

Harrison, R.G.. (1993). Hybrids and hybrid zones: Historical perspective. Hybrid Zones and the Evolutionary Process. 3-12.

Hewitt, G. M. "Hybrid Zones-Natural Laboratories for Evolutionary Studies." Trends in Ecology & Evolution 3, no. 7 (July 1988): 158–67. https://doi.org/10.1016/0169-5347(88)90033-X.

Hunt, Roderick. "Relative Growth Rates." In *Basic Growth Analysis: Plant Growth Analysis for Beginners*, edited by Roderick Hunt, 25–34. Dordrecht: Springer Netherlands, 1990. https://doi.org/10.1007/978-94-010-9117-6_3.

Koop, H. "Vegetative Reproduction of Trees in Some European Natural Forests." *Vegetatio* 72, no. 2 (October 1, 1987): 103–10. https://doi.org/10.1007/BF00044840.

Latva-Karjanmaa, Tarja, Leena Suvanto, Kari Leinonen, and Hannu Rita. "Sexual Reproduction of European Aspen (*Populus tremula L*.) at Prescribed Burned Site: The Effects of Moisture Conditions." *New Forests* 31, no. 3 (May 1, 2006): 545–58. https://doi.org/10.1007/s11056-005-2742-2.

Lexer C, Fay MF, Joseph JA, Heinze B, Sorin M-S. 2005. Barrier to gene flow between two ecologically divergent *Populus* species, *Populus alba* (white poplar) and *P. tremula* (European aspen): the role of ecology and life history in gene introgression. Molecular Ecology 14: 1045–1057.

Lexer C, Joseph J, van Loo M, Prenner G, Heinze B, Chase MW, et al. The use of digital image-based morphometrics to study the phenotypic mosaic in taxa with porous genomes. Taxon. 2009; 58:349–64

Lexer C, Joseph JA, van Loo M, Barbara T, Heinze B, Bartha D, et al. Genomic admixture analysis in European *Populus* spp. reveals unexpected patterns of reproductive isolation and mating. Genetics. 2010; 186:699–712. doi: 10.1534/genetics.110.118828 PMID: 20679517

Lexer, Christian & Buerkle, CA & Joseph, Jeffrey & Heinze, Berthold & Fay, Michael. (2007). Admixture in European *Populus* hybrid zones makes feasible the mapping of loci that contribute to reproductive isolation and trait differences. Heredity. 98. 74-84. 10.1038/sj.hdy.6800898.

Lindtke D, Buerkle CA, Barbara T et al. (2012) Recombinant hybrids retain heterozygosity at many loci: new insights into the genomics of reproductive isolation in *Populus*. Molecular Ecology, 21, 5042–5058.

Lindtke D, Gonzalez-Martinez SC, Macaya-Sanz D, Lexer C. Admixture mapping of quantitative traits in *Populus* hybrid zones: power and limitations. Heredity. 2013; 111:474–85. doi: 10.1038/hdy.2013.69 PMID: 23860234

Lindtke, Dorothea, Zachariah Gompert, Christian Lexer, and C. Alex Buerkle. "Unexpected Ancestry of *Populus* Seedlings from a Hybrid Zone Implies a Large Role for Postzygotic Selection in the Maintenance of Species." *Molecular Ecology* 23, no. 17 (2014): 4316–30. https://doi.org/10.1111/mec.12759.

Luquez, Virginia & Hall, David & Albrectsen, Benedicte & Karlsson, Jan & Ingvarsson, Pär & Jansson, Stefan. (2008). Natural phenological variation in aspen (*Populus tremula*): The SwAsp collection. Tree Genetics & Genomes. 4. 279-292.

Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team (2018). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-137, https://CRAN.R-project.org/package=nlme

Plett, Jonathan M., Olivia Wilkins, Malcolm M. Campbell, Steven G. Ralph, and Sharon Regan. "Endogenous Overexpression of *Populus* MYB186 Increases Trichome Density, Improves Insect Pest Resistance, and Impacts Plant Growth: Poplar MYB186 Impacts Growth and Pest Resistance." *The Plant Journal* 64, no. 3 (November 2010): 419–32. https://doi.org/10.1111/j.1365-313X.2010.04343.x.

R Development Core Team (2008). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org

Santos del Blanco, Luis & de-Lucas, Ana & González-Martínez, Santiago & Sierra-de-Grado, Rosario & Hidalgo, Elena. (2012). Extensive Clonal Assemblies in *Populus alba* and *Populus x canescens* from the Iberian Peninsula. Tree Genetics & Genomes. 9. 10.1007/s11295-012-0574-8.

Scarascia-Mugnozza, Giseppe & Ceulemans, R & Heilman, P.E. & Isebrands, J.G. & Stettler, Reinhard & Hinckley, Tom. (1997). Production physiology and morphology of *Populus* species and their hybrids grown under short rotation. II. Biomass components and harvest index of hybrid and parental species clones. Canadian Journal of Forest Research. 27. 285 - 294. 10.1139/x96-180.

