universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Evolution of the *Primula auricula* complex (Primulaceae) in the Eastern Alps"

verfasst von / submitted by

Michael Kurt Johann Janda, BSc

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

Master of Science (MSc)

Wien, 2021 / Vienna, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 832
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Master's degree programme Botany
Betreut von / Supervisor:	Assoz. Prof. Dipl.-Ing. Dr. Ovidiu Paun

Acknowledgements

First and foremost, I want to thank Ovidiu Paun for being the supervisor of my thesis and for guiding me from start to finish. I highly appreciate his expertise and knowledge and all the advice he was able to give me in every phase and aspect of my work.

I would also like to thank Karin Vetschera for initiating and Hochschuljubiläumsfonds der Stadt Wien for funding the project H-294808/2018.

I am very thankful for the contributions made by Christian Gilli and Lukas Grossfurthner by identifying and sampling the individuals. Christian further provided access to samples from the herbarium. Lukas supported me throughout the process of data analysis and helped me a lot with R.

I would like to thank Michael H. J. Barfuss, who was a tremendous help in the wet lab, especially by helping to create an extraction protocol that worked well with the samples.

Also, a big thank you goes to Clara Priemer who was able to help me with thought-provoking discussions towards the end of writing and provided some new information on the samples.

I also want to thank Anna-Sophie Hawranek and Dominik Metschina for helping me to answer some questions that emerged during the making of this thesis.

Last but not least I would like to express my deep gratitude to my parents, grandmother, and wife who supported and encouraged me throughout my studies.

Abstract

Recent efforts to elucidate the evolutionary relationships within *Primula* sect. *Auricula* - the largest endemic plant group of the European Alpine system - based on molecular methods, led to the separation of one of its most widespread species, *Primula auricula*, into two allopatric species, *P. auricula* and *P. lutea* largely occupying southern versus northern distribution areas. Traditionally, *Primula auricula* formed a morphologically diverse species complex that contained two subspecies, *ssp. balbisii* and *ssp. auricula*, known to occur largely in sympatry, but at different altitudes. The morphological characters used to identify either of these subspecies are very similar to those assigned to the two novel species. In this work we investigate individuals of the *P. auricula* complex and its two subspecies from two mountains in the Eastern Austrian Alps along altitudinal gradients, plus 12 herbarium accessions from elsewhere in the Alps. By employing RAD-sequencing, relationships among the samples could be inferred. It could be confirmed that individuals with morphological characters now associated with allopatric *P. auricula* and *P. lutea* can be found on the same mountain whilst not differing genetically. An Isolation-by-Distance analysis has shown that relatedness in the *P. auricula* complex strongly correlates with geographical proximity. Despite the lack of genetic differentiation between the subspecies, photosynthetic measurements taken in-situ indicate a subspecies-specific response to the amount of light available. However, the relative contribution of constitutive versus plastic factors to the observed phenotypic variation remains a subject for future research.

Zusammenfassung

Jüngste Bestrebungen die evolutionäre Beziehung innerhalb von *Primula* sect. *Auricula*, der größten endemischen Pflanzengruppe des Europäischen Alpensystems, basierend auf molekularen Methoden zu klären, führte zur Teilung einer ihrer weitverbreitesten Arten, *Primula auricula*, in zwei neue Spezies, *P. auricula* und *P. lutea*, die jeweils das nördliche bzw. südliche Verbreitungsgebiet besiedeln. Ursprünglich bildete *P. auricula* einen morphologisch diversen Artenkomplex, der die zwei, großteils auch sympatrisch, aber in unterschiedlichen Höhenstufen vorkommenden, Unterarten ssp. *balbisii* und ssp. *auricula* beinhaltete. Die bestimmungsrelevanten Merkmale dieser Subspezies ähneln stark jenen, die den zwei neuen Arten zugeordnet wurden. In der vorliegenden Arbeit werden Individuen der zwei Unterarten des *P. auricula* Komplexes von zwei Bergen in den östlichen österreichischen Alpen entlang eines Höhengradienten, sowie 12 Herbarbelege aus den Ostalpen untersucht. Die Verwandtschaftsverhältnisse innerhalb der Stichprobe wurden mittels RAD-sequencing untersucht. Es konnte bestätigt werden, dass sich Individuen, mit Merkmalen, die nun den allopatrisch vorkommenden Arten *P. auricula* und *P. lutea* zugeordnet werden, aber auf dem gleichen Berg vorkommen, genetisch nicht unterscheiden lassen. Eine Isolation-by-Distance Analyse konnte zeigen, dass der Grad der Verwandtschaft zwischen Individuen stark mit geographischer Nähe korreliert. Trotz fehlender genetischer Differenzierung zwischen den Subspezies deuten in-situ vorgenommene Photosynthese-bezogene Messungen auf eine Unterarten-spezifische Reaktion auf die Menge des verfügbaren Lichtes hin. Jedoch bleibt der relative Beitrag konstitutiver versus plastischer Faktoren, in Bezug auf die beobachtete phänotypische Variation, Gegenstand zukünftiger Forschungen.

Table of Contents

1. Introduction

The genus *Primula* within the family Primulaceae contains around 430 species that are common in the cooler habitats of the northern hemisphere (Richards, 2003). Its centre of diversity is found in the Himalayas region in Asia where approximately 75% of its species occur (Richards, 2003). In Europe, however, only about 34 species can be found mostly in Alpine habitats (Richards, 2003), of which 26 belong to the section *Auricula* (commonly referred to as Auriculas), ranging from the Cantabrian Mountains in northern Spain to the mountain ranges of the Balkans (Boucher et al., 2016; Aymerich et al., 2014). As Alpine endemics, Auriculas have been of particular interest as a model to investigate the history of the flora of the European Alpine system (Boucher et al., 2016; Zhang et al., 2004; Ozenda, 1995). Additionally, in the last two decades, a great effort has been made to resolve the phylogeny and species delimitations within sect. *Auricula* based on molecular approaches. This was initiated by Zhang (2002) and Zhang and Kadereit (2004, 2005) using ITS and AFLP data, and later revisited by Boucher et al. (2016) with RAD-seq (Restriction-site associated DNA sequencing) data. This confirmed the monophyly of *Primula* sect. *Auricula*, as was already proposed by different studies that focused on the whole genus (de Vos et al., 2014; Mast et al., 2001), as well as its relatively recent emergence as a group (around 3.6 to 6.6 mya). In Boucher et al. (2016) four delimited clades were proposed in opposition to the more traditional separation into seven subsections (Smith and Fletcher, 1949; Zhang, 2002), whilst Zhang and Kadereit (2004) suggested two clades. Auriculas are ancestrally hexaploid species with $2n = 6x = 66$, although hypohexaploid individuals are also known (as cited in Zhang, 2002). The origin of sect *Auricula* most likely lies in East Asia, but a North American origin has also been considered (Boucher et al., 2016; Zhang et al., 2004).

The species of interest in this work is one of the most widespread of section *Auricula* - the eponymous *Primula auricula*. This forms a species complex of morphologically diverse and taxonomically difficult entities with an extensive distribution area that covers the entirety of the European Alps and extends into the Tatra Mountains and parts of Hungary, with additional occurrences in the South-West Carpathians, Apennines and Balkans (Boucher et al., 2016; Zhang and Kadereit, 2004; Smith and Fletcher, 1949). Amongst the decisive diagnostic characters is the density and length of glandular hairs on the leaves, calyx and flower stalks, as well as the presence of a farina (i.e. a mealy excretion) that is exudated by them, although individuals without farina tend to have longer hairs and vice-versa (Smith and Fletcher, 1949). In the past, the species had been considered either highly polymorphic without further subdivisions, or divided into subspecies and varieties of unclear geographic boundaries (as cited in Aymerich et al., 2014). Together with the newly described *Primula subpyrenaica* in Catalonia, Spain (Aymerich et al., 2014), the *P. auricula* complex is the only

yellow-flowered group in the section. Based on their recent molecular analyses however, Zhang and Kadereit (2004, 2005) were able to genetically distinguish two separate species in the complex, granting two former subspecies species rank. They concluded that the new species must be allopatric, with the novel species *P. lutea* occupying the "northern" areas of traditional *P. auricula* (Western and Northern Alps and Tatra Mountains) and novel species *P. auricula* populating the "southern" part of the original range (Eastern and Southern Alps, Apennines, Carpathians and Balkans). This was again ascertained by Boucher et al. (2016). The geographical boundaries presented by the authors remained imprecise (Boucher et al., 2016; Somlyay and Bauer, 2010). Zhang and Kadereit (2004) also designated specific morphological characters to the new species, describing *P. lutea* as a sparsely- and short-haired (<0.2mm), usually efarinose, taxon and *P. auricula* as densely- and long-haired (>0.2mm) and usually farinose. This strict linkage they established between geography and distinct diagnostic characters in particular has not remained undisputed, not least because of the morphologically variable nature known and described in this species complex (Boucher et al., 2016; Aymerich et al., 2014; Somlyay and Bauer, 2010, Smith and Fletcher, 1949). On at least one occasion since the separation of traditional *P. auricula* into *P. auricula* and *P. lutea* it could be shown that long-haired individuals occur in the suggested distribution area of short-haired *P. lutea* and vice-versa (Somlyay and Bauer, 2010).

In this work, populations of *Primula auricula* on two altitudinal transects on two mountains in the Eastern Austrian Alps are examined and their genetic structure and relationships are being investigated. Identification of individuals in the field is based on the Excursion Flora for Austria, Liechtenstein and Southern Tyrol by Fischer, Oswald and Adler (2008), which uses the traditional taxonomy that sees one species *Primula auricula* separated into two morphologically distinct subspecies: *P. auricula* ssp. *balbisii* and *P. auricula* ssp. *auricula*. These subspecies are described to be mostly sympatric in the region although separated by altitude to some degree, with the former being predominantly found at lower elevations. Morphologically, *P. auricula* ssp. *balbisii* is very similar to the novel species *P. auricula*, as is ssp. *auricula* to *P. lutea*, concerning the density and length of glandular hairs and colour intensity of the corolla, albeit not necessarily in terms of abundance of the farina (Zhang and Kadereit, 2004). As a consequence of this, the subspecies have recently become synonymous to their novel counterparts at the species level in the excursion flora, which has led to further exacerbation of an already difficult species complex. In this work, sequencing of individuals uses the state-of-the-art RAD-sequencing approach (Baird et al., 2008) for inference of relationships. Twelve additional accessions from herbarium material that have been identified as *P. auricula* at their individual time of sampling and cover broadly the Eastern Alps are included in the genetic analysis. Furthermore, photosynthetic measurements were carried out on wild individuals in-situ with a field fluorometer/spectrometer to quantitatively characterise the phenotypes of the sampled

individuals. The over-all aim of this work is to contribute to the further understanding of the complicated *Primula auricula* complex. For this, a particular emphasis is directed towards whether genetically differentiated groups can be identified reflecting the morphologically characterised subspecies *sensu* Zhang and Kadereit (2004).

