CHARLES UNIVERSITY IN PRAGUE
FACULTY OF SCIENCE

Study programme: Biology
Curriculumr: Botanika

Bc. Jana Bayerová

Evolutionary and ecological consequences of polyploidization in *Arabidopsis arenosa* complex in the Western Carpathians

Diploma thesis

Supervisor: RNDr. Filip Kolář, Ph.D.

Praha, 2016

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 15.8.2016

Podpis

Acknowledgment

I would like to thank my advisor RNDr. Filip Kolář, Ph.D. to lead my work and patience and valuable advices in finishing my thesis.

I would also thank to Mgr. Libor Mořkovský for support during the work and also to help with work in R program.

I woul also thank all members of our team, who were cooperated in field collecting and measuring.

Abstract

Polyploidization is a key mechanism of rapid speciation, with many phenotypic consequences which extent, however, is poorly understood. A deeper understanding of the evolutionary implications of genome duplication is limited due to lack of knowledge of the links between changes in genome, the phenotype of the individual and environmental constraints. Natural lineages closely related to model species represent the ideal systems for addressing such questions. The thesis is thus focuses on highly promising yet overlooked di-polyploid system within Arabidopsis genus. In the western Carpathians morphologically distinct populations of diploid and tetraploid plants of *Arabidopsis arenosa* grow along a marked altitudinal gradient. Using high-throughput DNA sequencing, measuring morphological characteristics and collecting ecological data of high alpine and foothill populations I try to reveal main trends of genetic and morphological variability of these populations. Additionaly, using morphometrics of natural and experimentally planted populations we want to test the hypothesis whether morphological divergence of alpine and foothill populations has a genetic basis or is driven by phenotypic plasticity. The presented thesis is an important multidisciplinary combination of genetic research on natural related of model organism and field measurements. Close relationship *Arabidopsis arenosa* to model organism *Arabidopsis thaliana* allows to interpret the results in a broader framework.

Key words:

alpine, ecology, polyploidization, *Arabidopsis*, *A. arenosa*, Tatry Mountains, ecological consequences, RADSeq

Abstrakt

Polyploidizace je klíčový mechanismus rychlé speciace rostlin s řadou fenotypických důsledků, jejichž rozsah je stále nedostatečně prozkoumán. Hlubšímu porozumění evolučním důsledkům genomové duplikace brání hlavně nedostatečná znalost souvislostí mezi změnami na úrovni genomu, fenotypu jedince i ekologických vazeb. Tyto souvislosti je možné nejsnáze studovat v přírodních populacích blízkých příbuzných modelových druhů, pro něž je k dispozici nejvíce informací o struktuře a regulaci genomu a jeho interakcích s fenotypem. Tato práce se proto zaměřuje na slibný dosud přehlížený di-polyploidní systém v rámci rodu Arabidopsis. V oblasti západních Karpat nalézáme morfologicky odlišné populace diploidních i tetraploidních rostlin druhu *Arabidopsis arenosa* rostoucích podle výrazného výškového gradientu. Pomocí měření morfologie, sběru ekologických dat a genotypování SNP markerů (získaných high-throughput DNA sekvenováním) populací z vysokohorského prostředí a z podhůří jsem se pokusila poodhalit hlavní směry genetické a morfologické variability těchto populací. Zároveň pomocí morfometrického studia přírodních i pěstovaných populací jsem testovala hypotézu, zda výrazná morfologická odlišnost horských populací a populací z podhůří představuje geneticky či ploidně podmíněnou odlišnost nebo projev fenotypické plasticity jedné linie se širokou nikou. Předkládaná práce nabízí dosud ojedinělou kombinaci genetického výzkumu na divokých příbuzných modelového organismu a terénních datech. Blízká příbuznost *Arabidopsis arenosa* s modelovým organismem *Arabidopsis thaliana* umožňuje zasadit výsledky práce do širšího kontextu vědeckých poznatků.

Klíčová slova:

alpský, ekologie, polyploidizace, *Arabidopsis*, *A. arenosa*, Tatry , ekologicke dusledky, RADSeq

Contents

Abstract 4
Abstrakt 5
Contents 6
1. Introduction .... - 7 -
2. Literature review .... - 8 -
2.1. Polyploidization .... - 8 -
2.2. Study group .... - 9 -
2.3. Questions and hypotheses .... - 10 -
3. Methods .... - 12 -
3.1. Field sampling/ Collecting .... - 12 -
3.2. Flow cytometry (FCM) .... - 13 -
3.3. Morphometric measurement (and data analysis) .... - 15 -
3.3.1. ImageJ measurements .... - 15 -
3.3.2. Shape measurements .... - 17 -
3.3.3. Joining of data .... - 18 -
3.4. Environmental chracteristics .... - 18 -
3.5. Growing experiment .... - 19 -
3.6. Molecular data analysis .... - 19 -
3.7. Comparative analyses of genetic, morphological and ecological variation .... - 20 -
4. Results .... - 22 -
4.1. Morphological variation at original sites .... - 22 -
4.2. Ecological differentiation .... - 29 -
4.3. Molecular data .... - 32 -
4.4. Morphological variation of plants from experiment in common conditions .... - 36 -
4.5. Data synthesis (variation partitioning) .... - 39 -
5. Discussion .... - 42 -
A. Morphological differentiation and its causes. .... - 42 -
B. Colonization of alpine environment. .... - 43 -
C. What is the role of genome duplication in colonization of alpine environments? ..... - 44 -
D. Origin of alpine tetraploids. .... - 45 -
6. Conclusions .... - 46 -
Citations .... - 47 -

1. Introduction

My thesis addresses phenotypic consequences of whole genome duplication (polyploidy) in natural populations exhibiting altitudinal differentiation. As a suitable example I chose a polyploid complex of *Arabidopsis arenosa* agg. (Měsíček 1970, Měsíček a Goliášová 2002) in a small area of Western Carpathians (Tatras in Slovakia) where it encompasses diploid and autotetraploid populations both spanning over a remarkable altitudinal gradient of 1500 m. Populations of *A. arenosa* are very morphologically different, especially high altitude populations against lower populations. In the alpine belt itself, we can observe slight morphological differentiation even among populations of different ploidy level that were mirrored in their distinct (yet invalid) taxonomic treatments (Měsíček and Goliášová 2002, Schmickl et al 2012).

Therefore, I asked what the causes of this foothill-alpine morphotype differentiation are, as the principal question of my thesis. I hypothesized that this striking morphological differentiation could be caused by several (non-exclusive) explanations: a) ecological conditions, b) genetic background or c) ploidy level differentiation.

This whole complex in the Tatras is also well suitable for study of general consequences of polyploidization. Our question was, what the role of genome duplication in colonization of alpine environments is. Due to the availability of genetic resources for its close relatives with assembled reference genomes (*A. thaliana*, *A. lyrata*). I could easily leverage new modern methods based on high-throughput DNA sequencing (RADSeq – Restriction site associated DNA Sequencing) for addressing this question. Further, I analyzed ploidy level of the populations from High Tatras, Western Tatras and Low Tatras and in the surrounding foothill areas by means of flow cytometry. Finally, I screened the phenotypes of the natural populations (by measuring morphological characters) and described their niches by screen of ecological conditions at the original sites occupied by *A. arenosa* (e.g., soil sampling, orientation of slope, gradient).

2. Literature review

2.1. Polyploidization

Polyploidization, i.e. multiplication of the whole chromosome sets, is a key mechanism of speciation in plants (Soltis et al. 2009). It is an instant change and new species is often strongly reproductively isolated from its parental species (e.g., Coyne and Orr 2004). The expansion of genome is followed by various changes. For example polyploidisation may affect size and morphology of individual, in growth rates and viability of individual or in tolerance to ecological conditions or competitiveness (Levin 2002). Commonly debated result of polyploidization is that polyploids have benefit in extreme conditions and their occurrence is higher in alpine ecosystems (Husband et al. 2013). However, only a few empirical studies addressed such questions. To comprehensively assess this question, more polyploid complexes with altitudinal differentiation should be studied and the studies should combine phenotypic (ecological, morphological) and genetic approaches. The new modern ecological-evolutionary view can provide study of polyploid complexes of close relatives to model organism. For this reason, the focal species of my thesis, *Arabidopsis arenosa* is well suitable to address such question (see section 2.3.).

How can polyploids originate? Firstly, genome doubling may result from interaction of genetically close plant(s) within one species. The plants originated this way are called autopolyploids. They have two sets of the same homologous chromosomes (Parisod et al. 2010). The second way is formation a new individual due to fusion of chromosome sets from two different plant species (typically following their hybridization). This new individual is made from two different sets of chromosome and is called allopolyploid (Ramsy a Schemske 1998). The allopolyploids could have the problem with mating, because their new chromosomes might be not compatible (Comai 2005). This differences is need to consider especially in molecular studies, because in allopolyploids it is difficult to distinguish between consequences of hybridization or polyploidization.

With respect to polyploidization, *Arabidopsis* genus represents a well-studied model systems (e.g. Clauss and Koch 2006, Hunter and Bomblies 2010). However, molecular-evolution processes are usually well studied in *Arabidopsis*, but polyploidization in natural populations and its ecological-evolutionary consequences are virtually unknown.

In genus *Arabidopsis* we can find several examples of allopolyploids between diploid lineages, because of incomplete isolation between these species, for example allotetraploids *A. suecica* and *A. kamchatica* (Bomblies and Weigel 2007). Some changes in ecological niches were observed, but not statistically valuated.

Among the most significant natural polyploid belongs complicated di- tetraploid complex *A. arenosa* agg. (Kolář et al. 2015, Měsíček 1970, Schmickl et al. 2012). The benefit of these complex is its close related of cytotypes and containing of autotetraploids with tetrasomic inheritance (Arnold et al. 2015).

Alpine populations of the *A. arenosa* complex in the Western Carpathians (Tatry Mts.) are genetically distinct from alpine populations of remaining mountains areas occupied by *A. arenosa* (Arnold et al. 2015, Kolář et al. 2016, and unpublished data). In traditional concepts the alpine Tatrean populations were differentiated as a separate species *A. neglecta*.

While in most of mountains is just tetraploid cytotype, in the Tatry Mts. we can find both diploid and tetraploid cytotype. Diploid and tetraploid cytotypes also occur in the surrounding lower altitude. In this area we can observe mosaic of different cytotypes through different altitude. In our previous research we find some triploids and mixed populations in this area (Kolář et al. 2015).

The Western Carpathian area is well suitable for studying polyploidization and its consequences with relation to colonization of alpine areas. There is a big species diversity likely reflecting survival of *A. arenosa* populations in local glacial refugium (Kolář et al 2016), and this area is considered as a speciation area for this group (Schmickl et al. 2012).

2.2. Study group

The *A. arenosa* complex includes annuals, biennials or shortly multiannual herbs. They are 5 – 45 cm height and they have ground rosette. This plants grow in natural habitats rocky characters. Especially in shady places, wetting rocks or scree in substrate range from limestone to siliceous substrate. They can grow in the secondary places, like disrupted habitats, roadsides, railways or on alluvial gravels.

In this area take palace repetitive process of polyploidization and hybridization (Clauss a Koch 2006), which increases complexity of the species and demands further study.

Between alpine and foothill populations we can find at the first sight obviously differences (Fig. 1). The individuals from alpine populations are smaller, with bigger odorous flowers while the foothill populations include bigger individuals with smaller flowers and rich inflorescence.

Fig. 1: Illustration of alpine (the first part of picture) and foothill (the second part of picture) individual of *Arabidopsis arenosa*.

The causes of this striking differentiation are still unknown. We have some hypothesis that it could be caused by genetical differentiation, by ecological conditions or by phenotypic plasticity.

From the previous study (Kolář et al. 2016) we know that Western Carpathians populations are genetically related. The alpine diploid populations (growing typically at 1,500-2,500 m) are closely related with near foothill diploid populations *A. arenosa* agg. (growing typically at 600-900 m) with typically different morphology, but they don't create separated species.

2.3. Questions and hypotheses

A. What are the causes of the pronounced morphological differentiation between alpine and foothill populations?

(Hypothesis 1) Alpine morphotype reflects phenotypic plasticity.

(Morphology of cultivated plants will correspond with growing conditions and between alpine and foothill populations cultivated under the same conditions will not be found

significant genetically differentiation. Ecological factors will be the major drivers explaining the morphological variation of the field-collected plants)

(Hypothesis 2) Alpine morphotype is genetically determined. (Morphology of cultivated plants will correspond with originate of plants regardless growing conditions. Genetic structure will be the major driver explaining the morphological variation of the field-collected plants)

B. Was the alpine environment in the study area colonized during a single or multiple independent events?

(Hypothesis 3) Single origin of the alpine morphotype: All alpine populations are closer to ech other than to the foothill populations.

(Hypothesis 4) Multiple origins of the alpine morphotype: Certain (groups of) alpine populations are more closely related with foothill populations then among each other.