Schweitzer JA, Martinsen GD, Whitman TG. 2002. Cottonwood hybrids gain fitness traits of both parents: a mechanism for their long-term persistence? American Journal of Botany 89: 981–990.

Stettler R., Bradshaw T., Heilman P., Hinckley T.,(editors). Biology of *Populus* and Its Implications for Management and Conservation. Canada, National Research Council ,NRC Research Press, 1996.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A et al. 2006. The genome of black cottonwood*, Populus trichocarpa* (Torr. & Gray). Science 313: 1596–1604.

van Loo M, Joseph J, Heinze B, Fay M, Lexer C (2008) Clonality and spatial genetic structure in *Populus x. canescens* and its sympatric backcross parent *P. alba* in a Central European hybrid zone. New Phytologist , 177, 506–516.

Whitham, Thomas G & Bailey, Joseph & Schweitzer, Jennifer & Shuster, Stephen & Bangert, Randy & LeRoy, Carri & Lonsdorf, Eric & Allan, Gery & P DiFazio, Stephen & Potts, Bradley & Fischer, Dylan & Gehring, Catherine & Lindroth, Richard & C Marks, Jane & Hart, Stephen & Wimp, Gina & C Wooley, Stuart. (2006). A framework for community and ecosystem genetics: From genes to ecosystems. Nature reviews. Genetics. 7. 510-23. 10.1038/nrg1877.

Zuest, Tobias & Agrawal, Anurag. (2017). Trade-Offs Between Plant Growth and Defense Against Insect Herbivory: An Emerging Mechanistic Synthesis. Annual Review of Plant Biology. 68. 10.1146/annurev-arplant-042916-040856.

Appendix

Figure S1. Relationship between q and overall RGR calculated for 2018.
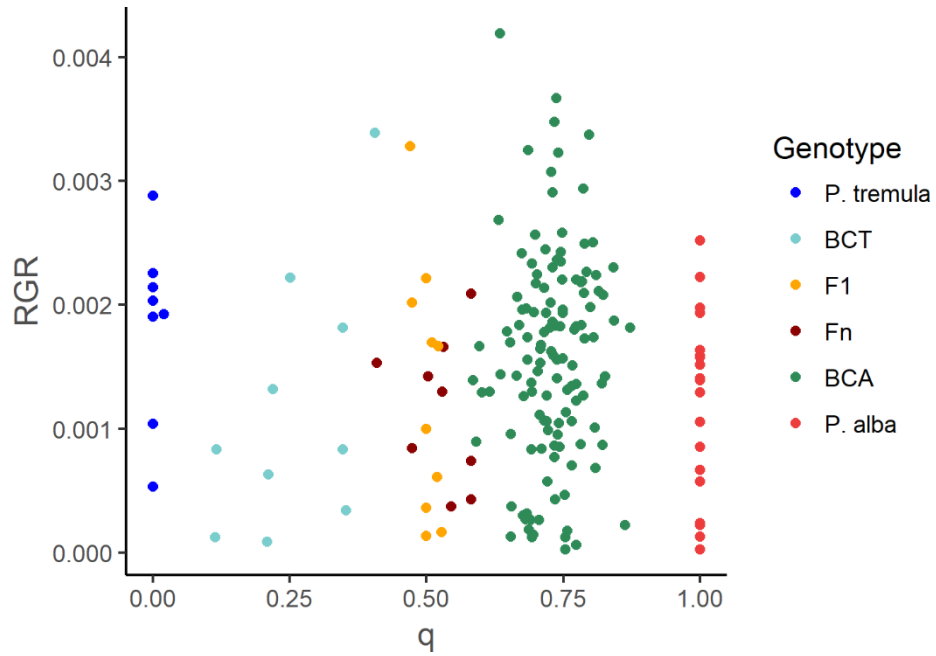
Table S1. Model comparison and Chi^2-test between the five GLMMs of seedling survivorship. In each model, maternal family was introduced as a random factor. Estimates of fixed coefficients evaluation of each model are added to the predictors.

Model	Intercept	q linear	q quad	Q12 linear	CG	df	logLik	AIC	delta AIC	p.value
GLMM1	2.91					2	-64.7	133.5		
GLMM2	1.34	3.95				3	-60.1	126.1	7.4	0.0022
GLMM3	2.04	-3.55	9.13			4	-58.4	124.7	1.4	0.0652
GLMM4	1.80	-31.33	32.63	16.78		5	-46.4	102.7	22.0	<0.0001
GLMM5	1.29	-31.57	33.19	16.94	1.05	6	-44.7	101.4	1.4	0.0672

Table S2. Model comparison and Chi^2-test between the eight height growth LMEMs. For height measurements of each month as a response variable, April was used as baseline for measurement time points. In each LMEM, location in the CG (block number) was added to the models as a random factor. Values of fixed coefficient evaluations of each model are added to the predictors.