2. Material and methods

2.1. Sampling and identification

Individuals of *Primula auricula* were identified in the field based on the determination key in Fischer, Oswald and Adler (2008) and classified accordingly as either of its two described subspecies, ssp. *balbisii* and ssp. *auricula*. The diagnostic characters upon which the subspecies were identified in the field are the density and length of glandular hairs on the leaves and the existence of a farina on the leaves, flower stalks and calyx, as well as the colour intensity of the flower and the presence of a scent of the same (Table 1, Fig. 1). Since not all individuals were in bloom at time of sampling, farina and glandular hairs were primarily used for identification. Individuals that exhibited intermediate morphology, i.e. couldn't be unambiguously identified as either subspecies in-situ, were recognised as potential hybrids and will be called indets (*indeterminable*) henceforth.

Table 1. Diagnostic characters to identify the two subspecies of *Primula auricula.*

	Primula auricula* ssp. *balbisii	***Primula auricula* ssp.*auricula***
Glandular Hairs	Hairs between 0.25-0.4(0.5) mm long, leaves densely covered	Hairs between 0.1-0.2 mm long, leaves only sparsely covered
Farina	Plant not farinous	Leaves, flower stalk and calyx farinous
Flowers	Dark yellow, usually unscented	Light yellow, scented

A total of 204 accessions of *Primula auricula* were sampled from fifteen different locations in the Central and Eastern Alps, with the focal point on the sites of Schneeberg along the Herminensteig hiking trail to the summit and at Leopoldsteinersee along the Hochblaser hiking trail where 101 and 90 individuals have been collected, respectively. Along the trails, populations of *P. auricula* were so chosen that, when possible, sampling stations were around 100m apart in terms of elevation, in order to obtain good representation of the two subspecies along an altitudinal gradient.

Fig. 1. Habitus of Primula *auricula* ssp. *balbisii* (left) and ssp. *auricula* (right). © L. Grossfurthner

Depending on size, we collected one to few leaves from each individual, thereby keeping the plants mostly intact. Sampled leaves were zip-bagged and subsequently dried with silica-gel. This way, we were able to collect tissue from an average of ten individuals per sampling station, resulting in nine sites at Lake Leopoldsteinersee and ten sites on Mount Schneeberg at altitudes between 700-1750m and 1200-2000m, respectively (for a full list see Table A in Appendix). The remaining 12 accessions were herbarium specimen from single locations ranging from eastern Switzerland and north-western Italy to eastern Austria and have been provided by the Department of Botany and Biodiversity Research (Fig. 2; for a list of the herbarium samples see Table B in Appendix).

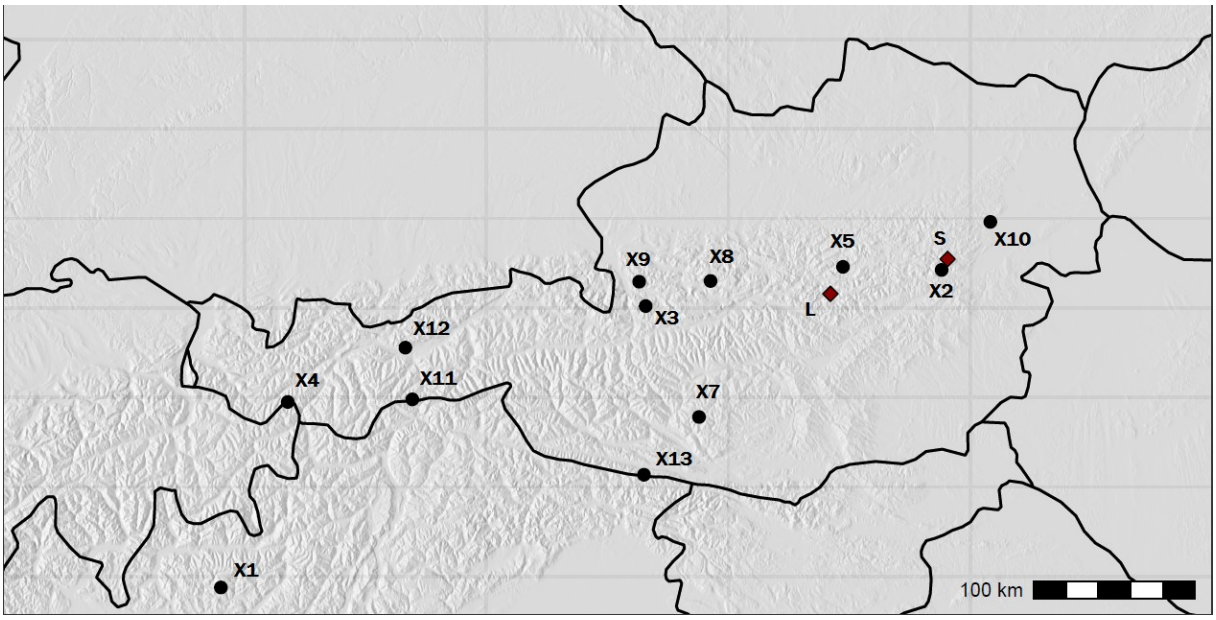

Fig. 2. Sampling sites for this study. Black dots depict the locations of herbarium samples; focal points of Schneeberg (S) and Leopoldsteinersee (L) are represented by red diamonds.

2.2. DNA extraction and RAD-seq library preparation

Several extraction protocols have been tried, many of which yielded unsatisfying or insufficient amounts of DNA potentially due to the high amount of secondary metabolites found in *Primula auricula*. Finally, a custom extraction protocol modified from the Stratec Invisorb Spin Plant Mini Kit (Stratec Molecular, Berlin, Germany) protocol was chosen, that included several additional cleaning steps. For total DNA extraction, 20mg±2mg of silica-dried leaves, or leaves taken from herbarium specimen (i.e. press-dried), were frozen with liquid nitrogen and subsequently ground with a Tissue Lyser II (Qiagen, Hilden, Germany) for two cycles of 2 min. To remove secondary metabolites, the ground tissue was incubated three times in 1,000µL of a Sorbitol solution (Sorbitol, 2% Polyvinylpyrrolidon, 1% β-Mercaptoethanol) for 10mins on ice and centrifuged for 10mins in a cold centrifuge (10,000 x g) disposing of the supernatant after each turn. The amount of Lysis Buffer for the next step was increased by a factor of 1.5 to 600µL and the incubation time was raised from 30 min to 60 min at 65°C. After this, another cleaning step was introduced by adding 400µL of SEWAG (Chloroform:Isoamyl alcohol, 24:1) to the Lysis mix. Samples were incubated on ice for 10mins, shaking regularly to assure complete mixing of all phases, before centrifuging in a cold centrifuge (10mins, 10,000 x g). The supernatant was transferred to the prefilter of the Invisorb Spin Plant Mini Kit and centrifuged for 1min at 11,000 x g. To the flow-through 4µL of RNase (100mg/mL) was added and the solution incubated at room temperature for 5mins. The amount of Binding Buffer was increased to 300µL from 200µl. Washing steps followed the extraction kit's manual, but two additional Ethanol washes (EtOH 96%, 550µL, 10.000 x g for 1min) followed. Elution of DNA was performed twice with 60µL prewarmed Elution Buffer, otherwise following the extraction kit manual. Quantity of DNA was assessed using a QuBit 2.0 Fluorometer (Thermofisher Scientific, Waltham, MA, USA). A total of 138 samples were finally included in the RAD-seq library preparation, i.e., 72 from Mount Schneeberg and 54 from Lake Leopoldsteinersee, averaging seven and six individuals per altitudinal station, respectively, as well as 12 herbarium samples.

RAD-sequencing followed a modified protocol starting from Paun et al. (2016). In total, two libraries were prepared; the first included 72 individuals solely from Mount Schneeberg, and the second 66 individuals from Leopoldsteinersee and the herbarium. For each sample, 150ng of gDNA (genomic DNA) was used as starting material. The samples were digested with PstI at 37°C for two hours, before heat inactivation of the enzyme at 80°C for 20mins. The resulting fragments were ligated with P1 adapters overnight at 16°C, before heat inactivation of the ligase at 65°C for 10mins. After pooling barcoded accessions into sub-libraries, the resulting DNA fragments were further broken up by sonication with a Bioruptor Pico (Diagenode SA, Belgium) with two cycles of 45s on/60s off. After sonication, pools were purified with a Qiagen MiniElute Reaction clean up column and then double-

size selected with SPRI beads (0.55x on the right and 0.7x on the left). The ends of the DNA fragments were polished using the Quick Blunting Kit (NEB, room temperature for 30mins) and again cleaned with a Qiagen MiniElute Reaction clean up column. Samples were prepared for the ligation of a second adapter by adding overhangs with (15U) Klenow exo- at 37°C for 30mins before another purification. The sub-libraries were quantified with a QuBit 2.0 Fluorometer to adjust the amount of DNA from each pool to the lowest value measured (i.e. it is necessary to continue with the same total amount for every sub-library). Distinct P2 adapters were ligated to the fragments for each sub-library at room temperature for 30mins. Both P1 and P2 adapters contain unique barcode sequences whose combinations facilitate recognition of individual samples after Illumina sequencing. Before amplification, sub-libraries were pooled and size selected (left side, 0.7x SPRI) and eluted in 50µL H_20. A total of 12µL elute was used for six PCR reactions and the amplified products checked on a 1% TBE gel alongside a positive control. Then, PCR products were cleaned up with a QIAquick PCR purification kit and size selected (left side with 0.7x SPRI, elute in 30µL H_20). Concentration of DNA was measured both before and after size selection with a QuBit 2.0 Fluorometer. In addition, samples were analysed with a Bioanalyzer (Agilent) with a High-Sensitivity DNA-Kit for a final quality check. The libraries were diluted to 5ng/µL and sent for sequencing in a Illumina HiSeq machine unit of the Vienna Biocenter Facility (www.vbcf.at/ngs).

2.3. Bioinformatics pipeline

Raw Illumina reads were demultiplexed first by index reads with BamIndexDecoder v.1.03, available from the Picard Illumina2Bam package (https://github.com/wtsi-npg/illumina2bam). The output bam files were converted to fastq files with Picard v.2.6.0 (available from https://github.com/broadinstitute/picard) and then filtered with process_radtags from STACKS 1.47 (Catchen et al., 2013) based on their inline barcodes. The step in process_radtags cleaned reads by removing those with an uncalled base, or if having low score, as well as rescue barcodes and cut sites with a maximum of one mismatch. The processed reads were further assembled *de novo* into a catalogue of loci using the denovo_map.pl wrapper from STACKS. Parameters were set to require at least four reads to create a stack (-m), allow a maximum one mismatch between stacks of an individual to create a locus (-M), and accept a maximum of two mismatches between stacks of two or more individuals (-n), whilst also allowing for indels to be considered (-gapped). These parameters were optimized by testing several different combinations, as yielding the highest number of loci that contain one to ten SNPs per locus occurring in at least 70% of individuals. We then prepared a pseudoreference from the created catalogue using the function export_sq.pl of Stacks by including as different 'chromosomes' haplotype consensuses of loci in a fasta file, following Heckenhauer et al. (2018).