C. What is the role of genome duplication in colonization of alpine environments?

(Hypothesis 5) Alpine tetraploids are able to reach more extreme environments in higher altitudes than alpine diploids.

(Hypothesis 6) Alpine tetraploids occupy wider ranges of ecological conditions (climatic, soil) than alpine diploids.

D. How did Tatras alpine tetraploids originated?

(Hypothesis 7) Alpine tetraploids from Tatry Mts. originated from local alpine diploids.

(Hypothesis 8) Alpine tetraploids from Tatry Mts. represent independent derivate type from the foothill tetraploids.

(Hypothesis 9) The origin of alpine tetraploids is more complicated process, which includes hybridization with more lineages of *A. arenosa* (e.g. between alpine diploids and other foothill morphologically differentiated individuals).

3. Methods

3.1. Field sampling/ Collecting

Plant material was collected from 2013 to 2015 in Slovakia and Poland. Under a collaborative project lead by K. Marhold, in which I took part. All fieldwork was done under permission issued by the local authorities in Slovakia and Poland. In total we sampled 54 populations. Our strategy was designed to cover Tatras' alpine populations and surrounding foothill populations (supplement Tab. 1). The localities were designed by expectations suitable conditions for *Arabidopsis* plants and chosen from previous data collection from Měsíček and Goliášová and unpublished materials kindly provided by M. Kolnik. We included 26 populations of diploids (2x), 25 populations of tetraploids (4x) and 3 mixed populations (2x, 4x) (Fig. 2). In several populations we founded potential aneuploid individuals (pops No. AA021, AA084), in three population (pop. No AA021, AA162, AA308) we founded triploid individual, these aberrant individuals were, however, excluded from the genetic analyses and for further analyses we assigned the population's major ploidy level.

Fig. 2: Map of collecting populations (red 2x, blue 4x, orange 2x/4x).

We sampled approximately 15 – 20 individuals at each locality with respect to population size. Each individual was labeled and stored in plastic bag in cold conditions until used in the FCM analysis. For morphometric analysis we collected the main stem and its rosette, then separated petals, sepals, stamens and pistil from one randomly selected flower per main inflorescence and stuck those together with one largest leaf of rosette and the second stem leaf from the base to black paper by tape (Fig. 4 B). For DNA extraction we picked leaves and stems to tea bags, put them to silica gel and let them dry. The stems with remaining leaves, flowers and fruits inflorescences were air-dried. Herbarium vouchers (Fig. 4 A) are deposited in the Herbarium of Charles University in Prague (PRC).

We recorded information about the localities - GPS co-ordinates and altitude, together with information about the habitat and environmental conditions. For more locality details see supplementary (Tab. 1).

At each site we also selected a 3×3 m plot with abundant *A. arenosa* and recorded there cover of rocks, scree and of all vegetation layers as well as presence of all vascular plant species (phytosociological relevés) at 3x3 m^2 area.

We collected and air dried the mixed rhizosphere soil samples (minimum from 4 places at the sites of vegetation samples) at 47 localities (supplement Tab. 1). The soil samples and measured following characteristics: pH and concentrations of selected elements (C, N, K, Ca, and Mg). In the Analytical Laboratory of the Institute of Botany, Průhonice CZ, following characteristics were analysed: available and exchangeable (extract using 0.1M KCl) pH measured by pH meter (WTW Multilab 540) and available concentrations of Ca determined with an atomic absorption spectrometer (Unicam 9200X; Unicam Ltd, Cambridge, UK) after 1 m ammonium acetate extraction (pH 7·0).

3.2. Flow cytometry (FCM)

DNA ploidy levels (Suda et al., 2006) were estimated by using the following simplified FCM two-step protocol (Doležel et al., 2007). We used a small part of fresh petals or an intact tissue from the leaves. The each plant was chopped together with the appropriate volume of the internal reference standard (*Solanum pseudocapsicum*, 2C = 2.59 pg, Temsch, Greilhuber & Krisai, 2010; the same individual was used for all measurements), using a sharp razor blade in a Petri dish containing 0.5 mL of ice-cold Otto I buffer (0.1 M citric acid, 0.5 % Tween-20). The suspension was filtered through a 42 mm nylon mesh and incubated for 10 min at

room temperature. Isolated nuclei were stained with 1 mL of Otto II buffer (0.4 M Na_2HPO_4 . 12 H_2O) supplemented with 4′, 6-diamidino-2-phenylindole (DAPI) at a final concentration of 4 µg mL^{-1} and β-mercaptoethanol (2 µL mL^{-1}). After a few minutes, the relative fluorescence intensity of 3000 particles was recorded using a Partec ML flow cytometer (Partec GmbH, Münster, Germany) equipped with a UV-led lamp. Histograms were evaluated using FloMax software, ver. 2.4d (Partec, Münster, Germany) (Fig. 3). Fresh petals were preferred due to the absence of endopolyploidy. We analyzed up to five individuals in a pooled petal samples together to reduce the analysis costs and time request. The pooled leaf samples of tetraploids were analyzed by the same approach, but stem parts of tetraploids or vegetative parts of diploids were analyzed separately for each individual. For the populations where the fresh material was not available, we used dry material from silica gel. Each plant was separately re-analyzed separately whenever mixed-ploidy samples were detected.

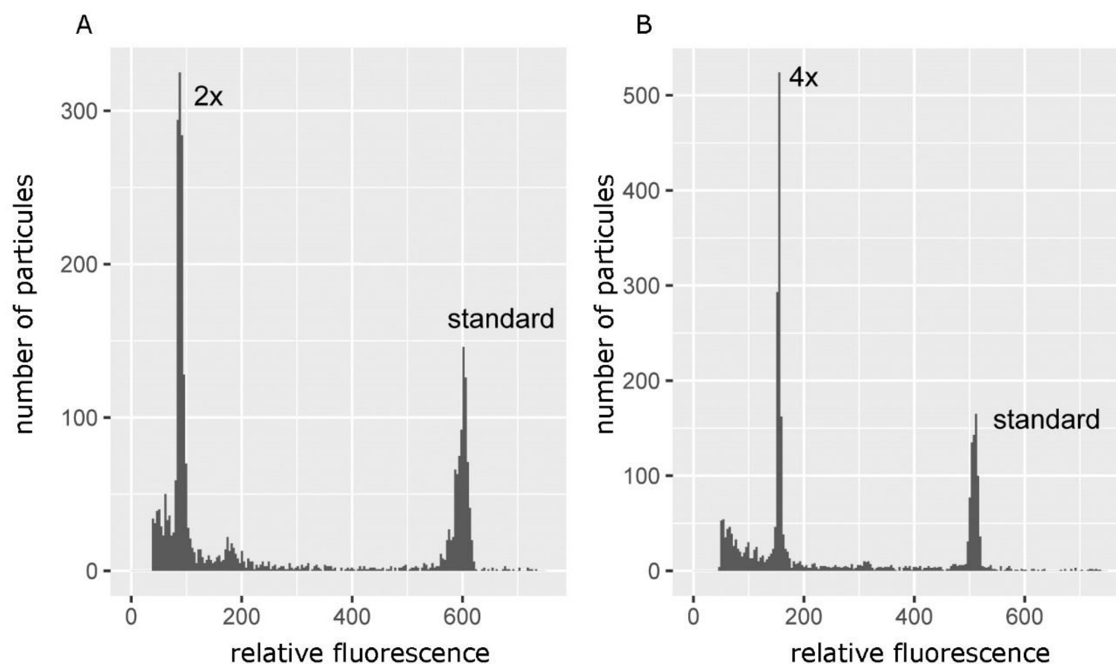

Fig. 3: Example histograms of flow cytometric analysis of fresh petal tissue of one individual of diploid (A) and tetraploid (B) individual of *A. arenosa* stained with DAPI and using *Solanum pseudocapsicum* used as internal reference standard

3.3. Morphometric measurement (and data analysis)

The pasted flowers and leaves and herbarium vouchers from filed collected individuals were scanned at 600 dpi to .tif file format. For better work with individuals, I manually edited scans in Photoshop CS2 and I cropped each individual into one .jpg format file.

3.3.1. ImageJ measurements

For measure length, width and angle I used program ImageJ (Radband, W.S. 1997-2016). I opened single individual .jpg files in program ImageJ, select measure tool (line or angle), set scale in menu *Analyze -> Set Scale…* for 600dpi pixels, *Known distance* for 25.4 (1 inch), *Pixel aspect ratio* for 1.0, *Union of length* for mm, tick this *Global*, check *Scale* – 23.6220 pixels/unit and click *"OK"*. Then I marked distances, push "Ctrl+m" which caused discovering values in special window. I save file after measuring set of characters for individual. For flower it was petal length/ width and sepal length/ width. For leaves it was different length, width and number of divisions and for herbarium voucher it was length of stem, length of inflorescence, angle between stem and silique spindle, width of silique and both for three measurements from randomly chosen silique (Fig. 4).

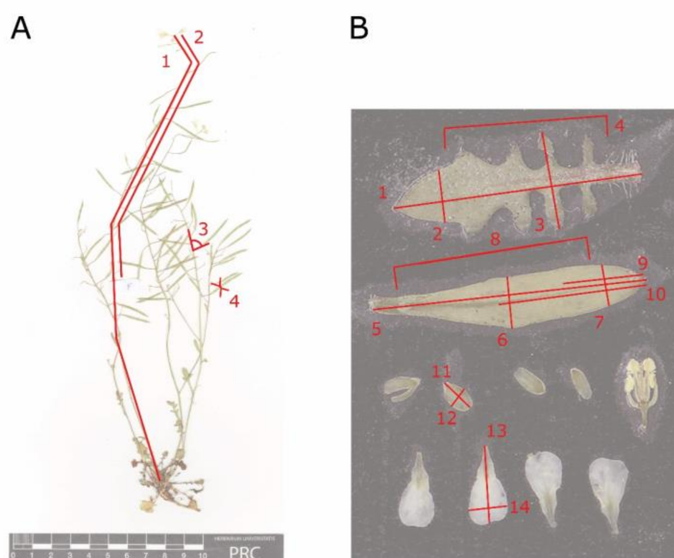

Fig. 4: Morphological characters measured. **A – Herbarium voucher:** 1 – stem (length of whole stem), 2 – inflorescence (length of inflorescence), 3 – mean_silique_angle (mean of three randomly measured angle between stem and silique spindle), 4 – mean_silique_width (mean of three randomly measured silique width); **B – pasted rosette leaf**: 1 – rosleaf_length (total length of leaf), 2 – rosleaf_w_termlobe (width of termolobe), 3 – rosleaf_w_divpart (biggest width in divided part of leaf), 4 – rosleaf_divisions (number of lobe), **pasted stem leaf:** 5 – stemleaf_length (total length of leaf), 6 – stemleaf_w_divpart (biggest width of divided part of leaf), 7 – stemleaf_w_termlobe (width of termolobe), 8 – stemleaf_divisions (number of lobe), 9 – stemleaf_length_termlobe (length of termolobe), 10 – stemleaf_1length (length from first division to top), **pasted flower:** 11 – sepal_length (length of sepal), 12 – sepal_width (width of sepal in the widest part), 13 – petal_length (length of petal), 14 – petal_width (width of petal in the widest part).

After measuring all individuals I processed and check dataset. Then I joined all morphologic characters to ono table for analysis.

WORKFLOW FOR FLOWERS

I had a set of the measurements for the flowers from ImageJ, which generate .tsv files (=tabulator separated). In the each file, there were name of individual, order of measurement, column “Angle” (not interested me) and column “Length”, where the length of petal, width of petal, length of sepal and width of sepal were. For completing all data to one data-set, I must filtered non complete measurements and join measurements for all individuals in one set. For this I use program Bash (Git for Windows).

After that I obtained four lists of file names (“files-all” with individuals with all four measurements, “files-petals” with individuals with measurements for petals, “files-sepals” with individuals with measurements for sepals and “files-none” with no measurements). Then I checked, if all files were processed and if data were unique. And then I export measurements to .tsv files in long format.

For next processing I used program R Studio. I extract the codes of individuals, made columns for measurements and join complete and incomplete measurements. This wide format file were prepared to connecting with other morphologic measurements.

WORKFLOW FOR LEAVES

I used quite similar process for leaves measurements as for flower measurements. However, as an input file I used table with all measurements and list of all individual names. It was processed in R Studio program, where I selected columns for leaves characters. Then I connected table with leaves data-set and checked if all data are unique. I recalculate measurements to same units – “mm” like in new conjunct data-set will be. Then I saved it and this new file were prepared to join to all other measurements.

WORKFLOW FOR HERBARIUM VOUCHERS

The herbarium measurements were processed analogically as flower measurements. I had a folder with single individual’s .jpg files obtained from cropped population’s scans from program Photoshop CS2. I import them and measure in program ImageJ and obtained set of

.tsv files for single individuals. Then I modify and filter data in program Bash (Git for Windows).

