# Sebastian Szczepański

# Genetic diversity of the hybrid zones and allopatric populations of closely related pine species (genus *Pinus*)

This doctoral dissertation consists of four research articles describing genetic studies of allopatric and hybrid populations of Scots pine (*Pinus sylvestris* L.), dwarf mountain pine (*Pinus mugo* T.), and peat-bog pine (*Pinus uliginosa* N.). The main research focus was on three contact zones between *P. sylvestris* and *P. mugo*, where hybrid forms sometimes exhibit a phenotype resembling *P. uliginosa*. These zones included two populations located in the Central Sudetes: Błędne Skały in the Stołowe Mountains and the Torfowisko pod Zieleńcem peat bog in the Bystrzyckie Mountains, as well as the Bór na Czerwonem nature reserve in Nowy Targ, situated between the Gorce and Tatra Mountains. The reference group comprised a cross-section of natural *P. sylvestris* populations, both lowland and montane, within the Polish and broader European range. Reference populations of *P. mugo* included all Polish mountain ranges where this species occurs, as well as selected European sites. Reference material for *P. uliginosa* was collected from its *locus classicus* – the Wielkie Torfowisko Batorowskie peat bog – and two forest stands near Węgliniec. Sampling was conducted between 2021 and 2023.

The first article, published in *Dendrobiology* and entitled *Development of a SNaPshot assay for the genotyping of organellar SNPs in four closely related pines*, presents a newly developed genotyping method targeting a panel of 13 maternally inherited mitochondrial SNPs and a diagnostic chloroplast marker inherited paternally, which allows the identification of *P. sylvestris* versus *P. mugo* complex origin. The aim of this study was to improve and simplify existing genotyping protocols and validate them across 12 populations representing four pine species, including *P. sylvestris*, *P. mugo*, and *P. uliginosa*. The method demonstrated high accuracy and efficiency, enabling the identification of genotypes for 146 out of 147 tested individuals from the European range of the selected species. Twelve distinct mitochondrial haplotypes were identified, with the highest average diversity observed in *P. uliginosa* and *P. sylvestris*. The study also confirmed the specificity of the chloroplast diagnostic marker: all *P. sylvestris* individuals carried the "C" variant, whereas all individuals from the *P. mugo* complex carried the "A" variant at the *trnL-trnF* locus.

The second article, published in *BMC Plant Biology* and entitled *Hybridization has localized effect on genetic variation in closely related pine species*, describes an assessment of genetic variation in hybrid pine populations compared to allopatric reference populations, using 13 mitochondrial markers generated with the SNaPshot method and 14 nuclear microsatellite markers (nSSRs). Individuals from contact zones were classified into species groups based on nuclear genotypes (nSSRs), morphological features, and the diagnostic chloroplast DNA marker. Both nuclear and mitochondrial genetic variation were estimated using metrics such as allelic richness and haplotype richness—defined as the average number of alleles or haplotypes in a population normalized by the smallest population size. Additionally, a mitochondrial haplotype analysis was performed across multiple European populations to establish a database and network of observed variants. The results revealed no significant increase in genetic variation in the hybrid zones as would be expected; in fact, mitochondrial genome diversity was often reduced. No novel variants were found in the contact zones, which instead appeared to be homogenized to one or two haplotypes, commonly observed within the Polish range of the studied species. Nearby allopatric populations appeared unaffected by the processes in hybrid populations, indicating that hybridization has a localized effect.

The third chapter corresponds to an article currently under review in *Molecular Ecology*, entitled *Repeatable genomic outcomes along the speciation continuum: insights from pine hybrid zones (genus Pinus)*. This study represents the first attempt to use genomic data (7,390 SNP markers) to assess genetic differentiation, population structure, and patterns of genetic ancestry in the examined contact zones. Although genetically pure *P. mugo* individuals were rare, ancestry from this species dominated in every sympatric population. Moreover, genetic diversity—estimated using thousands of markers located in functional genes—was higher in hybrid individuals than in the allopatric populations of the parental species.

The fourth and final chapter presents an article under review in *BMC Plant Biology*, entitled *Molecular signatures of introgressive divergence and selection in contact zones of closely related pine species (Pinus genus)*. It focuses on selection processes and genomic introgression occurring in hybrid zones. Individuals were grouped into genetic classes based on the proportion of their genomic ancestry, calculated from 7,390 nuclear SNPs. Three commonly used methods—pcadapt, OutFLANK, and BayeScan—were applied to identify loci under selection (outliers). Markers identified by each method were compared and only loci identified as outliers by at least two methods were selected. The aim of the study was to determine which genes were subject to selection within specific species classes in hybrid zones, relative to allopatric reference populations. The results revealed considerably stronger selective pressure in *P. sylvestris*-like hybrids compared to individuals with predominantly *P. mugo* ancestry. Furthermore, signs of selection were also found in numerous genes of *P. sylvestris* individuals occurring in contact zones that exhibited pure genetic ancestry, the typical phenotype of the species, and the diagnostic chloroplast marker variant.

In summary, this dissertation demonstrates that hybridization, gene flow, and selection play distinct yet complementary roles in shaping genetic variation in the contact zones of the studied pine species. While neutral markers point to homogenization of gene pools, functional *loci* reveal increased genetic diversity and differentiation between hybrids and genetically pure *P. sylvestris* individuals in sympatric populations. The prevalence of *P. mugo* genomic background among most hybrids and the low selective pressure acting on such individuals suggest adaptive introgression and better environmental fit of *P. mugo*-like hybrids. In contrast, pure *P. sylvestris* individuals from contact zones appear to be subject to selection linked to adaptation to environmental conditions that are challenging for typical representatives of the species. The findings presented in the four scientific articles comprising this dissertation contribute to our understanding of hybridization as both an adaptive mechanism and a driver of speciation. New insights into the genomics of forest tree hybrid zones hold importance not only for evolutionary biology, but also have practical implications for nature conservation, genetic resource management, and biodiversity preservation in forest ecosystems facing increasing environmental pressures.