Further, mapping of the samples was tried against two different references, an available genome of the congeneric *Primula vulgaris* (Cocker et al., 2018) and the pseudoreference that was constructed from the STACKS catalogue. In both cases a similar approach was applied that started by employing the BWA software package v.0.7.5 (Li and Durbin, 2010). This included an initial step of indexing the references with the commands *bwa index -a bwtsw* for the whole genome of *Primula vulgaris* and *bwa index -a is* for the pseudoreference. For mapping against the references, the *bwa mem* function with the option -M was used and was followed by sorting the aligned sample files by coordinates with Picard v.2.6.0 (available from https://broadinstitute.github.io/picard/), converting them to the bam format in the process. As the mapping rates were found to be considerably higher for mapping against the pseudoreference (see results), we continued with these mapping files hereafter.

Genotype likelihoods were inferred using the program ANGSD v.0.929 (Korneliussen et al., 2014) based on a GATK model (i.e., -GL 2). The software was chosen as a first analysis of the hexaploid dataset as it may reflect uncertainties of genotypes better than other methods, despite being developed exclusively for diploids (Záveská et al., 2019). The parameters were set to keep only positions with high likelihoods (SNP_p-val < 2e-6), a mapping and base quality threshold of 20 (minMapQ 20, minQ 20) and those that contained at least half of the individuals (minInd 69). In addition, positions were only retained when the minor allele was found in at least two individuals (minMaf 1.4%). Based on the genotype likelihoods estimated in ANGSD a PCA analysis was performed with PCAngsd (Meisner and Albrechtsen, 2018) to produce covariance matrices. Subsequently, each matrix was plotted as heatmaps in R v.3.6.3 to visualise population structure.

In order to take the hexaploid nature of *Primula auricula* into account, a SNP genotyping approach for polyploids using low-coverage sequencing data (polyploid genotyping) by Blischak et al. (2018) was chosen. As a first step, the workflow requires both the creation of an index using the *faidx* command of SAMtools v.1.3 (Li et al., 2009) and that of a sequence dictionary using Picard v.2.2.1 for the reference genome. Read group information was added with Picard v.2.6.0 and variant calls of the mapped reads were produced by running the GATK unified genotyper (McKenna et al., 2010) with the ploidy set to 6 (-ploidy 6). A pileup file was generated using the mpileup command from SAMtools v.1.3 (Li, 2011). Subsequently, filtering of the sites was achieved running an R script with settings as follows: variant quality score (QUAL score) greater than 100, a minimum read depth (DP) of at least five reads per site, and allowing only for a maximum of 13 missing individuals per position. Less stringent settings for missingness have been tried but were found to retain too many loci, rendering the next steps computationally expensive. In order to extract the total and alternative read counts, as well as the per locus error rates, three different python scripts were run using Python v.2.7.6. Genotype likelihoods were then calculated with 1,000 iterations using the program EBG

(Blischak et al., 2018) based on a GATK-like model. A workflow for polyploid genotyping including helper scripts for preparation of the raw data, as well as the program EBG for subsequent inference of likelihoods are available on https://github.com/pblischak/polyploid-genotyping.

Pairwise relatedness was then estimated with PolyRelatedness v.1.9 (Huang et al., 2014) under a maximum-likelihood model using the Ritland 1996 estimator (Huang et al. 2015). Since PolyRelatedness is limited to 65,536 loci in an analysis, the polyploid genotyping output had to be filtered by excluding any sites with missing data, whilst also reformatting it into a common genotype format suitable for the program using a custom R script following Grossfurthner et al. unpub. The resulting matrix was plotted in R v.3.6.3 using the heatmap.2 function contained in the gplots package v.3.1.1 (available from http://CRAN.R-project.org/package=gplots). The filtered PolyRelatedness input file was further used to perform a Principal Component Analysis (PCA) in R v.3.6.3.

For inference of population structure using multilocus data the program STRUCTURE v.2.3.4 (Pritchard et al., 2000) was employed. The output from polyploid genotyping was filtered, removing any sites with missing data, and only one random SNP per RADtag were further retained as putatively unlinked loci. Parameters for this analysis were set to assume between one and six populations (K=6) with ten runs for each K and 1,000,000 replicates after a burn-in of 100,000. To decide for the best scenario (K), the Evanno method was employed (Evanno et al., 2005). Visualization as bar plots was achieved using the POPHELPER package v.2.3.1 (Francis, 2017) in R v.3.6.3.

Isolation-by-distance (IBD) was tested using a Mantel test from the genetic and geographic distance matrices using the adegenet package (Jombart, 2008); the obtained relationship was further plotted (Jombart & Collins, 2015) in R v.3.6.3. Euclidean genetic distances were calculated from the GPS data obtained during sampling in the field. The genetic distances have been calculated from the filtered STRUCTURE input file using the program POLYGENE (Huang et al., 2020).

2.4. Photosynthesis data

A MultispeQ V2.0 field fluorimeter by PhotosynQ was used for photosynthetic measurements on Mount Schneeberg and Leopoldsteinersee. These measurements include fluorescence and ambiance based parameters like Linear Electron Flow (LEF), Photosynthetic active radiation (PAR), altitude and ΦII (yield of Photosystem II) among others. The device was applied before any leaves were collected in order to avoid any aberrations caused by handling the individuals. Measurements were taken over the course of a day at each location, starting in the morning at lower altitudes and completed in the afternoon at higher altitudes. The obtained data was automatically uploaded and stored in the

PhotosynQ network and was accessed via their webpage (https://photosynq.org). Data was plotted in R v.3.6.3 using the ggplot2 package (Wickham, 2016).

3. Results

3.1. Sampling

A substantial portion of sampled individuals on our focal points Schneeberg and Leopoldsteinersee were classified as indets, i.e. intermediates between the two subspecies *Primula auricula* ssp*. balbisii* and *P. auricula* ssp. *auricula* that could not be unambiguously identified in the field, making up 37.2% or 71 individuals of the dataset. Specimen identified as ssp. *auricula* are represented by 80 accessions (41.9%) and ssp. *balbisii* by 40 accessions (20.9%). Very similar proportions occur in the portion of the dataset that was ultimately chosen for sequencing (Table 2).

Table 2. Overview of sampled individuals and their share in the dataset. Numbers in brackets show the amount of individuals eventually chosen for sequencing.

	Stations/populations (range of altitude)	*ssp. balbisii*	indet	*ssp. auricula*	total
Schneeberg	1-5 (1200m - 1550m)	28 (*18*)	21 (*16*)	2 (*1*)	101 (*72*)
	6-10 (1650m - 2000m)	3 (*3*)	12 (*8*)	35 (*26*)	
Leopoldsteinersee	1-5 (700m - 1100m)	9 (*5*)	38 (*24*)	3 (*1*)	90 (*54*)
	6-9 (1500m - 1750m)	0 (*0*)	0 (*0*)	40 (*24*)	
	Total sampled	40 (20.9%)	71 (37.2%)	80 (41.9%)	191
	Total sequenced	*26 (20.6%)*	*48 (38.1%)*	*52 (41.3%)*	*126*

As Table 2 indicates, individuals identified as ssp. *balbisii* were predominantly found in the lower five sampling stations of each mountain, whereas individuals that morphologically fitted ssp. *auricula* were clearly dominant at higher altitudes. This pattern becomes very evident when the subspecies are listed for every station individually (Table 3), where numbers of individuals from ssp. *balbisii* gradually decrease with increasing altitude, being replaced first by indets and, ultimately, ssp. *auricula*. On both mountains the indets occupy an intermediate elevation. In this context it has to be noted however, that a large gap occurs between station 5 and 6 on Leopoldsteinersee in terms of altitude, whereas, in this regard, sampling stations are more or less spread out evenly on Schneeberg.

Table 3. Amount of ssp. *balbisii*, ssp. *auricula* and indets in every station. Dominant groups are highlighted in grey.

Schneeberg 1-10

ssp. balbisii	indet	*ssp. auricula*	*altitude (m)*
10	*0*	*0*	1200
6	4	*0*	1350
7	3	*0*	1450
1	**8**	*1*	1450
3	**6**	*1*	1550
2	**4**	**4**	1650
1	**6**	3	1750
1	2	**8**	1850
0	*0*	**10**	1850
0	*0*	**10**	2000

Leopoldsteinersee 1-9

ssp. balbisii	indet	*ssp. auricula*	*altitude (m)*
6	4	*0*	700
2	**8**	*0*	800
0	**10**	*0*	950
0	**9**	1	1050
1	**7**	2	1100
0	*0*	**10**	1500
0	*0*	**10**	1550
0	*0*	**10**	1650
0	*0*	**10**	1750

3.2. Sequencing, SNP calling, mapping and filtering

The number of quality reads retained per sample ranged from 427,883 (Ps4_3) to 4,253,569 (X11), averaging to 1,583,945 reads. The *Primula vulgaris* reference genome contained 67,491 contigs to a total length of 4.1 Gb, which resulted in mapping rates from 17.5% to 38.0%. The pseudoreference created from the STACKS catalogue comprised 59,858 loci and a total of 5.6 Mb. The mapping rates for the pseudoreference were generally higher than those for the *P. vulgaris* genome, ranging from 37.8% (X3) to 73.2% (Ps1_9). Read depth lied between 5.1 (Ps4_3) and 45.7 (X12) for the reads mapped against the pseudoreference, averaging at around 17.5.

Running ANGSD with both the mapped reads from the pseudoreference and those from the *Primula vulgaris* reference resulted in 275,250 and 29,473 retained variant sites, respectively. In order to maximise the information content we decided to continue all further analyses with the pseudoreference. Initially, the polyploid genotyping approach yielded 242,154 sites after running its filtering script. The number of SNPs retained after excluding those with missing data was 27,816, and this data was further used for analyses with PolyRelatedness.

3.3. Population structure

The resulting coancestry heatmaps (Fig. 3), generated from the ANGSD and PolyRelatedness matrices, both show that the individuals from Schneeberg and Leopoldsteinersee are strongly delimited from each other. The samples from each mountain are considerably homogenous, although for Schneeberg the individuals exhibit slightly more variation. On both mountains, some stations exhibit slightly higher levels of relatedness within, which is true regardless of assigned subspecies. Both the ANGSD and PolyRelatedness analysis exhibit these patterns. The herbarium samples primarily tend to cluster together or associate more with Leopoldsteinersee. However, individuals X2, X3, X5 and X10 form an exception, since they tend to relate comparatively more to individuals from either Schneeberg (X2, X5, X10) or Leopoldsteinersee (X3) rather than other herbarium samples. Individuals X2 and X5, as well as X10, are geographically very close to both Schneeberg and Leopoldsteinersee and X3 is closer to Leopoldsteinersee (Fig. 2). One has to be reminded, that the ANGSD analysis treated the hexaploid dataset as diploids and is arguably less reliable than the PolyRelatedness matrix (Fig. 3).

The Principal Component Analysis (PCA) of the first two eigenvectors show a similar distinction of Leopoldsteinersee and Schneeberg that delimits them clearly from each other with no significant pattern within each mountain (Fig. 4). Also, herbarium samples X2, X3, X8, as well as X5 and X10 cluster with accessions from Leopoldsteinersee and Schneeberg, respectively. This is different to the heatmaps, where X2 was arguably more related to individuals from Schneeberg. The PCA of the second and third eigenvectors groups the aforementioned herbarium samples and both mountains together. In total, the first three eigenvectors explain 14.3% of variation (Fig. 4).