I obtained seven lists of file names (files "files-silique-angle-all", "files-silique-angle-none", "files-silique-width-all", "files-silique-width-none", "files-stem", "files-stem-inflorescence" and "files-stem-none"). Than I check, if all files were processed and if data were unique. And then I export measurements to .tsv files in long format.

For next processing I used program R Studio. I extract the ID codes of individuals and made columns for measurements. Then I connect non complete data for stem to complete data and saved it in wide format. After that I calculate silique width mean and silique angle mean of three randomly measured siliques per individual ideally and saved this files in wide format. This assembled datasets were prepared to connecting with other morphologic measurements.

3.3.2. Shape measurements

To evaluate the shape of the biggest petal from individual flower I used the SHAPE 1.3 program (Iwata, H. and Y.Ukai 2002). First I converted the .jpg files of the individuals to .bmp in program IrfanView 4.37 by menu *File -> Batch Conversion/ Rename…* Then I used image analysis program ChainCoder from SHAPE package, which records contours of selected shapes to chain-code file (Freeman 1974). It was configured as follows: *Object Color* – Bright (White), *Scale included* - No, *Scan Direction* – Y. Then click to *"Proceed to Processing"* and chose the folder with files (in .bmp format) for one population, add selected files and click *OK*. For each image click *"Load Images*", throught *"Select Area"* selecting the biggest petal, make *"Grey Scale"*, make *"Histogram"* and set slider to get a nice looking contour (by guessing according to previous image and usually the minimum on histogram), click *"Binarize Image"*, check a sharp boundary, optionally make *Ero Dil Filter* for 1 and then make *Labeling Object* for 500, make *"Chain Coding"* and click *"Save to File"*.

I stored each of the population files to one folder. For next work I concatenate .chc population files in program Bash (Git for Windows) to one file.

Second I generate data in .nef format. I used program CHC2NEF from the SHAPE package, import .chc file (*Max Harmonic No.*: 20, *Normalization Method* – based on the First Harmonic), make folder for .nef format and analyze file. I obtained .nef format file.

In the next step I used program PrinComp from SHAPE package, open our .nef file and processed it. I used default settings and calculate PCA (Principal Component Analysis). It saves “all.pcs” file with the result of PCA analysis and other files with information about analysis.

The results of PCA analysis were used in the PrinPrint program from the SHAPE package to construct an explanatory image for each the PCA axes. To evaluate the differences in the shape with respect to external criteria like altitude or ploidy I created a plot in R Project (package “ggplot2”) which was later combined with the explanatory images from program PrinPrint using Adobe Photoshop CS2.

3.3.3. Joining of data

For completed final data-set I used program R Studio. As an input I used .tsv files. It contains measurements for flowers, measurements for leaves, measurements for herbarium vouchers, which contain three files – data about silique angle, data about silique width and data about stem and inflorescence. The other file was with information about region, altitude and main ploidy. There were other file with data about colour of flowers and about stage of plant.

With the use of commands from package ‘dplyr’ I joined data for all morphological measurements. Then I separated mixed populations to two parts for diploids and tetraploids and filtered only rows with our Tatry populations.

I obtained prepared morphological data-set for following analysis.

3.4. Environmental chracteristics

The filed collected data was stored in database Turboveg. For each population, I extracted the following environmental characteristics: cover of herb and moss layers and cover of bare rock, (ii) soil characteristics determined from soil samples (pH, Ca concentration), (iii) Ellenberg indicator values (generated through program Juice) based on the vascular plant species presence-absence data. Other ecological data contain soil samples for analyze. Ellenberg values provide estimates of environmental characteristics of the sites inferred from species composition data (Ellenberg, 1992). In addition, (iv) for each population we obtained twelve precipitation-related and twelve temperature-related climate data as well as slope

inclination data from high-resolution climate database operated by GeoModel Solar, Bratislava, Slovakia.

3.5. Growing experiment

Data to this dataset were measured on life plants cultivated in phytotrons in cooperation with Mgr. Terezie Mandáková, PhD in the Central European Institute of Technology (CEITEC), in Brno, CZ. We used a subset of eight populations (pop codes AA016, AA084, AA087, AA090, AA168, AA171, AA208, AA229), from each ideally ten plants were half of them grown in alpine conditions and half of plants in foothill conditions and entire individuals in reproductive stage were sampled and processed in a similar way as the field-collected ones (see chapter 3).

It was counted and measured manually by ruler and wrote to paper, then rewrite to computer and used to calculation analysis. For work with data I used program R Studio. I completed all measurements for individuals to one table, checked NAs and filled population mean and modus. This was only for individuals from populations with less than 33% NAs in all. Then I did basic analysis – boxplots, histograms and check correlations. After that I figured up PCA analysis and made plots.

3.6. Molecular data analysis

DNA was obtained from dried samples from silica gel and extracted according to a CTAB protocol (Doyle, JJ, Doyle, JL 1987). Next work was process according to protocol in Arnold et al. 2015 and sequenced on an Illumina HiSeq 2000 platform in a service laboratory EMBL Genomics Core Facility [GeneCore], Heidelberg, Germany. I worked with filtered dataset.

Prior my analyses, the raw data were processed by other members of our team, generally following Kolář et al. 2016. Briefly, the raw reads were demultiplexed, quality trimmed (>30 Phred quality score) and mapped using Stampy version 1.0.23 (Lunter & Goodson 2011) on a repeat-masked genome of *Arabidopsis lyrata* v. 1.0.25 (Hu *et al.* 2011). The Genome Analysis Tool Kit v3.3.0 (GATK) (McKenna *et al.* 2010) was used for realignment around indels (*IndelRealigner* tool) and for simultaneous SNP discovery and genotyping (*HaplotypeCaller* and *GenotypeGVCFs*) following the recommended best practice, performing SNP discovery and probabilistic genotype calling across all samples

simultaneously (www.broadinstitute.org/gatk/). Using GATK (*VariantFiltration* and *SelectVariants*) we retained only bi-allelic sites that mapped to nuclear chromosome scaffolds with a minimum mapping quality of 40, which did not show mapping quality bias for the reads supporting the non-reference allele (keeping variants with mapping quality rank sum test value above -12.5) and which were present in at least 50% of our individuals at a sequencing depth of 8× or greater. In addition, we excluded potentially paralogous sites by excluding regions in which eight diploid whole-genome sequenced *A. arenosa* individuals (Yant *et al.* 2013) were heterozygous in more than two positions within a < 2 kb region (following Arnold *et al.* 2015).

I further worked with a set of 85,979 filtered single-nucleotide polymorphism (SNP) data stored in a .vcf file. I subsetted the data for Tatry, diploids and tetraploids using *bcftools* in Bash (Git for Windows), ensuring that non-variant sites within each subset were removed. This resulted in three datasets hereafter called entire Tatrean (73,059 SNPs, 4.64% of missing data), diploid Tatrean (50,467 SNPs, 5.29% of missing data) and tetraploid Tatrean (60,831 SNPs, 3.98% of missing data). Then I made PCoA analysis for whole dataset, for Tatry and for Tatry diploids and Tatry tetraploids. Then I count DAPC analysis for the same subsets. Next analysis was StAMPP for the same subsets.

Finally, I searched for admixture among the five major groups of diploid *A. arenosa* using Treemix v1.12 (Pickrell & Pritchard 2012). Considering the groups (2xA, 2xF, 2x AL, 4x) as populations, I constructed a maximum likelihood population graph from allelic frequencies of 73,059 loci based on blocks of 100 SNPs and allowed for one migration edge in order to see the principal admixture event among the five groups. The trees were bootstrapped by 100 replicates.

3.7. Comparative analyses of genetic, morphological and ecological variation

I quantified the relative contribution of environment, ploidy, and genetic structure to the morphological variation of the *A. arenosa* populations through variation partitioning (function *varpart* in R package "vegan" 2.4, Oksanen et al. 2016) which is based on a direct ordination (redundancy analysis, RDA, function *rda* in vegan). Scaled matrix of the morphological characters was constrained by the following environmental predictors (selected to capture the whole scale of environmental predictors without imposing strong redundancy): (i) cover of

herb layer and bare rock in the 3x3 m plots, (ii) pH and Ca content (from soil samples), (iii) Ellenberg indicator values for light, and nutrient availability inferred from accompanying species data (recorded in the 3x3 m plots), (iv) average slope of the surrounding area (inferred from geographical model) and (v) scores on first axis of separate unconstrained ordinations (principal component analyses) of the twelve temperature-linked and twelve precipitation-linked bioclimatic variables, respectively (derived from high resolution climatic model). In addition, as a genetic constraint I used scores on the first four PCoA axes derived from matrix of inter-population genetic distances (Nei's distances inferred in StAMPP, see above).

Finally I also analysed the potential drivers of genetic variation (again represented by scores on the first four PCoA axes) that were constrained by (i) the same set of environmental variables (ii) altitude of the original sites of the populations and (iii) ploidy level.

4. Results

4.1. Morphological variation at original sites

First, I summarized the morphological variation of individuals from 45 field-collected populations in boxplot for whole dataset, standardized it by subtracting the mean and dividing by standard deviation and separated by either altitudinal group (Fig. 5) or ploidy level (Fig. 6). The altitudinal groups reflect three major ecologically and genetically characterized groups (see next chapters) of populations: (i) ALPINE (A) group of diploid and tetraploid populations from alpine environments (altitudes 1,625-2,488m), (ii) FOOTHILL (F) group of diploid and tetraploid populations from foothills environments (semi-shaded rocks, screes, river beds, altitude 437-844m) and (iii) LIMESTONE (LOWER)ALPINE (LA) group of diploid and tetraploid populations from limestone areas of Belanske Tatry Tatras and Western Tatras. Namely diploid populations Osobita mountain - pop AA086 and col Kopske sedlo - pop AA091 with separate genetic position and specific ecology; tetraploid populations Giewont mountain pop AA164, Kopa Kondraczka mountain pop AA165, Sivy Vrch mountain pop AA178 with specific ecology (preferring limestones in montane and subalpine zones in altitudes 1,552-1,953m , i.e. higher than the foothill populations).

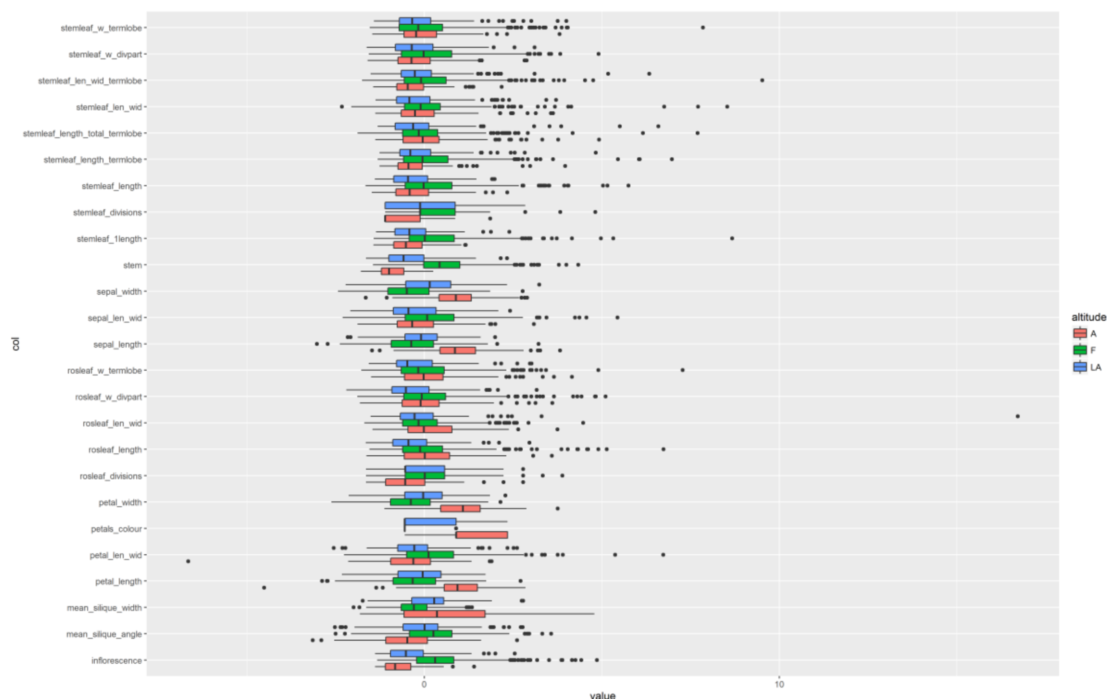

Fig. 5: Variation (after standardization) in 25 measured morphological characteristics of *A. arenosa* populations from Tatry Mts. and surrounding foothill areas in a total dataset of 45 populations separated by altitudinal groups (A=alpine, F=foothill, LA=limestone (lower) alpine).