Model	Intercept	time linear	time quad	time cubic	q linear	Q12 linear	GY	df	logLik	AIC	delta AIC	p.value
GLS1	205.42							2	-5687.1	11378.2		
LME1	205.35							3	-5678.7	11363.5		
LME2	163.55	11.94						4	-5553.6	11115.2	248.3	<0.0001
LME3	162.77	12.53	-0.08					5	-5553.6	11117.1	-2.0	0.8625
LME4	192.19	-24.49	12.18	-1.17				6	-5548.6	11109.1	8.0	0.0015
LME5	173.62	-24.49	12.18	-1.17	27.39			7	-5536.0	11086.0	23.1	<0.0001
LME6	5206.76	-24.49	12.18	-1.17	31.53		-2.50	8	-5530.5	11076.9	9.1	0.0009
LME7	5005.07	-24.49	12.18	-1.17	30.05	-6.56	-2.40	9	-5529.6	11077.1	-0.2	0.1790

Table S3. Model comparison and Chi^2-test between the eight RGR LMEMs, RGR calculated for each month with April as reference. In each LMEM, location in the CG (block number) was added to the models as a random factor. Values of fixed coefficients evaluation of each model are added to the predictors.

Model	GLS1	LME1	LME2	LME3	LME4	LME5	LME6	LME7
Intercept	0.000952	0.000949	-0.000095	-0.000439	0.000406	0.000427	-0.023056	-0.020571
time linear			0.000348	0.000643	-0.000545	-0.000545	-0.000545	-0.000545
time quad				-0.000049	0.000404	0.000404	0.000404	0.000404
time cubic					-0.00005	-0.00005	-0.00005	-0.00005
q linear						-0.000031	-0.00005	-0.000032
Q12 linear								0.000081
GY							0.000012	0.00001
df	2	3	4	5	6	7	8	9
logLik	5146.7	5175.8	5399.8	5407.9	5416.9	5417.0	5417.4	5417.9
AIC	-10289.3	-10345.7	-10791.5	-10805.8	-10821.8	-10819.9	-10818.8	-10817.7
delta AIC		56.3	445.8	14.3	16.0	-1.9	-1.1	-1.0
p.value		<0.0001	<0.0001	0.0001	<0.0001	0.7344	0.3549	0.3261

Table S4. . Model comparison and Chi^2-test between the RGR LMEMs, RGR was calculated for whole the vegetation period of 2018, location in the CG (block number) was added to the LMEMS as a random factor. Values of fixed coefficients evaluation of each model are added to the predictors.

Model	Intercept	q linear	Q12 linear	GY	df	logLik	AIC	delta AIC	p.value
GLS1	0.0014938				2	1023.6	-2043.2		
LME1	0.0003748				3	1037.6	-2069.2	26.1	<0.0001
LME2	0.0014579	0.0000449			4	1037.6	-2067.3	-2.0	0.8609
LME3	-0.0291600	0.0000198		0.0000152	5	1037.7	-2065.5	-1.8	0.6683
LME4	-0.0263138	0.0000406	0.0000924	0.0000138	6	1037.8	-2063.6	-1.8	0.6896

Table S5. Model comparison and Chi^2-test between the RGR of hybrids LMEMs, RGR was calculated for whole the vegetation period of 2018, location in the CG (block number) was added to the LMEMS as a random factor. Values of fixed coefficients evaluation of each model are added to the predictors.

Model	Intercept	q linear	Q12 linear	GY	df	logLik	AIC	delta AIC	p.value
GLS1	0.00151				2	868.7	-1733.4		
LME1	0.00151				3	880.8	-1755.6	22.2	<0.0001
LME2	0.00073	0.00116			4	884.5	-1761.0	5.4	0.0067
LME3	-0.08603	0.00113		0.00004	5	885.1	-1760.3	-0.7	0.2468
LME4	-0.08540	0.00113	0.00002	0.00004	6	885.1	-1758.3	-2.0	0.9667

Table S6. Model comparison and Chi^2-test between the five GLMMs of clonability. Germination year of propagation source was considered as a random factor in the models. Estimates of fixed coefficients evaluation of each model are added to the predictors.

Model	Intercept	q	Q12	DC	df	logLik	AIC	delta AIC	p.value
GLMM1	-0.56				2	-700.6	1405.1		
GLMM2	-1.90	1.84			3	-655.4	1316.8	88.4	<0.0001
GLMM3	-1.68	1.69	-0.25		4	-654.3	1316.6	0.1	0.1428
GLMM4	-1.61	1.70	-0.25	-0.03	5	-654.1	1318.2	-1.6	0.5348

Table S7. Analysis of variance and F test between the four linear models of leaf reflectance. Estimates of fixed coefficient evaluations of each model are added to the predictors.

Model	Intercept	q linear	q quad	Q12 linear	df	RSS	delta RSS	p.value
LM1	161.05					142654.3		
LM2	94.80	97.88			1	49861.2	92793.1	<0.0001
LM3	102.26	58.70	37.56		1	48273.3	1587.9	0.0123
LM4	102.23	36.38	60.31	7.50	1	48198.9	74.3	0.5850