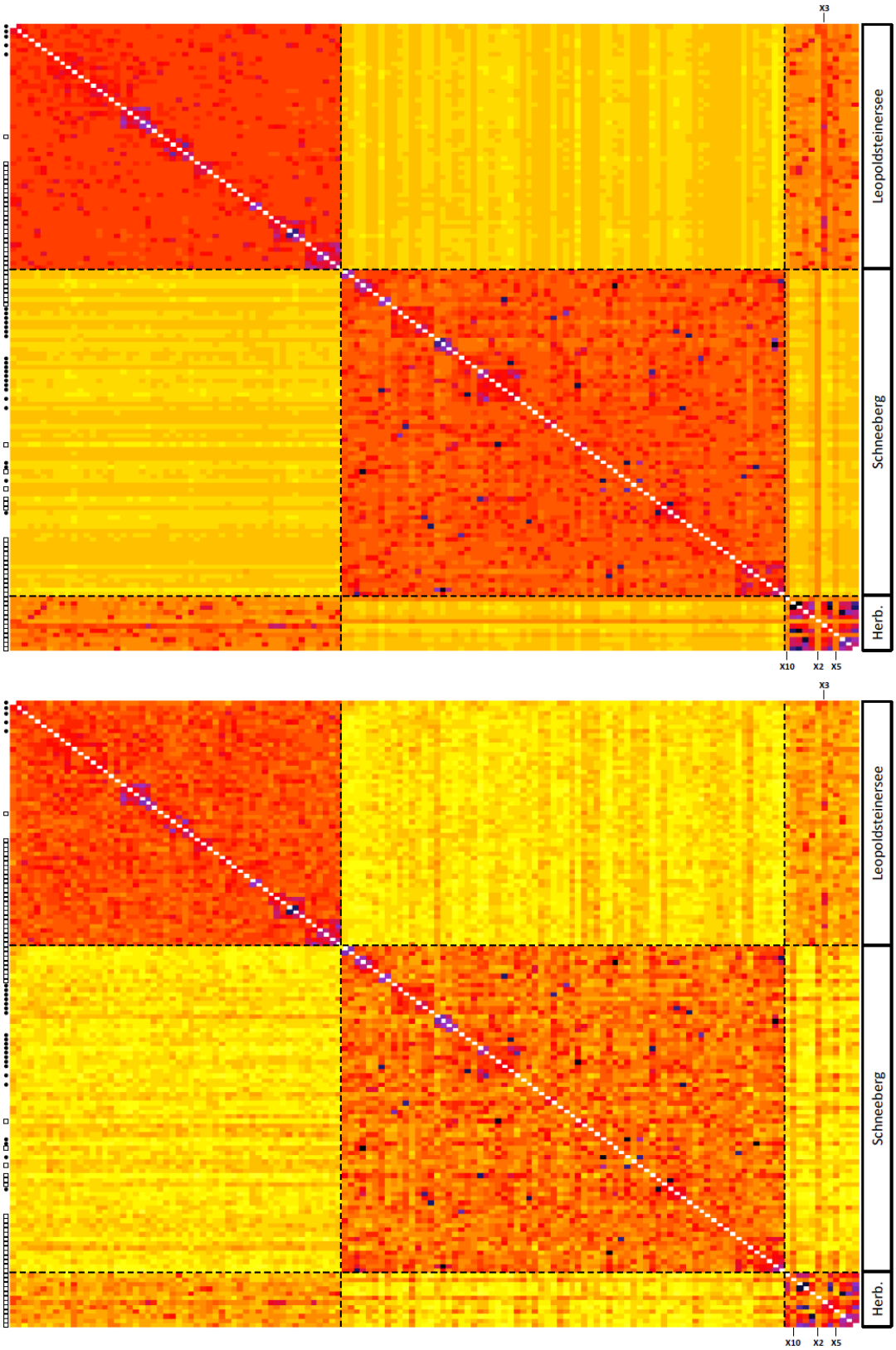

Fig. 3. Coancestry heatmaps generated from the ANGSD (top) and PolyRelatedness (bottom) analysis. Brighter colours show lower levels of relatedness, darker colours higher levels. Black dots, ssp. *balbisii*; white squares, ssp. *auricula*; unmarked individuals are indets. Values on the diagonal have been removed to increase contrast.

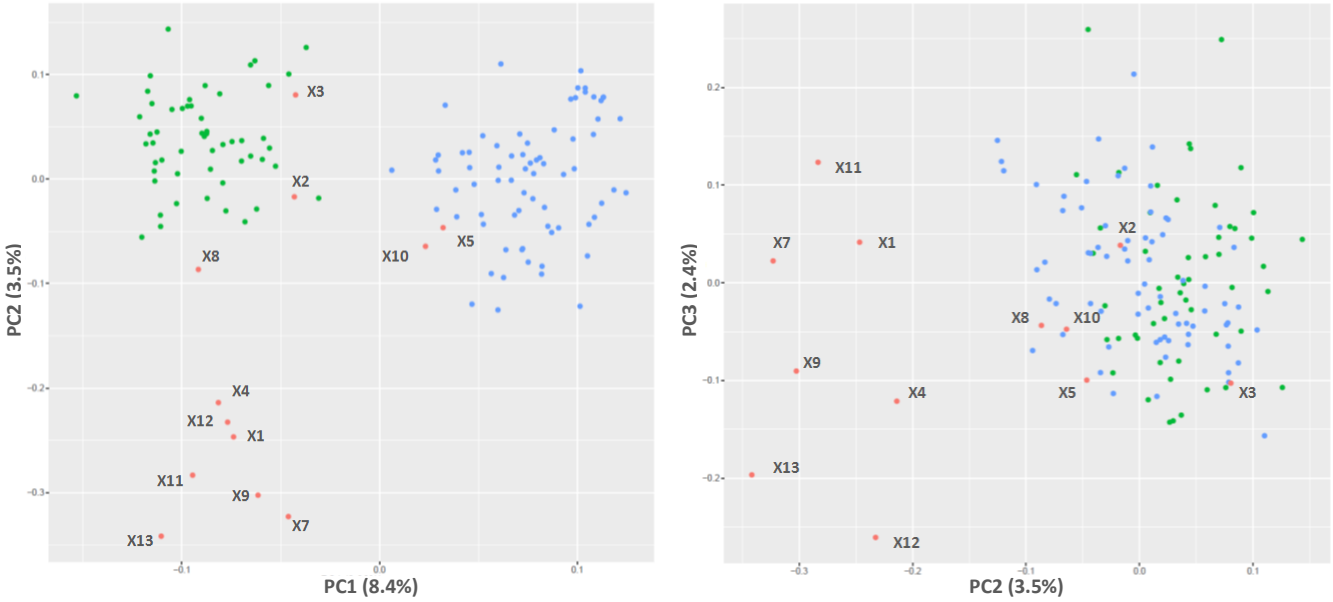

Fig. 4. Principal Component Analysis (PCA) of the first two (left) and the second and third eigenvectors (right), based on results from polyploid genotyping. Green dots, samples from Leopoldsteinersee; blue dots, accessions from Schneeberg; red dots, herbarium samples.

The amount of loci that were left for the STRUCTURE analysis after filtering and chosing one SNP per RAD locus was reduced to a total of 2,560. The Evanno method indicated K=2 being the optimal and K=3 being suboptimal (Fig. 5). Plots with four predefined populations or more showed the accumulation of 'ghost' clusters (i.e. clusters that are not represented by any full individual). As seen before, the individuals show genetic clustering by mountain and lack any significant structure within groups (Fig. 5). There is some admixture from Leopoldsteinersee in individuals from Mount Schneeberg in both K=2 and K=3. The cluster created by individuals from Leopoldsteinersee is very homogenous apart from a small fraction that is associated with the herbarium samples or Mount Schneeberg. This compares well with the pattern already observed in the plotted relatedness matrices (Fig. 3). Herbarium specimen constitute a separate cluster in K=3 with high levels of admixture from Leopoldsteinersee. When assuming only two populations in the dataset, this influence from Leopoldsteinersee becomes even more prominent. In K=3, it can be seen that only three individuals in the herbarium cluster exhibit introgression from Schneeberg, which are X2, X5, X10. Simultaneously, X3 shows a very strong influence from the Leopoldsteinersee cluster. The only individual that is entirely comprised of this herbarium cluster is X11, which was collected near the Brenner pass in Tyrol.

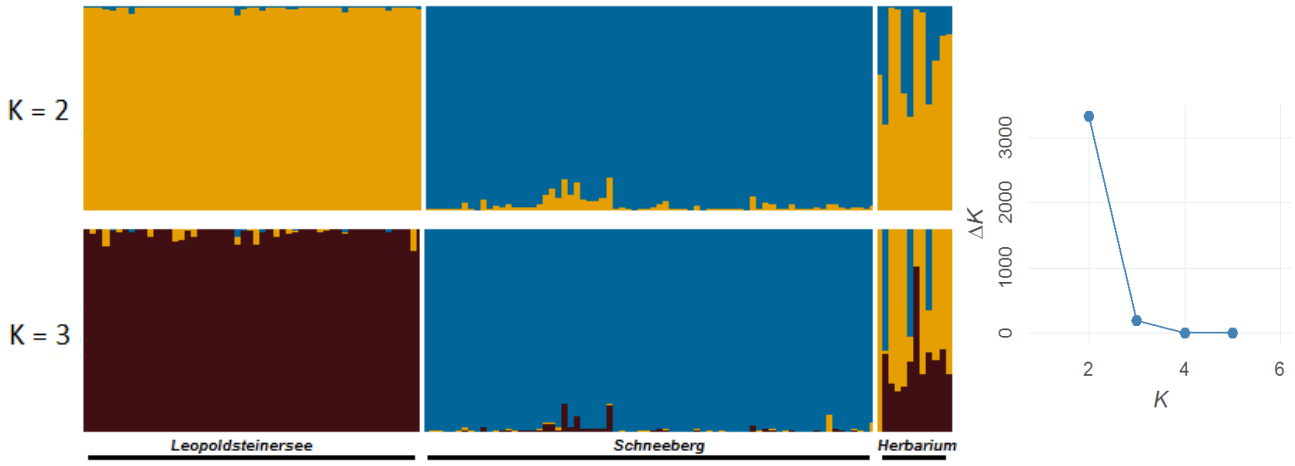

Fig. 5. Left, barplots of the STRUCTURE analysis when assuming two (top) and three populations (bottom). Right, result from the Evanno Method for DeltaK to choose best scenario.

3.4. Isolation-by-Distance

When testing the correlation between the genetic and geographical distances within each mountain, and between all individuals including herbarium samples, for Isolation-by-Distance (IBD) the underlying Mantel test is highly significant in all cases (p-value = 0.001). On both Schneeberg and Leopoldsteinersee, individuals are stretched out on the geographical axis, with very little deviations in genetic distances (Fig. 6). In each plot, one cluster with minute levels of genetic and very little geographical distance exists. The herbarium samples exhibit higher genetic distances in this analysis and, geographically speaking, are more spread out. The regression lines in each plot represent the range of genetic distances between the two closest individuals of the plot and those farthest away.