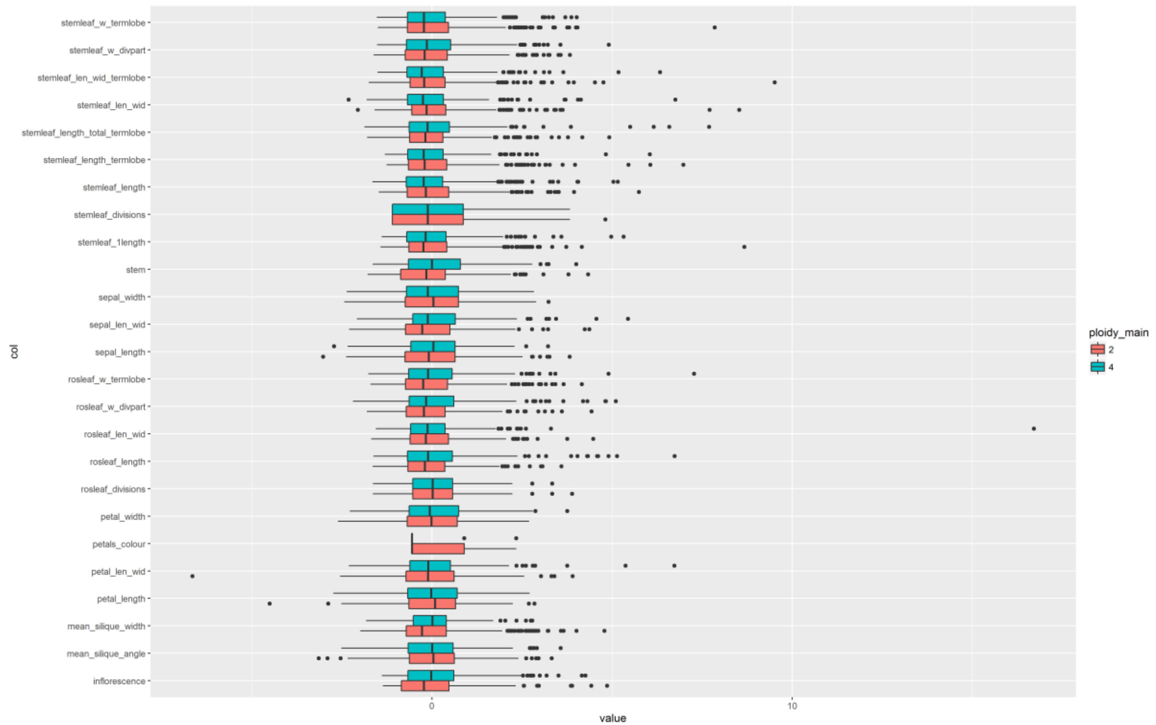

Fig. 6: Variation (after standardization) in 25 measured morphological characteristics of *A. arenosa* populations from Tatry Mts. and surrounding foothill areas in a total dataset of 45 populations separated by ploidy level.

Then I made histograms for each character and colored them either by altitudinal groups (Fig. 7) or ploidy level (Fig. 8).

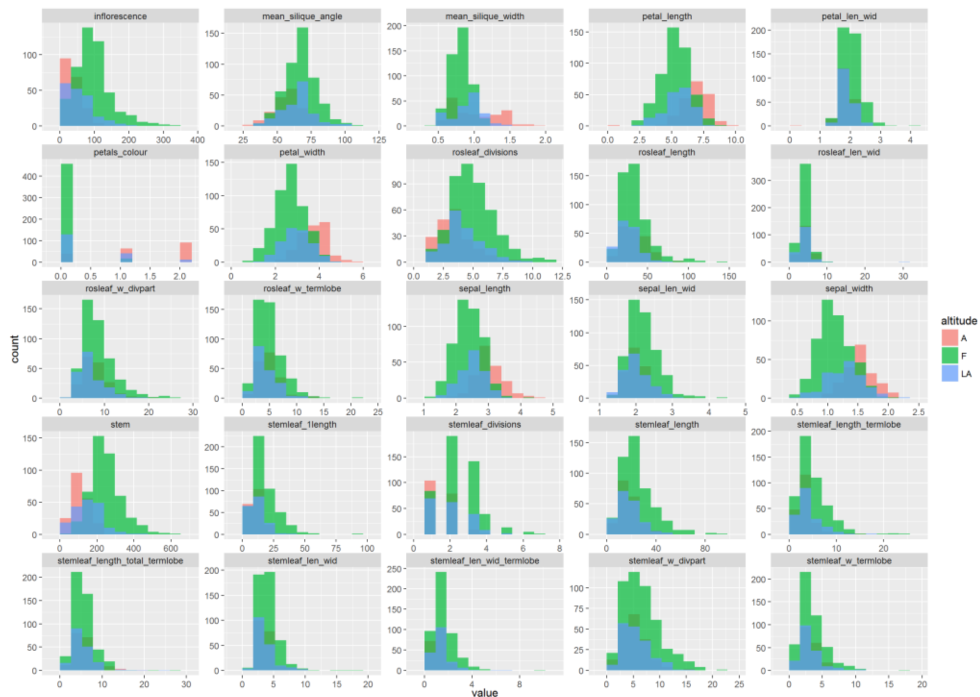

Fig. 7: Distribution of variation in 25 measured morphological characteristics of *A. arenosa* populations from Tatry Mts. and surrounding foothill areas in a total dataset of 45 populations colored by altitudinal groups (A=alpine, F=foothill, LA=limestone (lower) alpine).

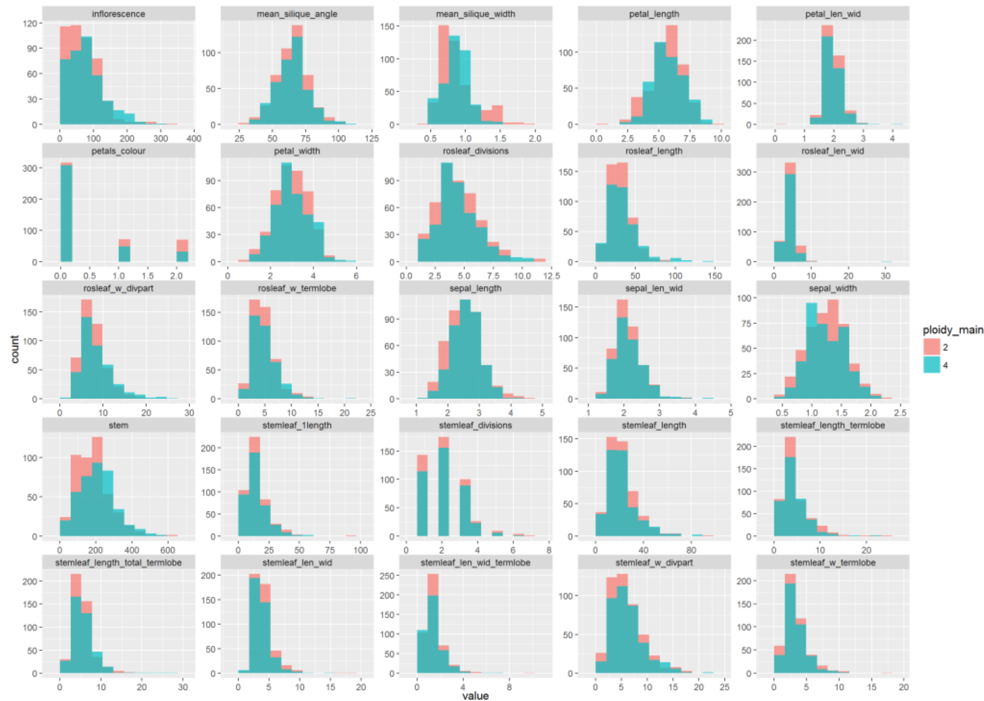

Fig. 8: Distribution of variation in 25 measured morphological characteristics of *A. arenosa* populations from Tatry Mts. and surrounding foothill areas in a total dataset of 45 populations colored by ploidy.

Then I standardized the dataset by subtracting the mean and dividing by standard deviation and calculated PCoA analysis colored by altitudinal group (Fig. 9 A) or by ploidy level (Fig. 9 B). In general, plants from the highest stands (alpine group) were well separated from populations occupying lowest elevations (foothill groups); the populations occupying limestone (sub)alpine stands exhibited intermediate morphology. Alpine populations exhibit in general larger flowers, wider silique, but lower stature and less divided stems (Fig. 10). In contrast, diploid and tetraploid populations were morphologically very similar as documented by nearly complete overlap in the PCA (Fig. 9 B).

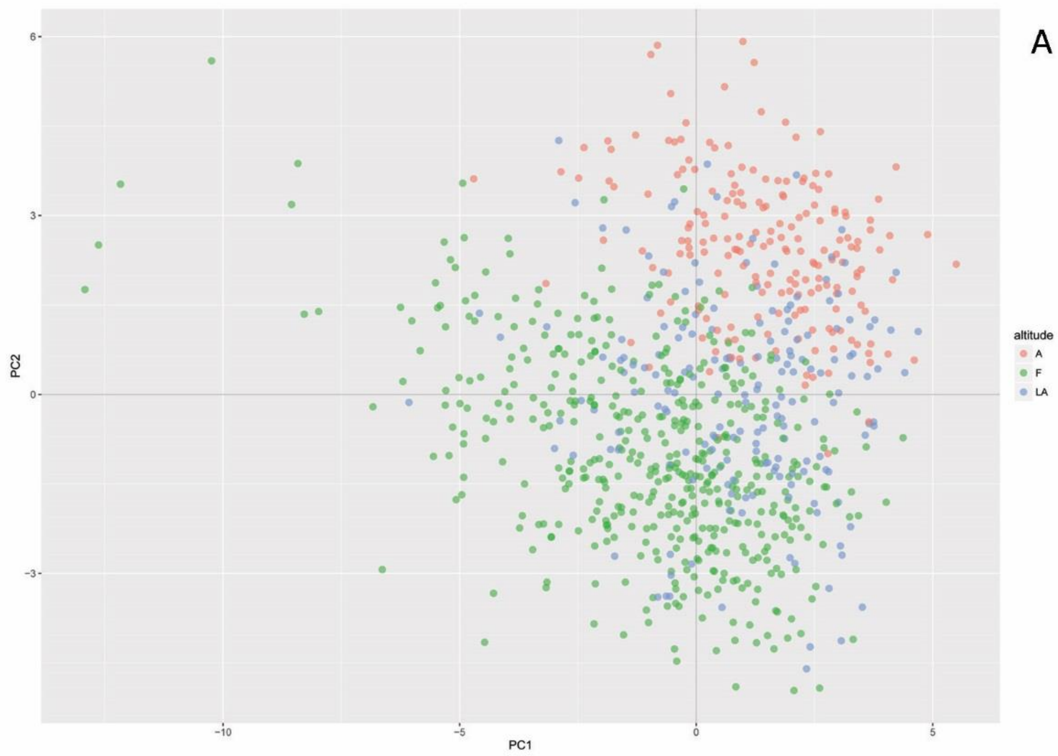

Fig. 9: Plot of PCoA analysis for 25 measured morphological characteristics of 45 *A. arenosa* populations from Tatry Mts. and surrounding foothill areas (A = colored by altitudinal groups, B = colored by ploidy level of populations)

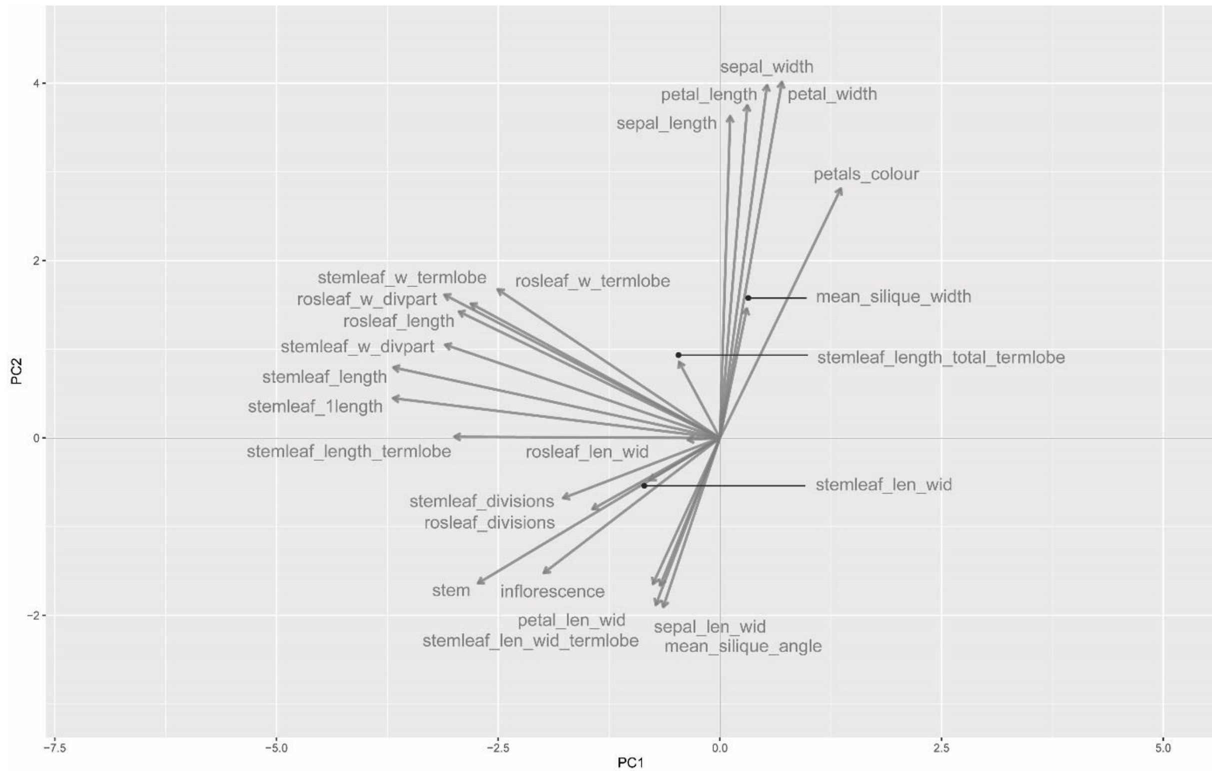

Fig. 10: Plot of PCoA analysis for 25 measured morphological characteristics of 45 *A. arenosa* populations from Tatry Mts. and surrounding foothill areas and plotted contributions of the measured characteristics to the first two PC axes.