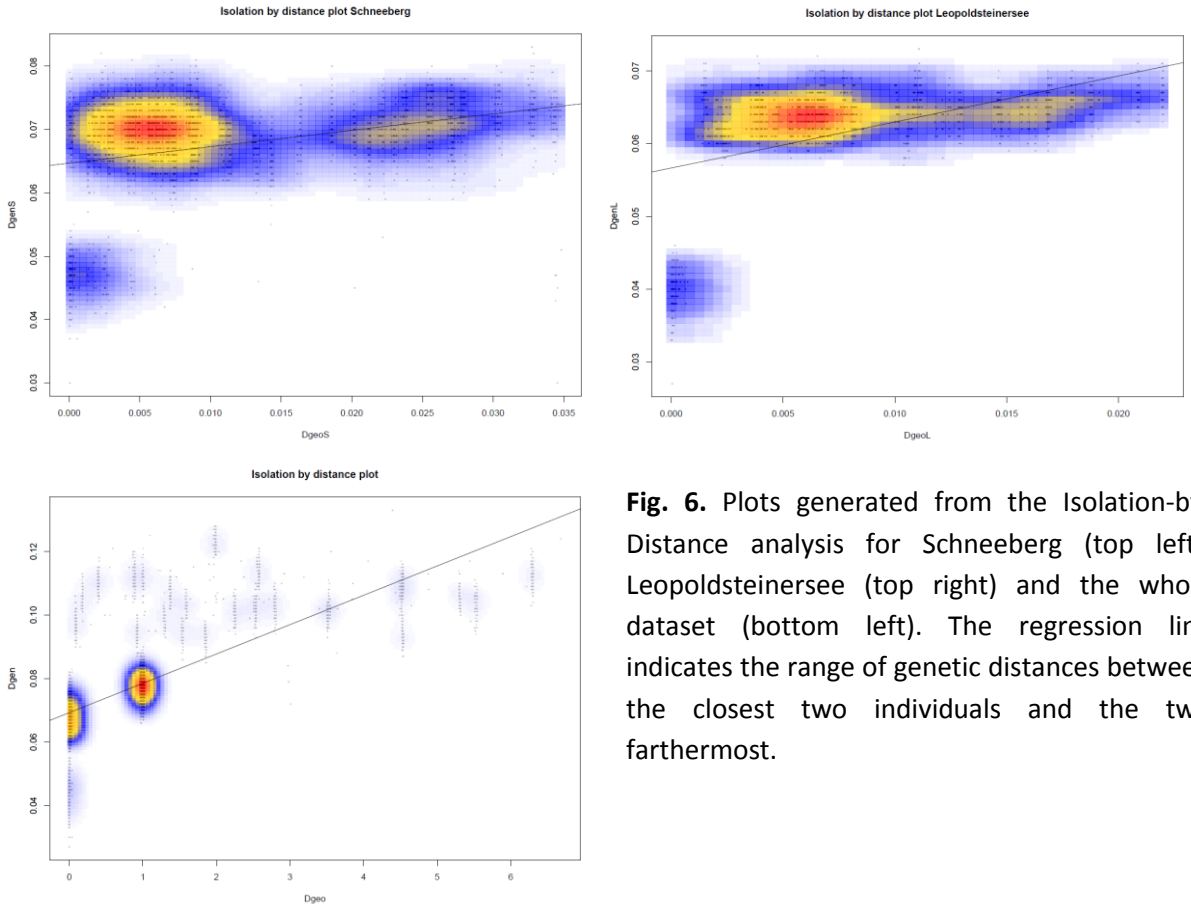

Fig. 6. Plots generated from the Isolation-by-Distance analysis for Schneeberg (top left), Leopoldsteinersee (top right) and the whole dataset (bottom left). The regression line indicates the range of genetic distances between the closest two individuals and the two farthermost.

3.5. Photosynthesis data

Evaluation of photosynthesis data reveals correlations between the proposed subspecies and the examined environmental factors. The majority of individuals from ssp. *balbisii* received only little amounts of Photosynthetic Available Radiation (PAR, below 300 mol photons m^{-2} s^{-1}) at time of measuring except for some outliers, compared to individuals from ssp. *auricula*, as well as the indets, which exhibit a much higher overall variation of received PAR (100 to 1900 µmol photons m^{-2} s^{-1}, Fig. 7a). The subspecies show a similar pattern when looking at Linear Electron Flow (Fig. 7b). This parameter is estimated by the Multispeq 2.0 field fluorimeter from PAR and ΦII, and represents the flow of electrons from H_20 to the Calvin Cycle. When looking at the results that describe whether incoming light ends up in Photosynthesis (ΦII) or is directed towards either regulated (ΦNPQ) or non-regulated (ΦNO) mitigation of excess light it can be seen that ssp. *balbisii* has a slightly higher ratio of ΦII than the other two groups ssp. *auricula* and indets which, again, exhibit a much wider range of values (Fig. 7c-e). When performing Wilcoxon tests ssp. *auricula* and ssp. *balbisii* are significantly different ($p < 0.05$) in all five examined parameters.

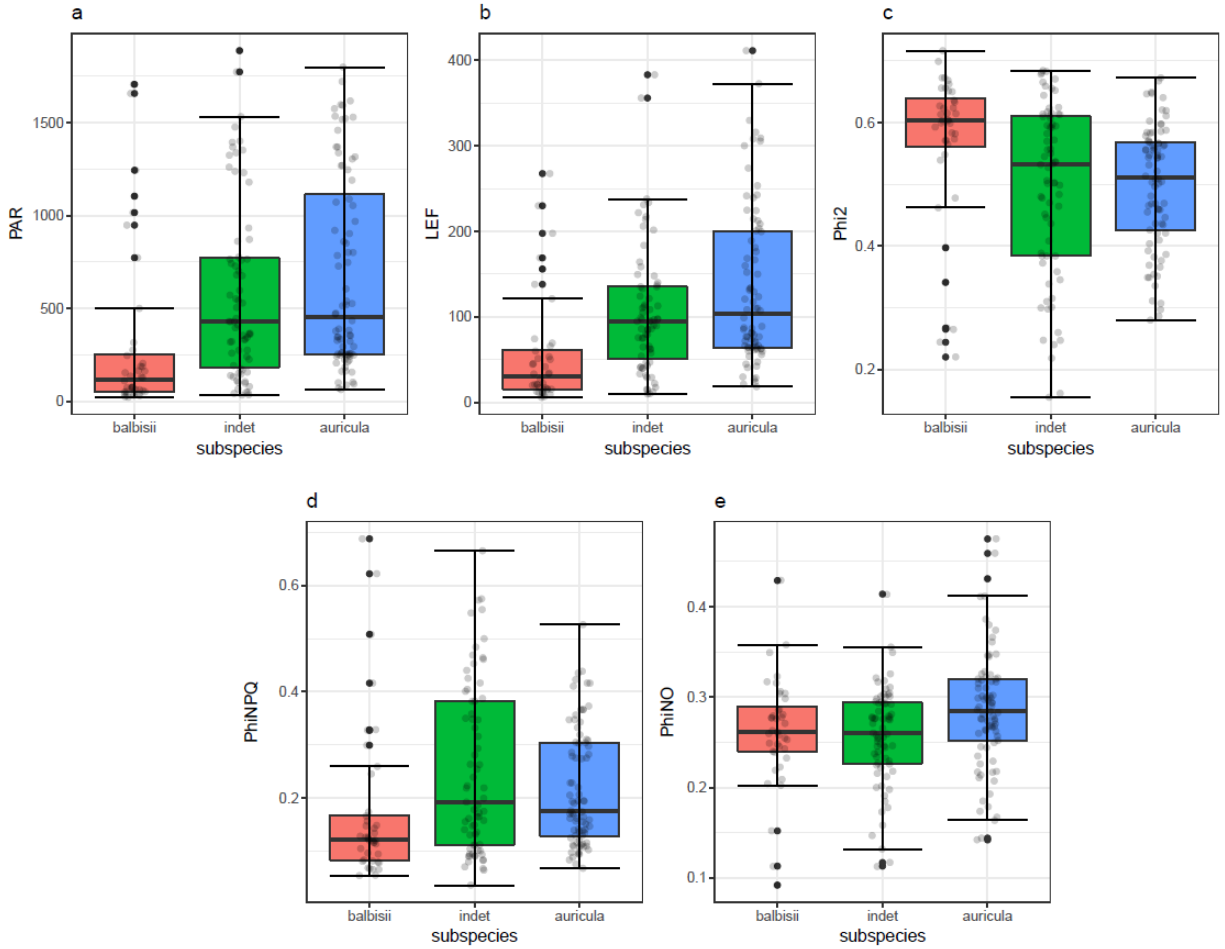

Fig. 7. Boxplots for PAR (a), LEF (b), ΦII (c), ΦNPQ (d) and ΦNO (e), separated into ssp. *balbisii* (red), indet (green) and ssp. *auricula* (blue). Black dots represent outliers, grey dots depict the density of individuals along the respective axis.

The effect increasing levels of PAR has on the parameters LEF, ΦII and ΦNPQ differs for both subspecies, as can be observed in figure 8. In general, higher amounts of incoming light lead to decreased percentages of it being used in Photosystem II (i.e. lower Quantum yield of Photosystem II, ΦII), whilst more light energy is being quenched non-photochemically (ΦNPQ). This trend is more pronounced in ssp. *balbisii* than in ssp. *auricula*, and can be seen on both mountains together and each mountain individually.

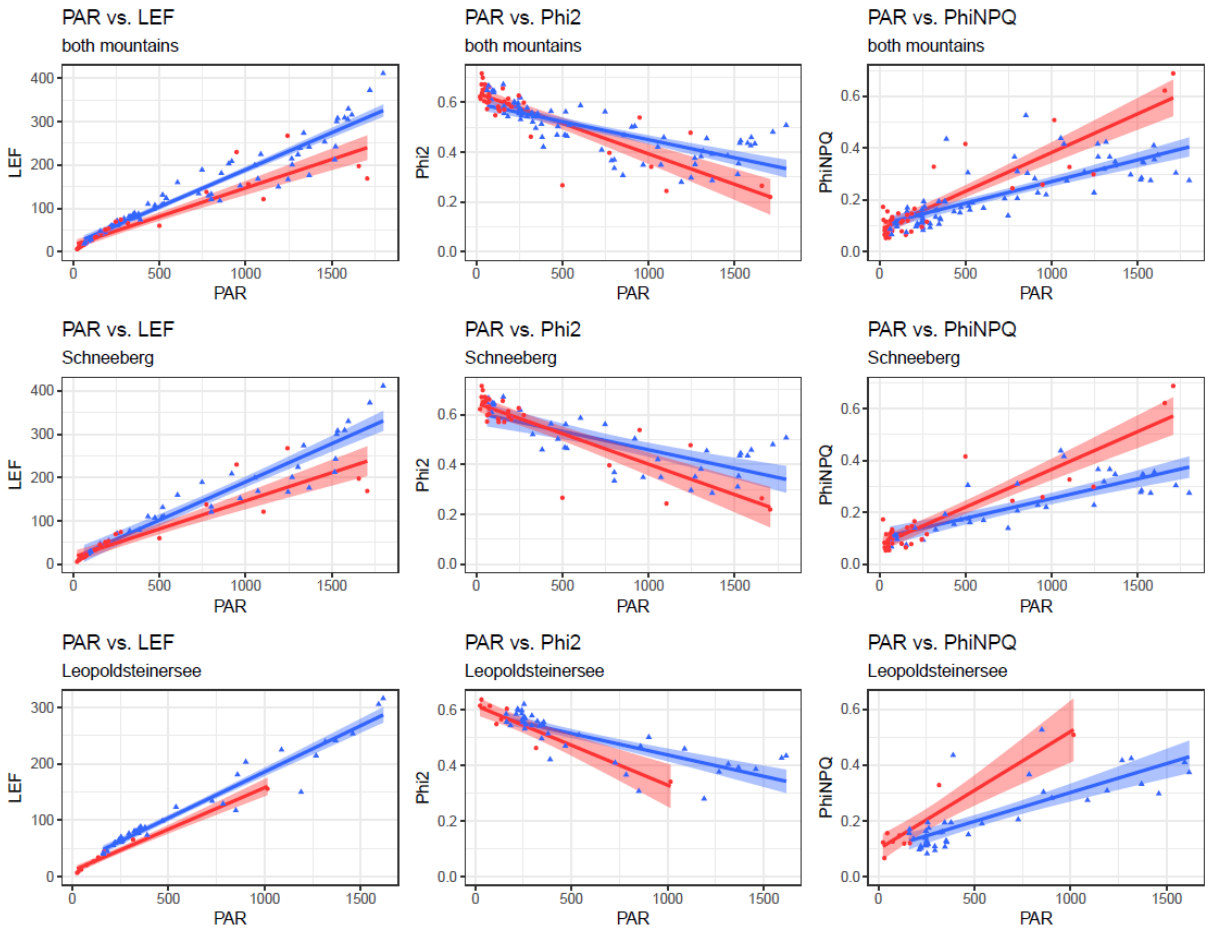

Fig. 8. Scatterplots with lines of best fit (95% confidence interval) that show how LEF (left column), ΦII (centre column) and ΦNPQ (right column) are affected by increasing levels of PAR for ssp. *balbisii* (red) and ssp. *auricula* (blue). Top row, both mountains combined; middle row, Schneeberg; bottom row, Leopoldsteinersee.

4. Discussion

4.1. Population structure, morphology and geography

There is no conspicuous genetic pattern that emerges within each focal sampling region, Schneeberg and Leopoldsteinersee, as could be seen in the ANGSD and polyploid genotyping analyses and their respective matrices, i.e. coancestry heatmaps. Moreover, no genetic distinction can be made between individuals of *P. auricula* ssp. *balbisii* and *P. auricula* ssp. *auricula*. Both relatedness matrices and Principal Component Analyses show the absence of this delimitation. The presumption that the traditional subspecies correspond to the genetically distinct novel species *P. auricula* and *P. lutea* because of morphological characters cannot be justified here. This further confirms the reported discrepancies concerning the morphological and geographical boundaries assigned by Zhang and Kadereit (2004, 2005) when they separated the *P. auricula* complex into two species (Boucher et al., 2016; Aymerich et al., 2014; Somlyay and Bauer, 2010). Both Zhang and Kadereit (2004) and Boucher et al. (2016) could assert this split into *P. auricula* and *P. lutea* genetically, but in both cases the sample size of the concerned species complex was very low, considering the extensive distribution area, with thirteen and seven accessions, respectively. As mentioned in the introduction, Somlyay and Bauer (2010) indicated an erroneous connection of morphological characters and distribution area made by Zhang and Kadereit (2004) by presenting contradictory cases where long-haired individuals from the presumed range of short-haired *P. lutea* were described and vice-versa. They pointed out that morphological characteristics in the complex are often dependent on the phenological state of the respective plant, and that this might have influenced the conclusion Zhang and Kadereit (2004, 2005) came to. Also, this variability in morphology has been known and described in the past (e.g Smith and Fletcher, 1949) and not being necessarily connected to geographical areas. With this in mind, the relevance of morphological characters to distinguish the novel species *P. lutea* and *P. auricula* remains problematic and, in this regard, makes its usefulness in the field questionable.

As can be inferred from the analyses that are based on genetic data, there is a general tendency in the set of samples that correlates geographic proximity with higher levels of relatedness. This trend was found to be highly significant (p-value = 0.001) in the IBD analyses and is represented visually in the regression line of the plot that contains all individuals of the dataset, which shows a large genetic difference between the closest two samples and those farthest away from each other (Fig. 6). The genetic analyses would associate samples X2, X5 and X10 with Schneeberg. This would make sense geographically, since those specimen are the closest to Schneeberg. The PCA of the second and third eigenvectors would cluster herbarium samples X2, X3, X5, X8 and X10, as well as Schneeberg and

Leopoldsteinersee together. This cluster only contains individuals from the Northern Eastern Alps apart from X9 which is located near X8 and X3. Although no final conclusion can be drawn from these findings, it is clear that geography plays an important part in the relatedness patterns of the *Primula auricula* complex. The strong patterns of Isolation-by-Distance recovered point to a considerable effect of drift in shaping the genetic patterns of this group. This is surprising, given that drift is expected to be less efficient in polyploids due to their increased effective population sizes (Balao et al., 2016). Both Zhang and Kadereit (2004) and Boucher et al. (2016) mentioned the importance of the Brenner zone as a phylogeographic barrier in connection with the *Primula auricula* complex. They suggest that this North-South line intersecting the Brenner pass at the Austrian-Italian border separates *P. lutea* to the west from *P. auricula* to the east. This would imply that herbarium sample X4 from Switzerland is in fact *P. lutea*. Since individuals X11 and X12 were gathered along this barrier, they might be part of either species or even admixed. However in the genetic analyses presented in this work no such pattern emerged that would indicate a delimitation of these three accessions from the rest. Having said this, the boundaries of novel species *P. lutea* and *P. auricula* were never precisely established by the authors and thus some uncertainty remains (Boucher et al., 2016).

4.2. Phenotypic plasticity

The altitudinal gradient along which *P. auricula* ssp. *balbisii, P. auricula* ssp. *auricula* and intermediate forms occur is part of the determination key by Fischer, Oswald and Adler (2008). Usually ssp. *balbisii* grows at lower elevations and ssp. *auricula* at higher altitudes. This pattern that resembles an almost gradual transition could be confirmed for the individuals in this work. Although not always significant, the intensity of incoming light as well as the amount of photosynthetic stress caused by it correlates with the subspecies. In general, ssp. *auricula*, which grows at higher altitudes, receives heightened levels of sunlight and consequently has to deal with more stress. Additionally, the amount of farina found on the leaves from individuals examined in various locations in the eastern Austrian Alps generally increases with altitude, probably hinting at a functional role in connection with exposed habitats (C. Priemer, pers. comm.). Thus, the phenotypic variation is most likely part of adapting to local habitats, although the exact mechanisms that lead to such phenotypic differences often remain obscure (Gratani, 2014). Recent re-measurements of leaf material from individuals of this work have revealed that the majority of samples have glandular hair lengths <0.2mm (C. Priemer, pers. comm.). Since glandular hair length is an important diagnostic character both in a traditional determination key and those established for novel species *P. lutea* and *P. auricula* these results would increase the ratio of ssp. *auricula* to ssp. *balbisii* significantly. This contradicts the assumptions of Zhang and Kadereit (2004, 2005), as short-haired populations occur where they would predict the distribution range of long-haired *P. auricula*.

4.3. Outlook

To gain further insight into the genetic structure of the complicated *Primula auricula* complex, a more comprehensive study with denser sampling and increased sample sizes that cover the whole distribution area is required. This should also include a focus on individuals along the putative hybridisation zones of the two species *P. auricula* and *P. lutea* that were given in broad outline by Zhang and Kadereit (2004, 2005) and, again, by Boucher et al. (2016). In addition to this, studying in greater detail the effect of past glaciations on this particular species complex could prove to be useful as well. To investigate the morphological spectrum more that is displayed by the species complex a common garden experiment to test for phenotypic plasticity will be necessary. Until then, the *Primula auricula* complex remains a topic for further debate.

References

Aymerich, P., López-Alvarado, J., & Sáez, L. (2014). *Primula subpyrenaica* (Primulaceae) a new species from the Pyrenean range (south-western Europe). *Phytotaxa*, *163*(2), 077-090.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. and Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one*, *3*(10), p.e3376.

Balao F, Tannhäuser M, Lorenzo MT, Hedren M, Paun O. 2016. Genetic differentiation and admixture between sibling allopolyploids in the *Dactylorhiza majalis* complex. *Heredity* 116: 351–361.

Beck-Mannagetta, G. (1893). *Flora von Nieder-Österreich* (Vol. 2). C. Gerold.

Boucher, F. C., Zimmermann, N. E., & Conti, E. (2016). Allopatric speciation with little niche divergence is common among alpine Primulaceae. *Journal of Biogeography*, *43*(3), 591-602.

Boucher, F. C., Casazza, G., Szövényi, P., & Conti, E., 2016. Sequence capture using RAD probes clarifies phylogenetic relationships and species boundaries in *Primula* sect*. Auricula. Molecular phylogenetics and evolution, 104*, 60-72.

Paul D Blischak, Laura S Kubatko, Andrea D Wolfe, SNP genotyping and parameter estimation in polyploids using low-coverage sequencing data, *Bioinformatics*, Volume 34, Issue 3, 01 February 2018, Pages 407–415, https://doi.org/10.1093/bioinformatics/btx587

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Mol Ecol*, 22: 3124-3140. https://www.doi.org/10.1111/mec.12354

Cocker, J. M., Wright, J., Li, J., Swarbreck, D., Dyer, S., Caccamo, M., & Gilmartin, P. M. (2018). *Primula vulgaris* (primrose) genome assembly, annotation and gene expression, with comparative genomics on the heterostyly supergene. *Scientific reports*, *8*(1), 1-13.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, *14*(8), 2611-2620.

Fischer, M.A., Oswald, K., Adler, W., 2008*. Exkursionsflora für Österreich, Liechtenstein, Südtirol.* 3rd ed. Linz, OÖ.

Francis, R. M. (2017). POPHELPER: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources, 17*(1), 27-32. DOI: 10.1111/1755-0998.12509

Gratani, L. (2014). Plant phenotypic plasticity in response to environmental factors. *Advances in botany, 2014.*

Heckenhauer, J., Samuel, R., Ashton, P. S., Salim, K. A., & Paun, O. (2018). Phylogenomics resolves evolutionary relationships and provides insights into floral evolution in the tribe Shoreeae (Dipterocarpaceae). *Molecular Phylogenetics and Evolution, 127,* 1-13.