VARIATION IN PETAL SHAPE

In PCA analysis of petal shape the diploid and tetraploid populations are not separated (Fig. 11). Neither do the three altitudinal groups (Fig. 12), which were, however, distinct in other morphological characters (see the previous chapter).

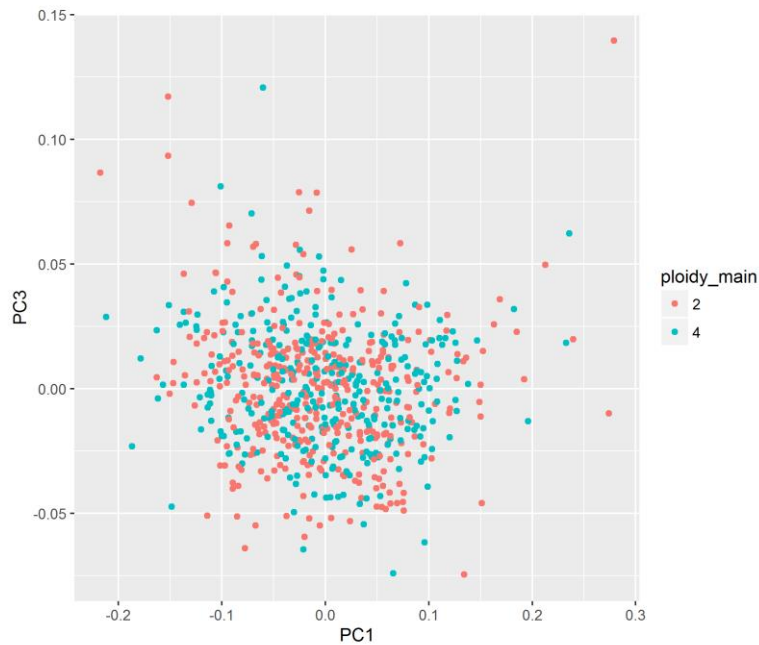

Fig. 11: PCA analysis of petal shape measured on 45 populations *A. arenosa* from Tatry Mts. and surrounding foothill areas, colored by ploidy level of populations.

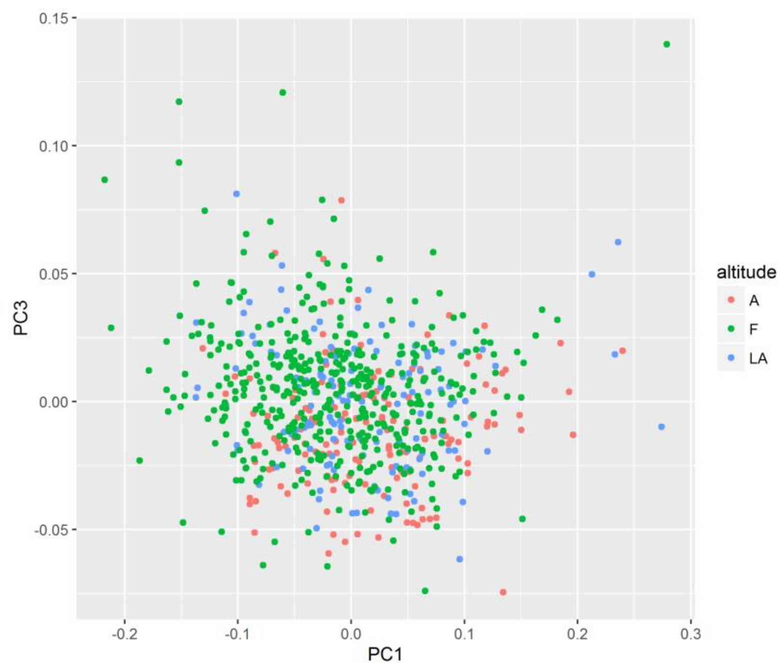

Fig. 12: PCA analysis of petal shape measured on 45 populations *A. arenosa* from Tatry Mts. and surrounding foothill areas, colored by altitudinal groups (A=alpine, F=foothill, LA=limestone (lower) alpine).

Contours reconstructed from the PC axes (PrinPrint program from the SHAPE package) showed that PC1 axis explain size of petals (width of head part) and PC3 axis explain notches at the base of petals (Fig. 13). PC2 axis explained shape of the hooked base of petals (not shown). However, I consider this variation rather an artifact caused by sticking on paper

during collecting then actual shape variation of petals. We complemented petal pictures for PC1 and PC3 axis by boxplots separately for ploidy level and for altitudinal groups (Fig.13).

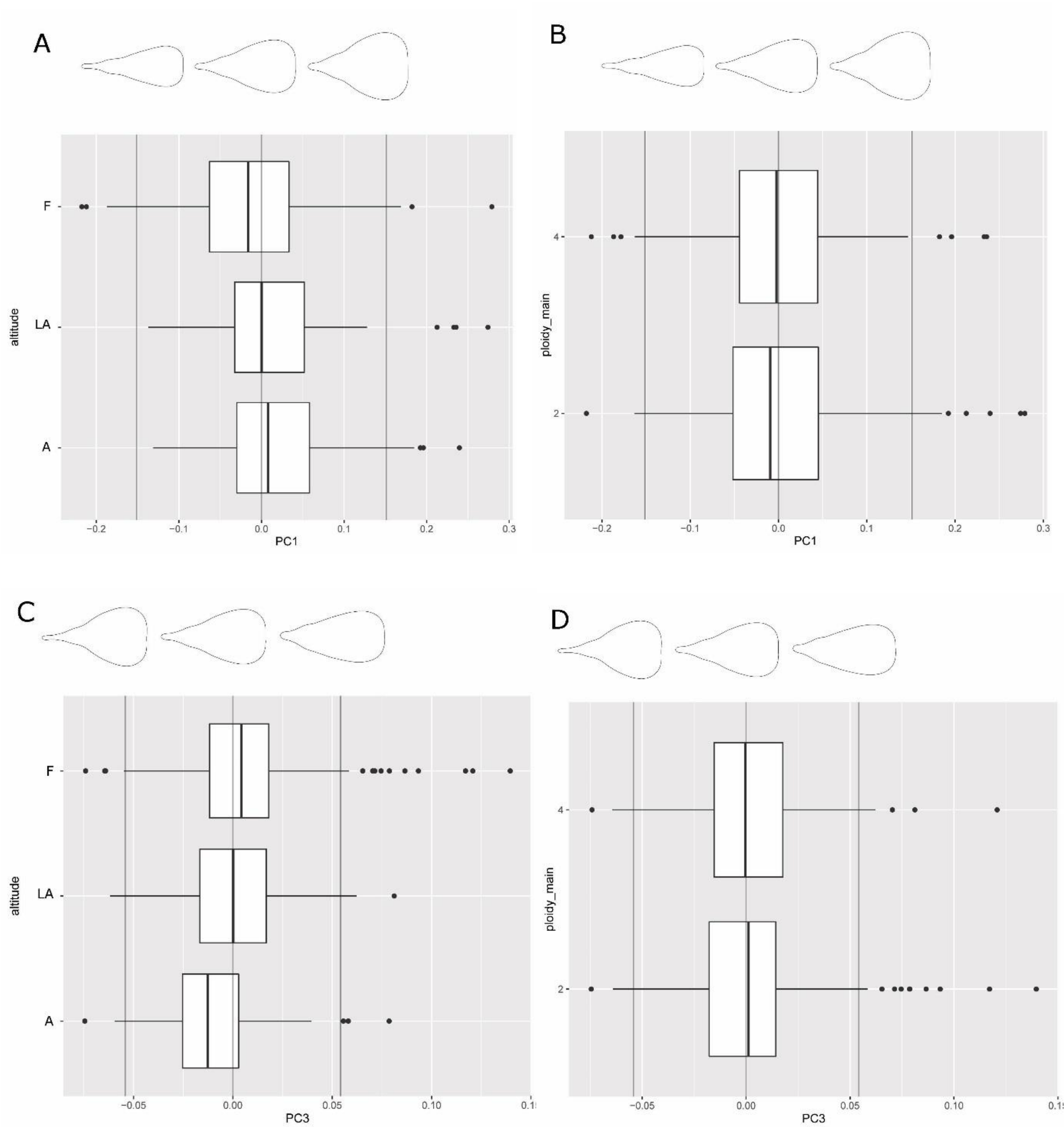

Fig. 13: Variation in petal shape of populations of Tatrean *A. arenosa* along first (A, B) and third PC axes (C, D) in populations categorized according to altitudinal group (A, C) and ploidy level (A, D).

4.2. Ecological differentiation

I characterized the ecological niche of the studied plants by various characteristics of the original collecting sites of 53 populations. Although the characteristics were quite variable, spanning from bioclimatic characteristics and soil data to species composition of the plots surrounding *A. arenosa* and Ellenberg indicator values derived from these species data. The overall pattern was similar. Diploids and tetraploids are not differentiated in ecological preferences in neither of these characteristics (Figs. 14 - 19). Importantly, the used exploratory techniques even do not suggest difference in total variation of the se characteristics what points to similar span of the ecological niche of both cytotypes.

In contrast, two of the three altitudinal groups (alpine and foothill) are clearly differentiated in their stands as documented by Ellenberg indicator values (lower values for soil reaction and temperature in alpine ones, Fig. 14), pH and Ca content (alpine populations only in slightly acidic or neutral stands while the foothill populations showed larger span in soil preferences) and species composition of the surrounding vegetation (Fig. 19). The limestone (lower)alpine populations either occupied sites with intermediate characteristics (Ellenberg values) or those similar to their alpine counterparts (soil characteristics, surrounding species composition),

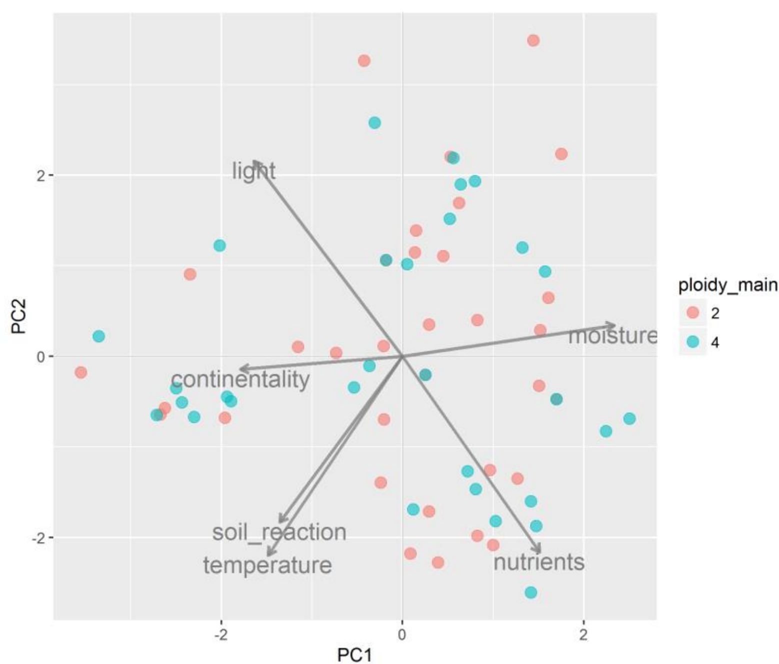

Fig. 14: Plot of PCA analysis of Ellenberg values for 53 populations of *A. arenosa*, colored by ploidy level.

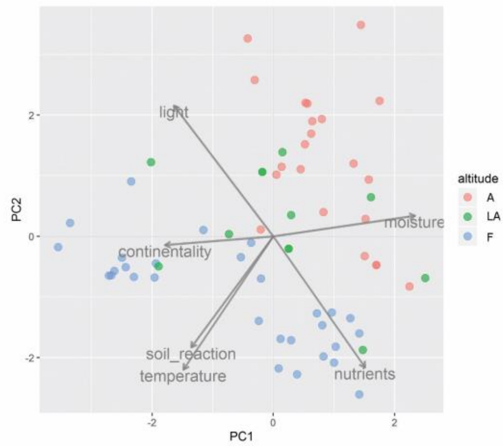

Fig. 15: Plot of PCA Ellenberg values for 53 populations of *A. arenosa*, colored by altitudinal group.

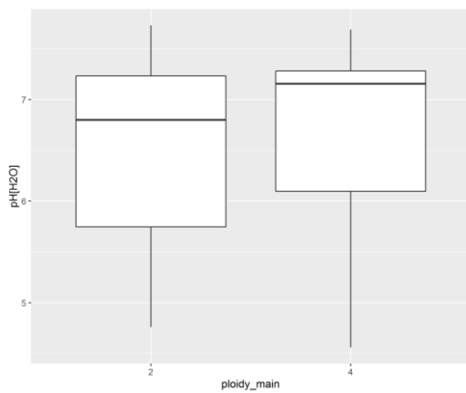

Fig: 16. Variation of 53 populations of *A. arenosa* in pH of soil where grown, colored by ploidy level of populations.

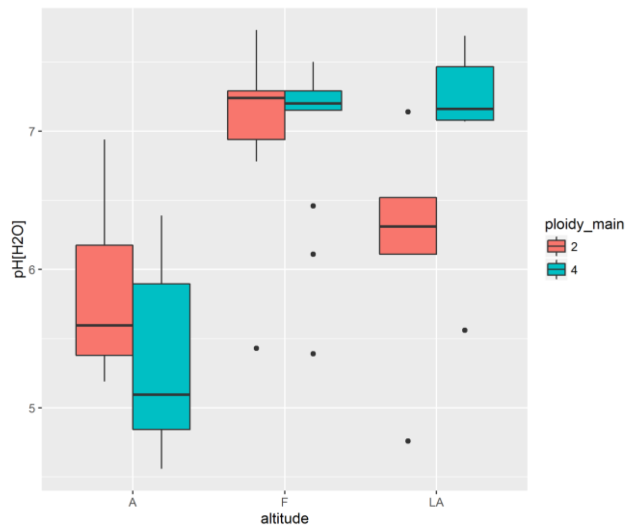

Fig: 17. Variation of pH content for 53 populations of A. arenosa, divided according to altitudinal groups and colored by ploidy level of populations.

Then I calculate the DCA analysis and it does not shown any trend for ploidy level groups (Fig. 18). In the plot where are the data colored by altitudinal group is shown differentiation between alpine and foothill populations, while alpine and limestone (lower) alpine populations are mixed (Fig. 19).

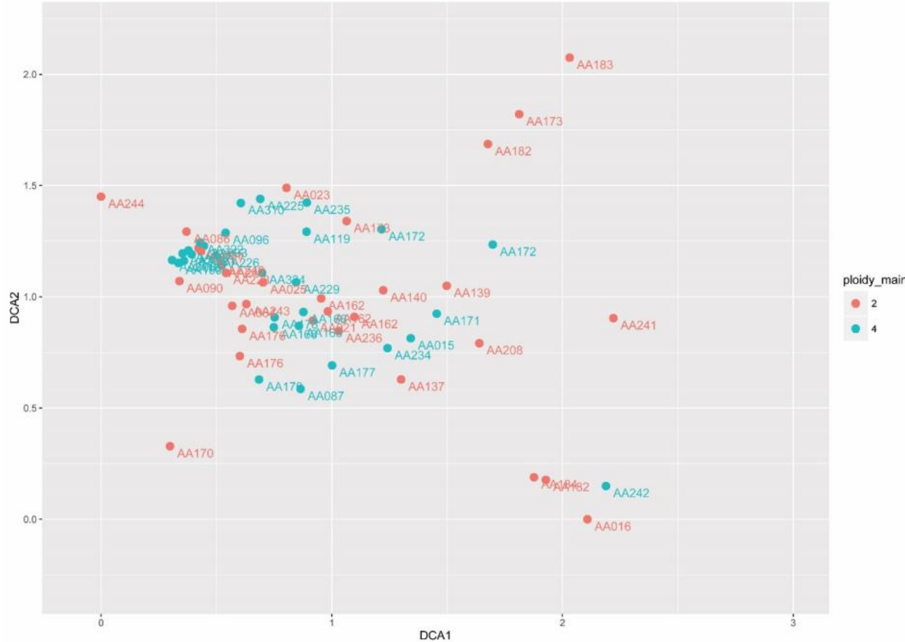

Fig. 18: Variation in species composition in plots in sites occupied by *A. arenosa*, colored by ploidy levels of the *A. arenosa* populations. Indirect ordination (DCA) of vascular plant species composition in 3x3 m plots surrounding *A. arenosa*.

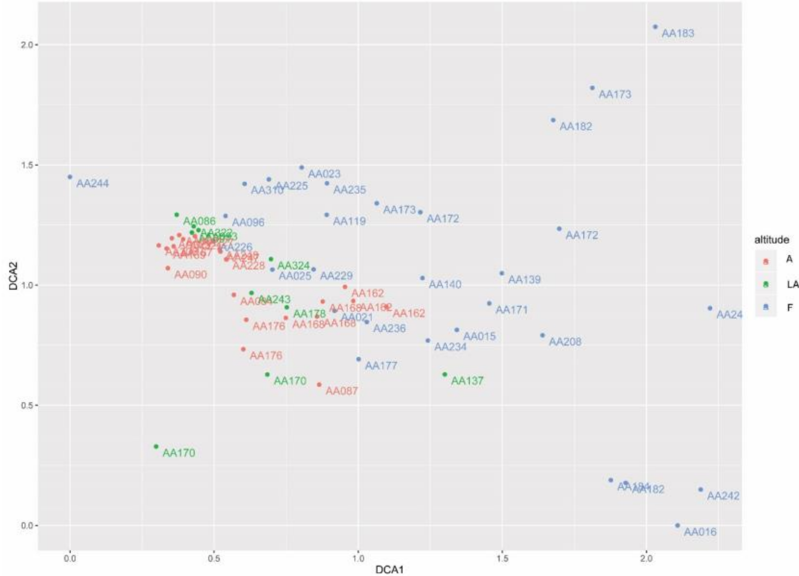

Fig. 19: Variation in species composition in plots in sites occupied by *A. arenosa*, colored by altitudinal groups of the *A. arenosa* populations. Indirect ordination (DCA) of vascular plant species composition in 3x3 m plots surrounding *A. arenosa*.

4.3. Molecular data

In molecular data I observed differentiation between Tatraean foothill and alpine diploids (Fig. 21), while the tetraploids are not differentiated through altitudinal groups (Fig. 22). In whole dataset in comparison of PCoA analysis colored by altitudinal groups (Fig. 20 A) and by ploidy level of populations (Fig. 20 B) I can observe that diploid groups are less genetically related than tetraploid group (Fig. 20. B) and tetraploids are closer related with two lineages of diploids (which is shown in treemix graph Fig. 25).

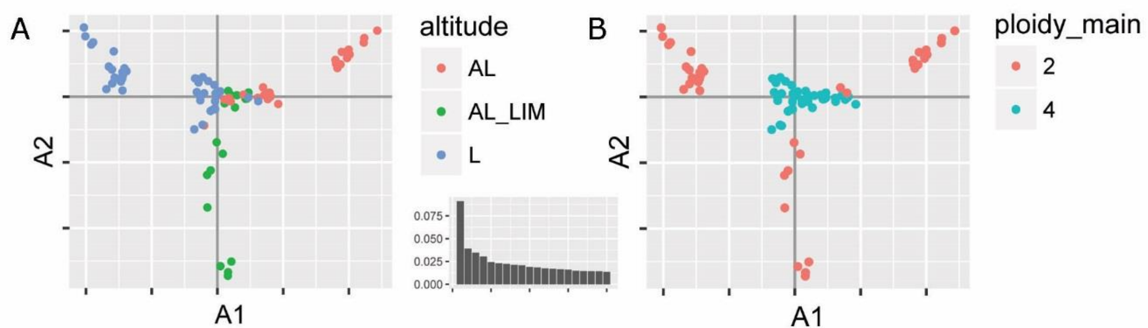

Fig. 20: Plot of PCoA analysis for Tatry dataset, axes A1, A2 shown PC1 and PC2. It is colored by altitude (A) and by ploidy (B), histogram shows proportional contribution in explaining variance by the first 20 axes.

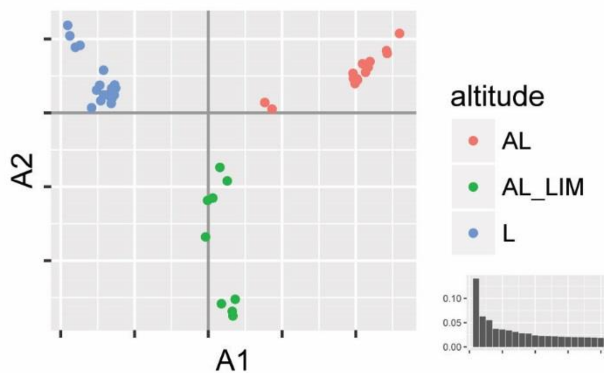

Fig. 21: Plot of PCoA analysis for Tatry diploid dataset, axes A1, A2 shown PC1 and PC2. , histogram shows proportional contribution in explaining variance by the first 20 axes.

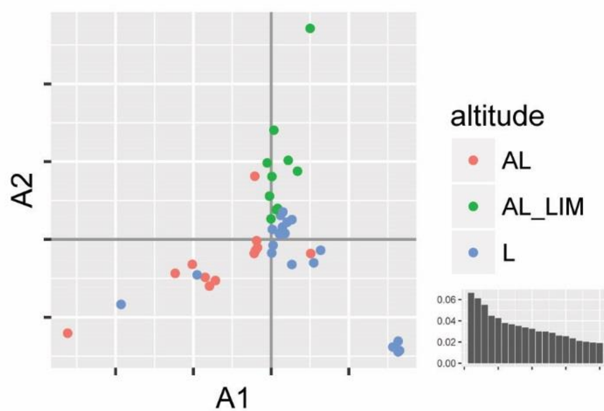

Fig. 22: Plot of PCoA analysis for Tatry tetraploid dataset, axes A1, A2 shown PC1 and PC2. It is colored by altitude, histogram shows proportional contribution in explaining variancethe first 20 axes

In K-means clustering there is shown for tetraploids just differentiation of specially most distant population AA235 Kamenica (Fig. 23).

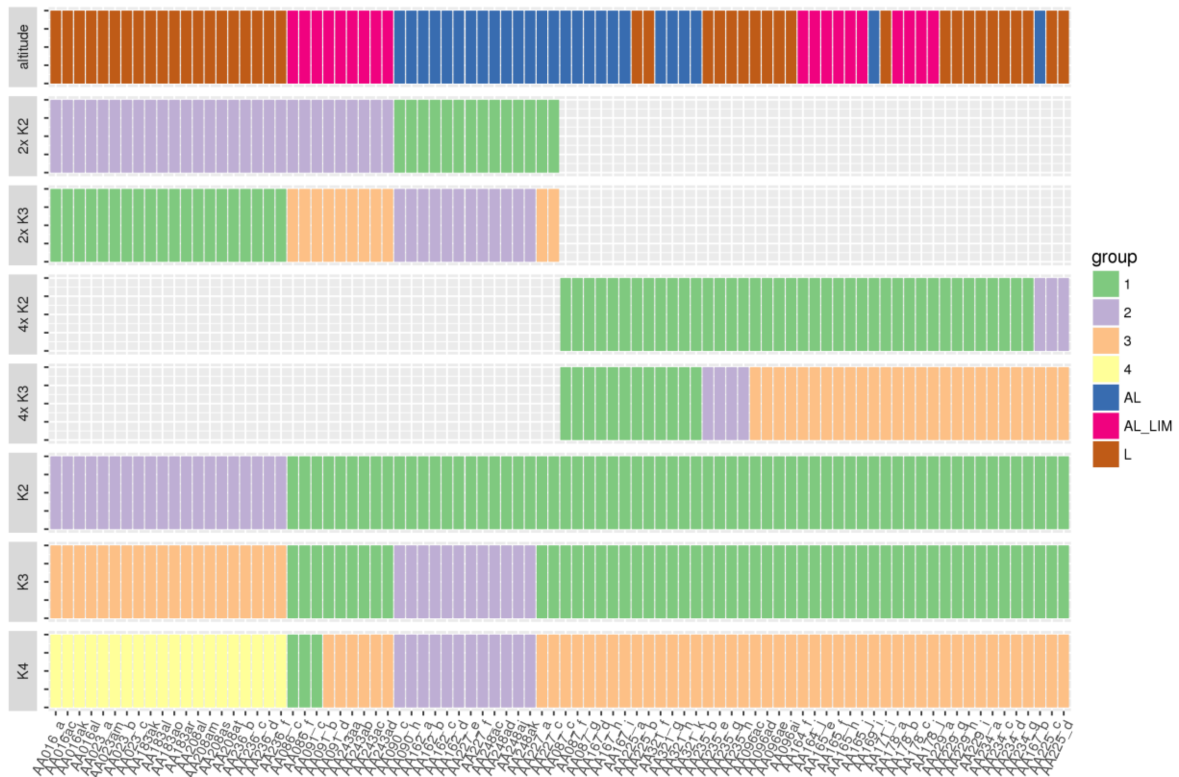

Fig. 23: Cluster assessment figure for Tatry dataset (K2 – K4), for diploid dataset (2x K2, 2x K3) and for tetraploid dataset (4x K2, 4x K3). Altitude shows Al – alpine population, AL_LIM – alpine limestone population and L – foothill populations.