Hughes, C. E., & Atchison, G. W. (2015). The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytologist, 207*(2), 275-282.

Huang K, Ritland K, Guo S, Shattuck M, Li B. (2014) A pairwise relatedness estimator for polyploids. *Mol Ecol Resour. 2014 Jul;14*(4):734-44. doi: 10.1111/1755-0998.12217. Epub 2014 Jan 27. PMID: 24460904.

Huang K, Guo ST, Shattuck MR, Chen ST, Qi XG, Zhang P, Li BG (2015) A maximum-likelihood estimation of pairwise relatedness for autopolyploids. *Heredity, 114*, 133-142. DOI: https://doi.org/doi:10.1038/hdy.2014.88

Huang K, Dunn DW, Ritland K, Li BG (2020) POLYGENE: population genetics analyses for autopolyploids based on allelic phenotypes*. Methods in Ecology and Evolution, 11*, 448-456. DOI: https://doi.org/10.1111/2041-210X.13338

Huang, H., & Knowles, L. L. (2016). Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic biology, 65*(3), 357-365.

Jombart, Thibaut. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics. 24*. 1403-1405. 10.1093/bioinformatics/btn129/.

Jombart, T., & Collins, C. (2015). An introduction to adegenet 2.0. 0. Imp Coll London-MRC Cent Outbreak Anal Model, 43.

Korneliussen, T.S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics 15***,** 356 (2014). https://doi.org/10.1186/s12859-014-0356-4

Li, H. & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics*, Epub. [PMID: 20080505]

Li, H.(2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics, 27*, 2987–2993.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics, 25*(16), pp.2078-2079.

Mast, A. R., Kelso, S., Richards, A. J., Lang, D. J., Feller, D. M., & Conti, E. (2001). Phylogenetic relationships in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. *International Journal of Plant Sciences*, *162*(6), 1381-1400.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data**.** *Genome Res*, 20:1297-303. DOI: 10.1101/gr.107524.110.

Meisner J., Albrechtsen A. (2018). Inferring population structure and admixture proportions in low depth NGS data. *Genetics*, *210*(2), 719-731.. https://doi.org/10.1534/genetics.118.301336

Ozenda, P. (1995). L'endémisme au niveau de l'ensemble du Système alpin. *Acta Botanica Gallica*, *142*(7), 753-762.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

Richards, J. (2003). *Primula*. Batsford, London.

Smith, W. W., & Fletcher, H. R. (1949). XXII.—The Genus *Primula*: Sections *Cuneifolia, Floribundae, Parryi*, and *Auricula*. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, *61*(3), 631-686.

Somlyay, L., & Bauer, N. (2010). Nomenclatural and taxonomic notes on two eastern taxa of the *Primula auricula* complex. *Biologia, 65*(5), 784-788.

de Vos, J. M., Hughes, C. E., Schneeweiss, G. M., Moore, B. R., & Conti, E. (2014). Heterostyly accelerates diversification via reduced extinction in primroses. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1784), 20140075.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Záveská, E., Maylandt, C., Paun, O., Bertel, C., Frajman, B., Schönswetter, P., & Steppe Consortium. (2019). Multiple auto-and allopolyploidisations marked the Pleistocene history of the widespread Eurasian steppe plant *Astragalus onobrychis* (Fabaceae). *Molecular phylogenetics and evolution*, *139*, 106572.

Zhang, L-B. (2002) Phylogeny, biogeography and systematics of *Soldanella* L. and *Primula* L. sect. *Auricula* Duby (Primulaceae) based on molecular and morphological evidence. [Doctoral dissertation, Johannes-Gutenberg Universität]

Zhang, L. B., Comes, H. P., & Kadereit, J. W. (2004). The temporal course of Quaternary diversification in the European high mountain endemic *Primula* sect. *Auricula* (Primulaceae). *International Journal of Plant Sciences*, *165*(1), 191-207.

Zhang, L-B. & Kadereit, J. W. (2004). Classification of *Primula* sect. *Auricula* (Primulaceae) based on two molecular data sets (ITS, AFLPs), morphology and geographical distribution. *Botanical Journal of the Linnean Society 146*: 1–26

Zhang, L-B. & Kadereit, J. W. 2005. Typification and synonymization in *Primula* sect. *Auricula* (Primulaceae). *Taxon 54* (3): 775–788.

Appendix

Table A. List of sampled individuals from Schneeberg (S)and Leopoldsteinersee (L).

Name	Individual	Population	Location	subsp.	Longitude	Latitude	altitude (m)	sequenced
Ps1_1	1	1	S	balbisii	15.8443473	47.7632256	1193	yes
Ps1_2	2	1	S	balbisii	15.8440401	47.7631627	1229	no
Ps1_3	3	1	S	balbisii	15.8440685	47.7627530	1259.6	yes
Ps1_4	4	1	S	balbisii	15.8442890	47.7630458	1202.6	no
Ps1_5	5	1	S	balbisii	15.8443361	47.7631329	1195.7	yes
Ps1_6	6	1	S	balbisii	15.8445404	47.7629391	1212	yes
Ps1_7	7	1	S	balbisii	15.8442249	47.7631081	1203.9	yes
Ps1_8	8	1	S	balbisii	15.8440492	47.7630919	1227.8	no
Ps1_9	9	1	S	balbisii	15.8441609	47.7631653	1207.6	yes
Ps1_10	10	1	S	balbisii	15.8440442	47.7631847	1228.1	yes
Ps2_1	1	2	S	indet	15.8423629	47.7623958	1323.3	yes
Ps2_2	2	2	S	indet	15.8423178	47.7624633	1323.9	yes
Ps2_3	3	2	S	indet	15.8423525	47.7623468	1325.7	yes
Ps2_4	4	2	S	indet	15.8423834	47.7624130	1321.5	yes
Ps2_5	5	2	S	balbisii	15.8421691	47.7621795	1337.8	yes
Ps2_6	6	2	S	balbisii	15.8424199	47.7627362	1319.4	yes
Ps2_7	7	2	S	balbisii	15.8423573	47.7620726	1337.4	yes
Ps2_8	8	2	S	balbisii	15.8422878	47.7622858	1330.1	no
Ps2_9	9	2	S	balbisii	15.8422329	47.7624767	1329.5	no
Ps2_10	10	2	S	balbisii	15.8423008	47.7622611	1330.3	no
Ps3_1	1	3	S	balbisii	15.8400737	47.7629918	1444.3	yes
Ps3_2	2	3	S	balbisii	15.8401642	47.7630490	1437.8	yes
Ps3_3	3	3	S	balbisii	15.8399259	47.7629204	1454.1	yes
Ps3_4	4	3	S	indet	15.8398766	47.7631303	1441	no
Ps3_5	5	3	S	balbisii	15.8397695	47.7630740	1449	yes
Ps3_6	6	3	S	balbisii	15.8399863	47.7630199	1446.3	yes
Ps3_7	7	3	S	indet	15.8399894	47.7629717	1448.9	no
Ps3_8	8	3	S	indet	15.8399774	47.7630628	1444	yes
Ps3_9	9	3	S	balbisii	15.8399921	47.7630506	1444.5	yes
Ps3_10	10	3	S	balbisii	15.8399878	47.7630124	1446.7	no
Ps4_1	1	4	S	indet	15.8384432	47.7673853	1457	yes
Ps4_2	2	4	S	balbisii	15.8385029	47.7672847	1461.7	yes
Ps4_3	3	4	S	indet	15.8385415	47.7671909	1467.3	yes
Ps4_4	4	4	S	indet	15.8385412	47.7672029	1466.4	yes
Ps2_5	5	4	S	indet	15.8386050	47.7671860	1466.8	yes
Ps4_6	6	4	S	indet	15.8386135	47.7671808	1467.1	yes
Ps4_7	7	4	S	indet	15.8385943	47.7671152	1472.3	no
Ps4_8	8	4	S	indet	15.8385986	47.7671115	1472.5	yes
Ps4_9	9	4	S	indet	15.8385477	47.7671830	1467.8	no
Ps4_10	10	4	S	auricula	15.8385481	47.7671673	1469	no
Ps5_1	1	5	S	indet	15.8375220	47.7663977	1549.6	yes
Ps5_2	2	5	S	balbisii	15.8375294	47.7663709	1550.3	no
Ps5_3	3	5	S	indet	15.8375288	47.7663857	1549.9	yes
Ps5_4	4	5	S	auricula	15.8375230	47.7663858	1550.2	yes
Ps5_5	5	5	S	indet	15.8375027	47.7663354	1552.4	yes
Ps5_6	6	5	S	indet	15.8374213	47.7663430	1555.1	yes
Ps5_7	7	5	S	indet	15.8374537	47.7663305	1554.4	yes
Ps5_8	8	5	S	balbisii	15.8374795	47.7662918	1554.6	yes
Ps5_9	9	5	S	indet	15.8374509	47.7663062	1555.2	no
Ps5_10	10	5	S	balbisii	15.8374698	47.7662630	1555.9	no
Ps6_1	1	6	S	balbisii	15.8358160	47.7657003	1653.5	yes
Ps6_2	2	6	S	auricula	15.8358323	47.7658225	1643.8	yes
Ps6_3	3	6	S	indet	15.8358463	47.7657439	1649.7	yes
Ps6_4	4	6	S	balbisii	15.8358735	47.7657460	1648.8	yes
Ps6_5	5	6	S	indet	15.8358755	47.7656019	1659.3	yes
Ps6_6	6	6	S	auricula	15.8359016	47.7658035	1644.1	yes