In the neighbor-joining network for Tatry Mts. population we can see genetic differentiation between diploids and tetraploids (Fig. 24).

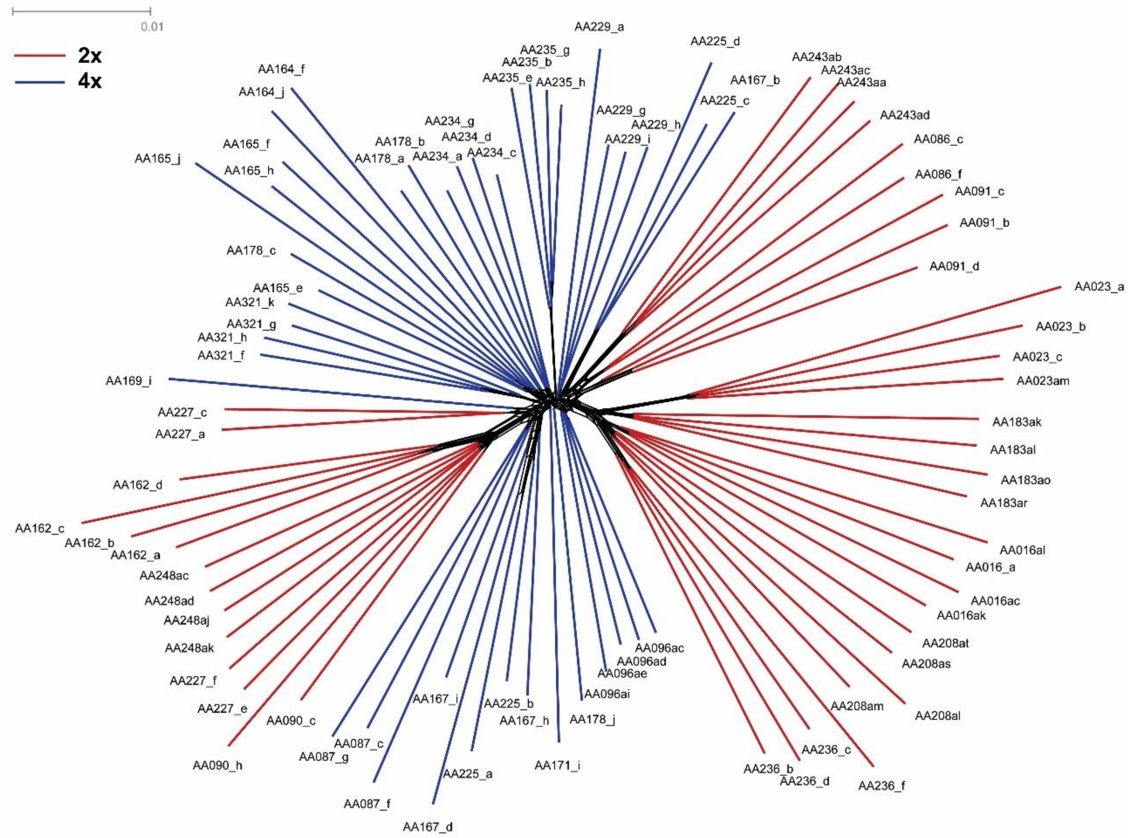

Fig. 24: Genetic distances between Tatras diploids and tetraploids shown in neighbor-joining network (diploids are red, tetraploids are blue).

Table 2: AMOVA analysis

percentage of variation among populations	percentage of variation among groups	p-value_level1	p-value_level2	grouping	dataset
47.6	0	<0.001			Tatry2x
28.2	19.3	0.037	<0.001	A vs. AL_LIM vs. F	Tatry2x
39.7	0	<0.001			Tatry4x
31.6	8.1	<0.001	0.045	A vs. AL_LIM vs. F	Tatry4x
46.1	0	<0.001			Tatry
37.5	8.5	<0.001	<0.001	A vs. AL_LIM vs. F	Tatry
43.5	2.5	<0.001	0.004	2x vs. 4x	Tatry

In the Treemix population graphs, the tetraploid group was sister to the alpine diploids (with high bootstrap support) but was also linked to the foothill diploid group by a migration edge, suggesting admixture (Fig. 25).

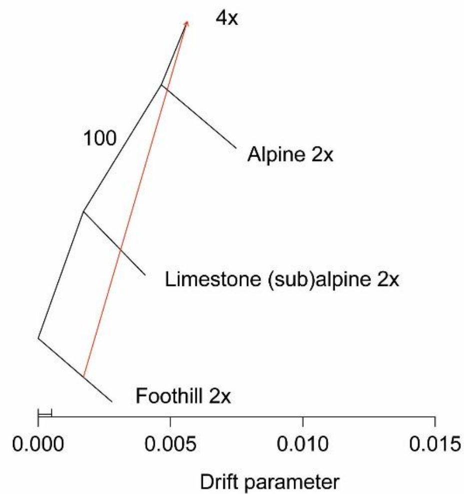

Fig. 25: Treemix maximum likelihood graph (73,059 SNPs) showing relationships among main lineages of *A. arenosa* with one migration edge (the single bootstrap support of > 50% is above the corresponding branch).

4.4. Morphological variation of plants from experiment in common conditions

I inspected data from 8 populations cultivated in standardized conditions in growth chamber through boxplot (Fig. 26) and histogram colored by altitude (Fig. 27) and by ploidy (Fig. 28). Importantly, plants from the experimental populations confirmed the observations made in the field-collected accessions. Plants cultivated from seeds collected in both alpine and foothill stands exhibited contrasting morphology even in common cultivations, as is apparent from non-overlapping values for certain characters (Fig. 27) and also from indirect ordination (Fig. 29). Importantly, the characters differentiating alpine and foothill (limestone (lower)alpine populations were not included in this experiment) are the same as those recorded in the morphological screen of the filed-collected plants. Ploidy level did not played role in morphological differentiation (Fig. 30) similarly as was observed in the individuals collected in field.

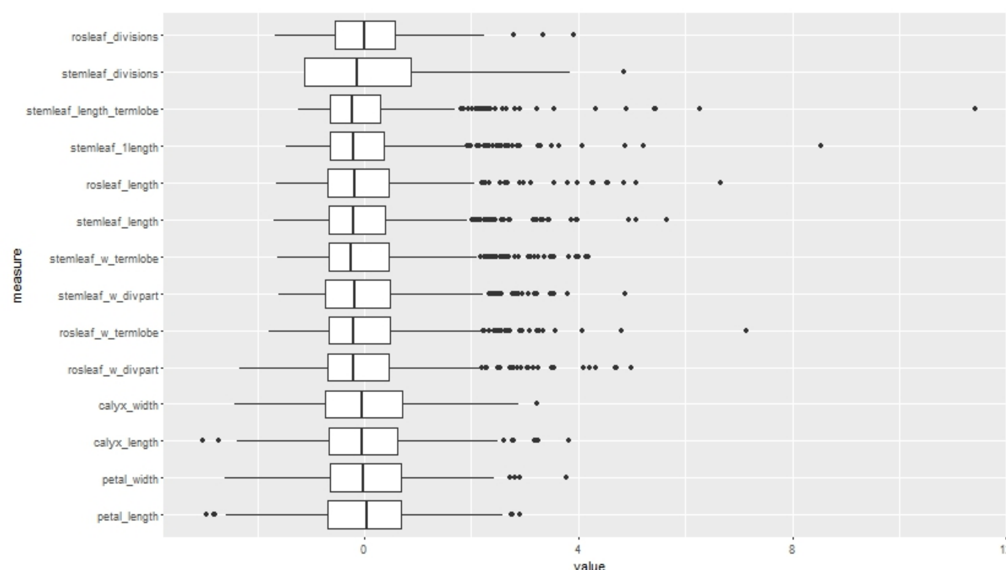

Fig. 26: Data summarize in boxplot standardized for chosen 14 measured characters.

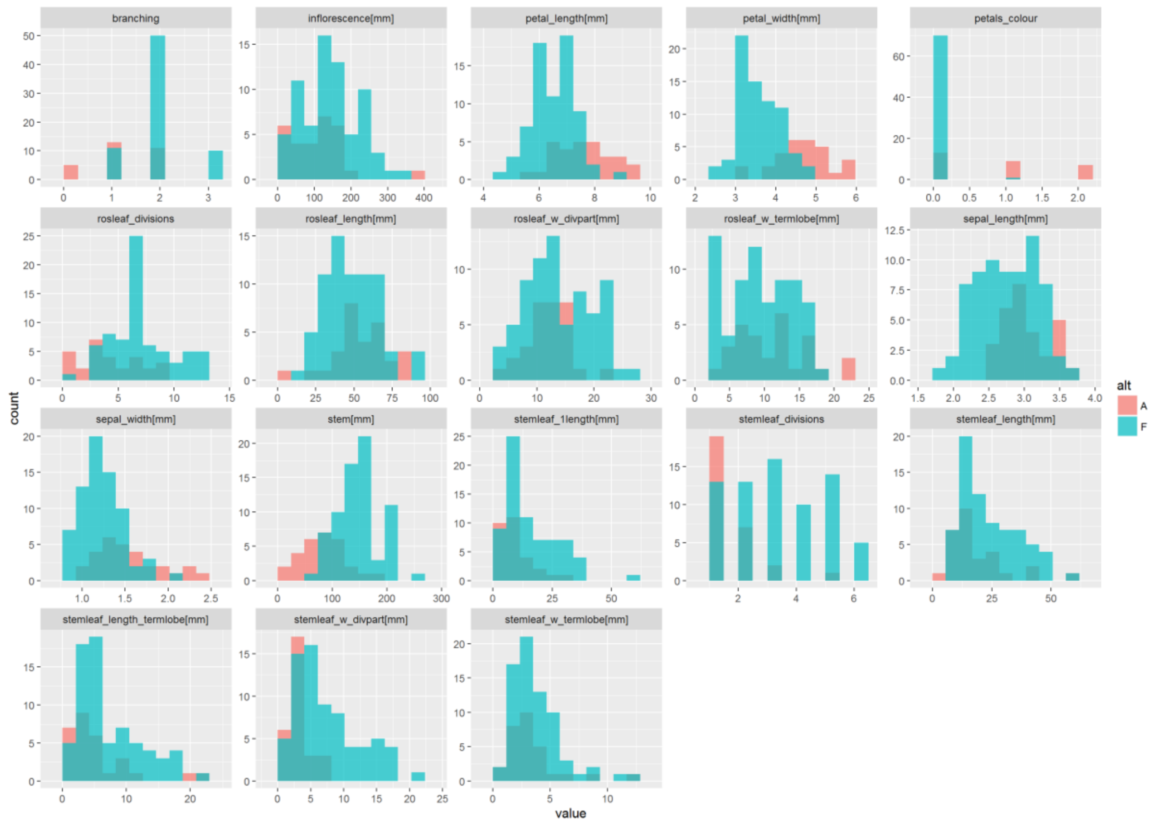

Fig. 27: Overview histograms colored by altitudinal groups.