Ps6_7	7	6	S	indet	15.8358566	47.7657188	1651.4	yes
Ps6_8	8	6	S	indet	15.8358500	47.7657329	1650.5	no
Ps6_9	9	6	S	auricula	15.8358857	47.7657577	1647.7	no
Ps6_10	10	6	S	auricula	15.8358459	47.7655982	1661	no
Ps7_1	1	7	S	auricula	15.8334689	47.7643098	1754.4	yes
Ps7_2	2	7	S	auricula	15.8334196	47.7642419	1760.2	yes
Ps7_3	3	7	S	auricula	15.8335189	47.7642186	1755.3	yes
Ps7_4	4	7	S	balbisii	15.8334616	47.7642168	1759	yes
Ps7_5	5	7	S	indet	15.8334530	47.7642476	1758	yes
Ps7_6	6	7	S	indet	15.8334977	47.7641945	1757.4	yes
Ps7_7	7	7	S	indet	15.8332199	47.7640969	1770.3	yes
Ps7_8	8	7	S	indet	15.8332154	47.7639909	1767.2	no
Ps7_9	9	7	S	indet	15.8332910	47.7640359	1766.2	no
Ps7_10	10	7	S	indet	15.8331935	47.7640262	1769	no
Ps8_1	1	8	S	indet	15.8311151	47.7627515	1853.8	yes
Ps8_2	2	8	S	indet	15.8310418	47.7627777	1855	yes
Ps8_3	3	8	S	auricula	15.8310213	47.7628109	1855.4	yes
Ps8_4	4	8	S	auricula	15.8310666	47.7628144	1854.7	yes
Ps8_5	5	8	S	auricula	15.8309531	47.7627907	1856.3	yes
Ps8_6	6	8	S	auricula	15.8309541	47.7628219	1856.4	yes
Ps8_7	7	8	S	auricula	15.8310191	47.7628539	1855.5	yes
Ps8_8	8	8	S	auricula	15.8310389	47.7628529	1855.2	no
Ps8_9	9	8	S	auricula	15.8310266	47.7628401	1855.4	no
Ps8_10	10	8	S	auricula	15.8310226	47.7628283	1855.4	no
Ps8_11	11	8	S	balbisii	15.8310736	47.7627588	1854.4	no
Ps9_1	1	9	S	auricula	15.8173093	47.7668402	1846.7	yes
Ps9_2	2	9	S	auricula	15.8173077	47.7668597	1847.2	yes
Ps9_3	3	9	S	auricula	15.8173143	47.7668309	1846.5	yes
Ps9_4	4	9	S	auricula	15.8173279	47.7668462	1846.8	yes
Ps9_5	5	9	S	auricula	15.8173123	47.7668537	1847.1	yes
Ps9_6	6	9	S	auricula	15.8172824	47.7668742	1847.7	yes
Ps9_7	7	9	S	auricula	15.8172959	47.7668632	1847.4	yes
Ps9_8	8	9	S	auricula	15.8172979	47.7669307	1849.1	no
Ps9_9	9	9	S	auricula	15.8172767	47.7669377	1849.4	no
Ps9_10	10	9	S	auricula	15.8173010	47.7668967	1848.2	yes
Ps10_1	1	10	S	auricula	15.8114588	47.7727671	1999.5	yes
Ps10_2	2	10	S	auricula	15.8114860	47.7728174	1999.5	yes
Ps10_3	3	10	S	auricula	15.8112582	47.7728550	2006	yes
Ps10_4	4	10	S	auricula	15.8113040	47.7728474	2004.7	yes
Ps10_5	5	10	S	auricula	15.8112769	47.7728587	2005.5	yes
Ps10_6	6	10	S	auricula	15.8112164	47.7728673	2007.2	yes
Ps10_7	7	10	S	auricula	15.8111453	47.7728884	2009.3	yes
Ps10_8	8	10	S	auricula	15.8111616	47.7728637	2008.6	yes
Ps10_9	9	10	S	auricula	15.8112047	47.7728806	2007.6	no
Ps10_10	10	10	S	auricula	15.8112269	47.7728698	2006.9	no
11_1	1	11	L	balbisii	14.8486109	47.5774179	722.3	yes
11_2	2	11	L	balbisii	14.8492605	47.5775234	718.7	yes
11_3	3	11	L	balbisii	14.8487352	47.5772218	701.7	yes
11_4	4	11	L	indet	14.8493397	47.5775094	718.1	yes
11_5	5	11	L	indet	14.8488746	47.5775079	724.5	no
11_6	6	11	L	balbisii	14.8489333	47.5774367	716.3	yes
11_7	7	11	L	indet	14.8492094	47.5771797	692.6	yes
11_8	8	11	L	balbisii	14.8492089	47.5774593	711.9	no
11_9	9	11	L	indet	14.8490165	47.5774597	715.7	no
11_10	10	11	L	balbisii	14.8493593	47.5775658	724.1	no
12_1	1	12	L	balbisii	14.8475875	47.5780464	809.3	yes
12_2	2	12	L	balbisii	14.8473359	47.5777896	789.2	no
12_3	3	12	L	indet	14.8474469	47.5781709	826.8	yes
12_4	4	12	L	indet	14.8474125	47.5781727	826.4	yes
12_5	5	12	L	indet	14.8474040	47.5781005	817.5	yes
12_6	6	12	L	indet	14.8474247	47.5781532	824.2	yes
12_7	7	12	L	indet	14.8473372	47.5780760	814.3	yes

12_8	8	12	L	indet	14.8472849	47.5781344	820.1	no
12_9	9	12	L	indet	14.8472389	47.5781097	817.2	no
12_10	10	12	L	indet	14.8472135	47.5781465	820.9	no
13_1	1	13	L	indet	14.8450862	47.5783446	924.8	yes
13_2	2	13	L	indet	14.8448286	47.5781875	909.6	yes
13_3	3	13	L	indet	14.8453407	47.5781964	912	yes
13_4	4	13	L	indet	14.8451592	47.5783708	934.6	yes
13_5	5	13	L	indet	14.8452389	47.5783994	944.6	no
13_6	6	13	L	indet	14.8451298	47.5783025	918.7	yes
13_7	7	13	L	indet	14.8451626	47.5783616	931.8	yes
13_8	8	13	L	indet	14.8451242	47.5783982	942.5	no
13_9	9	13	L	indet	14.8451114	47.5784147	947.5	no
13_10	10	13	L	indet	14.8451320	47.5782758	915.7	no
14_1	1	14	L	indet	14.8428090	47.5784601	1029.5	yes
14_2	2	14	L	indet	14.8427953	47.5784354	1028.6	yes
14_3	3	14	L	indet	14.8427391	47.5784926	1035.4	yes
14_4	4	14	L	indet	14.8427233	47.5784965	1036.2	yes
14_5	5	14	L	auricula	14.8427754	47.5783989	1027	no
14_6	6	14	L	indet	14.8427225	47.5784882	1035.6	yes
14_7	7	14	L	indet	14.8426663	47.5786011	1046.2	yes
14_8	8	14	L	indet	14.8426958	47.5784751	1035.5	no
14_9	9	14	L	indet	14.8426152	47.5783935	1031.9	no
14_10	10	14	L	indet	14.8425864	47.5784052	1033.8	no
15_1	1	15	L	auricula	14.8413547	47.5789774	1095.7	yes
15_2	2	15	L	indet	14.8413778	47.5790556	1106.2	yes
15_3	3	15	L	indet	14.8413360	47.5789823	1096.3	yes
15_4	4	15	L	indet	14.8413656	47.5790023	1099.1	yes
15_5	5	15	L	indet	14.8413599	47.5790155	1100.9	yes
15_6	6	15	L	indet	14.8413549	47.5790111	1100.3	yes
15_7	7	15	L	indet	14.8413986	47.5789790	1096.2	no
15_8	8	15	L	indet	14.8413480	47.5791235	1115.7	no
15_9	9	15	L	auricula	14.8413685	47.5789925	1097.8	no
15_10	10	15	L	balbisii	14.8413715	47.5790133	1100.6	no
16_1	1	16	L	auricula	14.8476710	47.5834510	1493.4	yes
16_2	2	16	L	auricula	14.8476286	47.5834577	1491.7	yes
16_3	3	16	L	auricula	14.8476450	47.5834614	1492.7	yes
16_4	4	16	L	auricula	14.8478185	47.5834939	1501.6	yes
16_5	5	16	L	auricula	14.8476620	47.5834728	1494.1	yes
16_6	6	16	L	auricula	14.8477104	47.5834522	1495.4	yes
16_7	7	16	L	auricula	14.8476409	47.5834729	1493.1	no
16_8	8	16	L	auricula	14.8475454	47.5834233	1485.5	no
16_9	9	16	L	auricula	14.8476894	47.5834084	1492.3	no
16_10	10	16	L	auricula	14.8477189	47.5833664	1491.9	no
17_1	1	17	L	auricula	14.8514346	47.5834893	1570.1	yes
17_2	2	17	L	auricula	14.8515433	47.5835188	1573.6	yes
17_3	3	17	L	auricula	14.8516728	47.5833857	1541.3	yes
17_4	4	17	L	auricula	14.8516756	47.5833505	1533.5	no
17_5	5	17	L	auricula	14.8516245	47.5833295	1530	yes
17_6	6	17	L	auricula	14.8515693	47.5832814	1523.4	yes
17_7	7	17	L	auricula	14.8515263	47.5834409	1559.6	yes
17_8	8	17	L	auricula	14.8515544	47.5834241	1555.2	no
17_9	9	17	L	auricula	14.8517570	47.5834858	1576.8	no
17_10	10	17	L	auricula	14.8518206	47.5834959	1586.4	no
18_1	1	18	L	auricula	14.8581115	47.5854758	1658.1	yes
18_2	2	18	L	auricula	14.8580565	47.5854918	1660.7	yes
18_3	3	18	L	auricula	14.8580946	47.5854930	1660	no
18_4	4	18	L	auricula	14.8580595	47.5854511	1657.3	yes
18_5	5	18	L	auricula	14.8580318	47.5854795	1659.6	yes
18_6	6	18	L	auricula	14.8580355	47.5854341	1656	yes
18_7	7	18	L	auricula	14.8580463	47.5854727	1659.1	yes
18_8	8	18	L	auricula	14.8580450	47.5854419	1656.6	no
18_9	9	18	L	auricula	14.8579533	47.5854242	1655.3	no

18_10	10	18	L	auricula	14.8581929	47.5855376	1662.3	no
19_1	1	19	L	auricula	14.8618455	47.5868802	1750.4	yes
19_2	2	19	L	auricula	14.8618896	47.5868669	1748.7	yes
19_3	3	19	L	auricula	14.8618653	47.5868923	1751.2	yes
19_4	4	19	L	auricula	14.8618419	47.5868732	1749.8	yes
19_5	5	19	L	auricula	14.8618301	47.5869044	1752.8	no
19_6	6	19	L	auricula	14.8618902	47.5868361	1746.1	yes
19_7	7	19	L	auricula	14.8618917	47.5868594	1748.1	yes
19_8	8	19	L	auricula	14.8618642	47.5869048	1752.3	no
19_9	9	19	L	auricula	14.8619109	47.5869455	1754.7	no
19_10	10	19	L	auricula	14.8618449	47.5868920	1751.4	no

Table B. List of herbarium samples.

X1	**WU 066770** Lombardei 2000m **https://wu.jacq.org/WU0066770**
X2	CP414 Raxalpe **https://www.jacq.org/detail.php?ID=1541683**
X3	Salzburg Ginzberger 1915 or 1916 **https://www.jacq.org/detail.php?ID=1541682**
X4	**WU 030607** Graubünden (CH) **https://wu.jacq.org/WU0030607**
X5	**WU 068383** Scheibbs **https://wu.jacq.org/WU0068383**
X6 *(not used for sequencing)*	E. Walitzi Kärnten (Hochobir) **https://www.jacq.org/detail.php?ID=1541681**
X7	**WU 0082545** Kärnten (Nockberge) **https://wu.jacq.org/WU0082545**
X8	**WU 0107690** Liezen **https://wu.jacq.org/WU0107690**
X9	**WU 0096917** Hallein **https://wu.jacq.org/WU0096917**
X10	070804 Baden (400m-500m, Schwarzföhre)
X11	Tirol (Schönswetter) (12.10.19) **https://www.jacq.org/detail.php?ID=1541676**
X12	**WU0123413** Tirol (Innsbruck) (16.4.19) (Nr. 1880) **https://wu.jacq.org/WU0123413**
X13	Kärnten (Gartnerkofel) (4.8.19) (CG-20191128-01) (Sample 7/10)