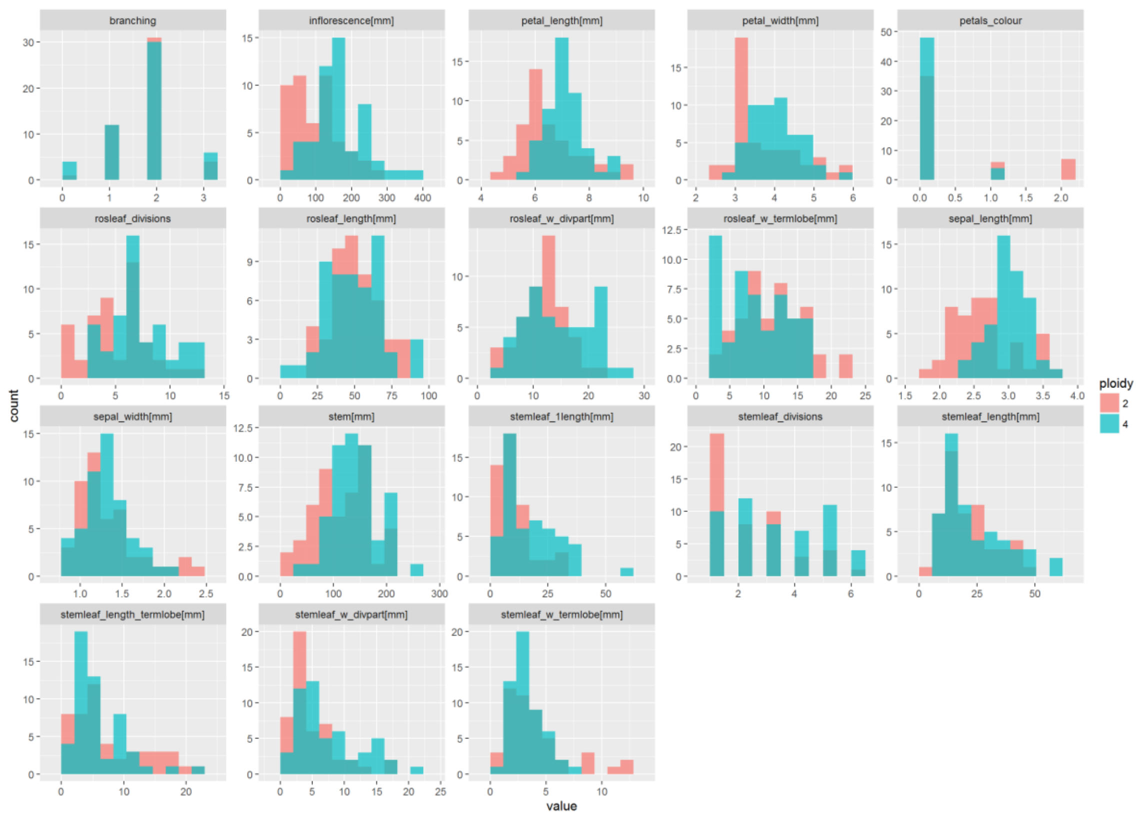

Fig. 28: Overview histograms colored by ploidy levels of populations.

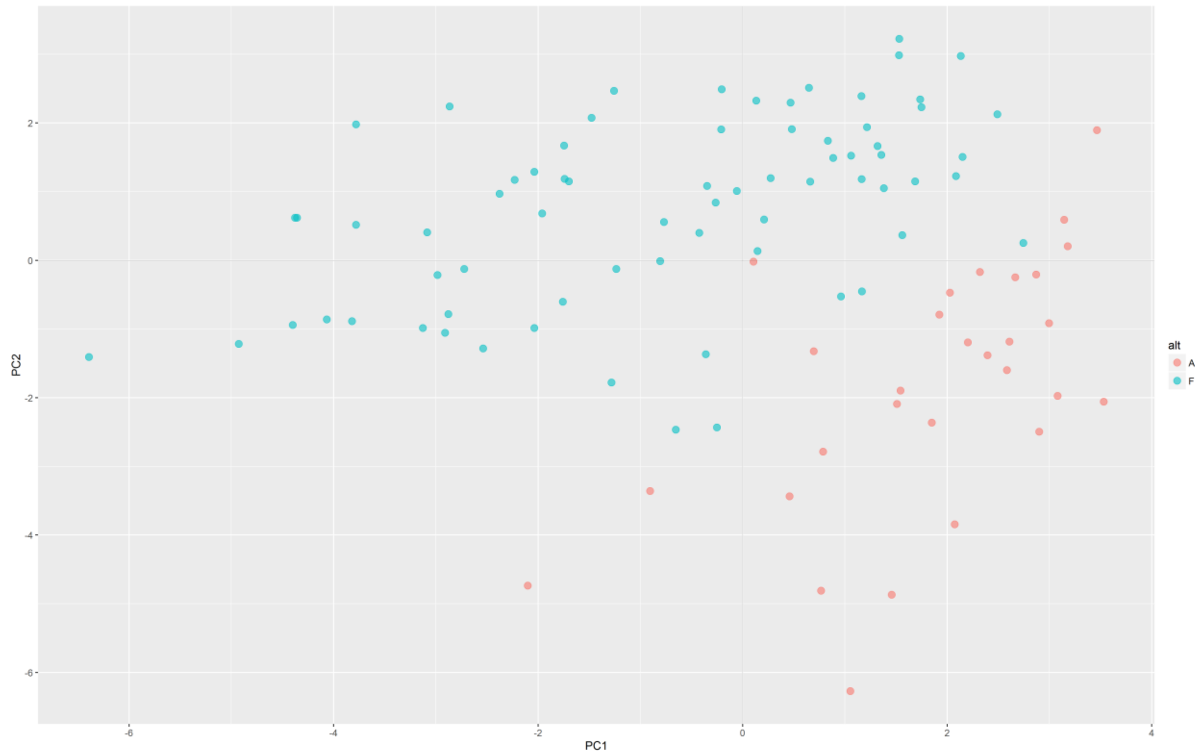

Fig. 29: PCA analysis plotted for 8 experimented populations, colored by altitudinal groups.

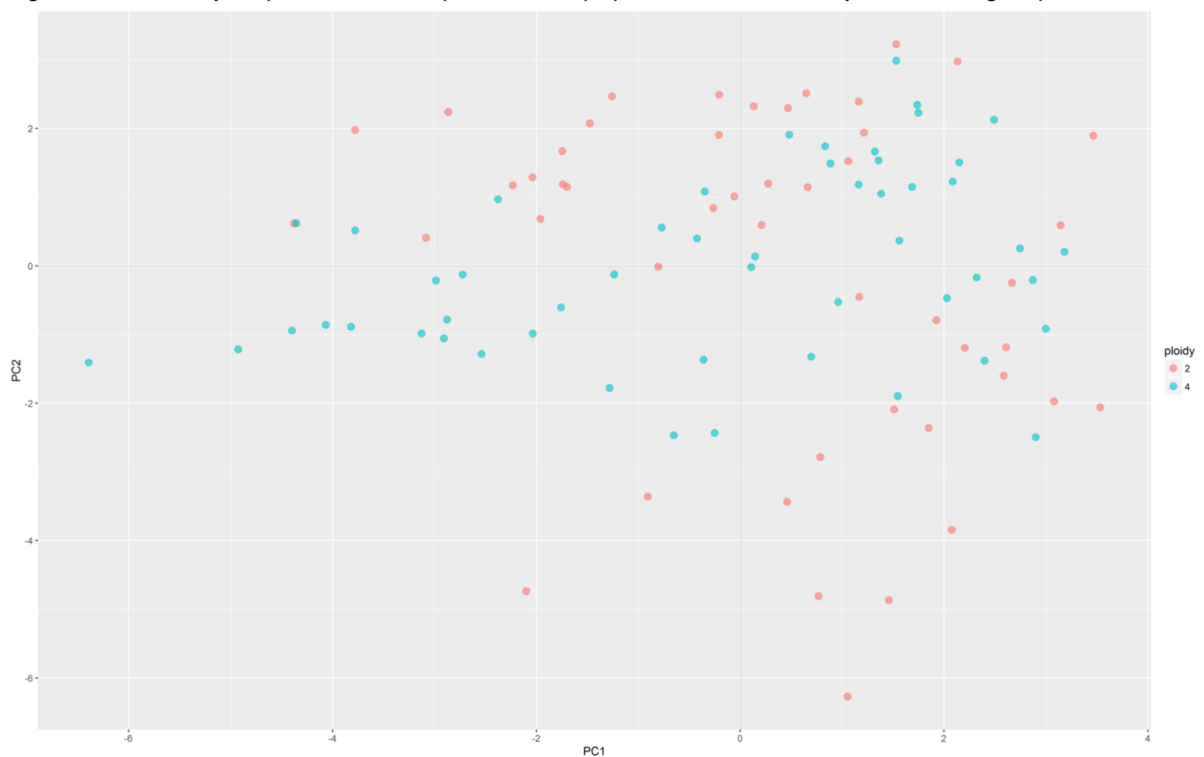

Fig. 30: PCA analysis plotted for 8 experimented populations, colored by ploidy level of populations.

4.5. Data synthesis (variation partitioning)

Here are displayed results from data synthesis for relative contribution of environmental parameters to the morphological differentiation.

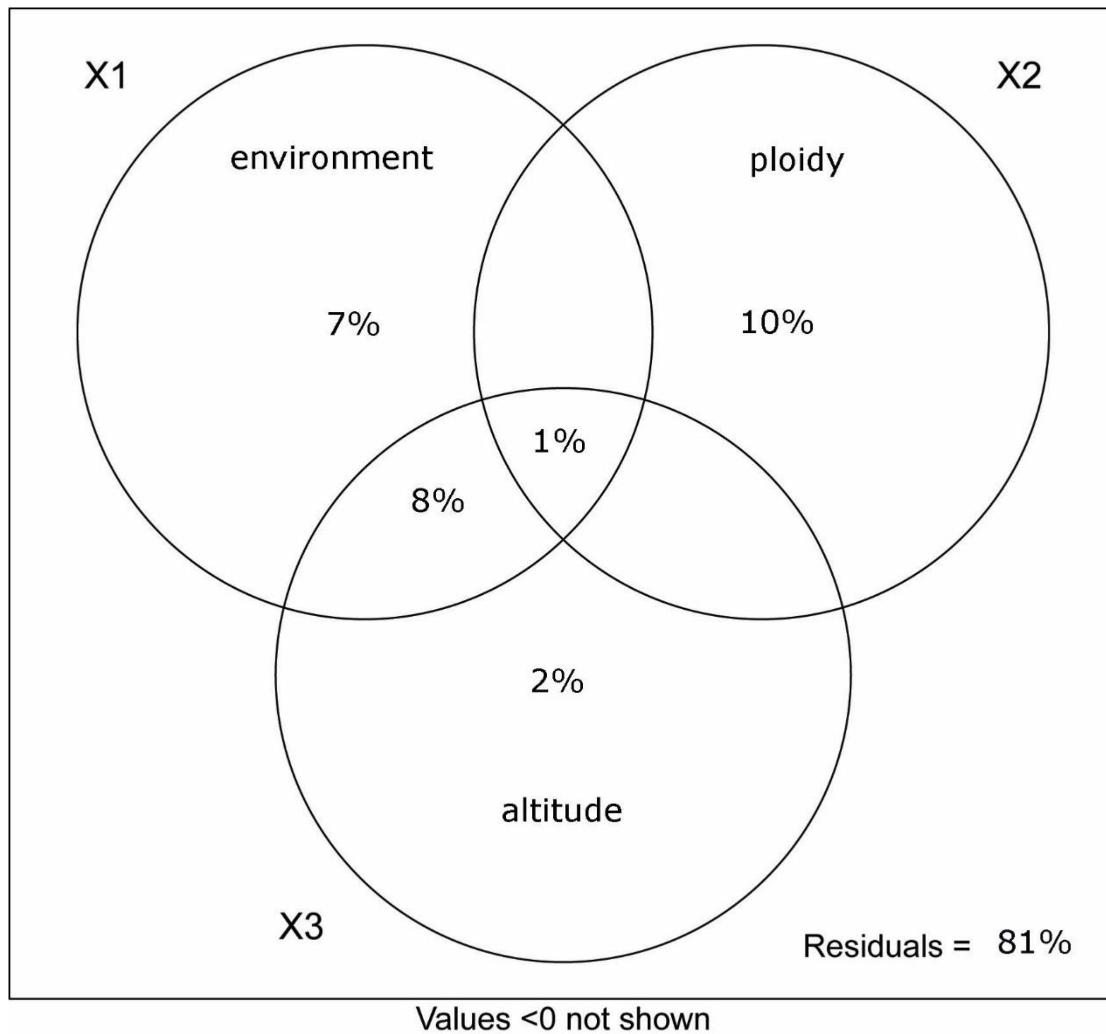

Fig. 31: Relative contribution of environmental parameters (bioclimatic, soil and indicator values derived from species composition data), ploidy level and altitude to the SNP-based genetic variation among 23 *A. arenosa* populations inferred by variation partitioning through direct ordinations (redundancy analyses).

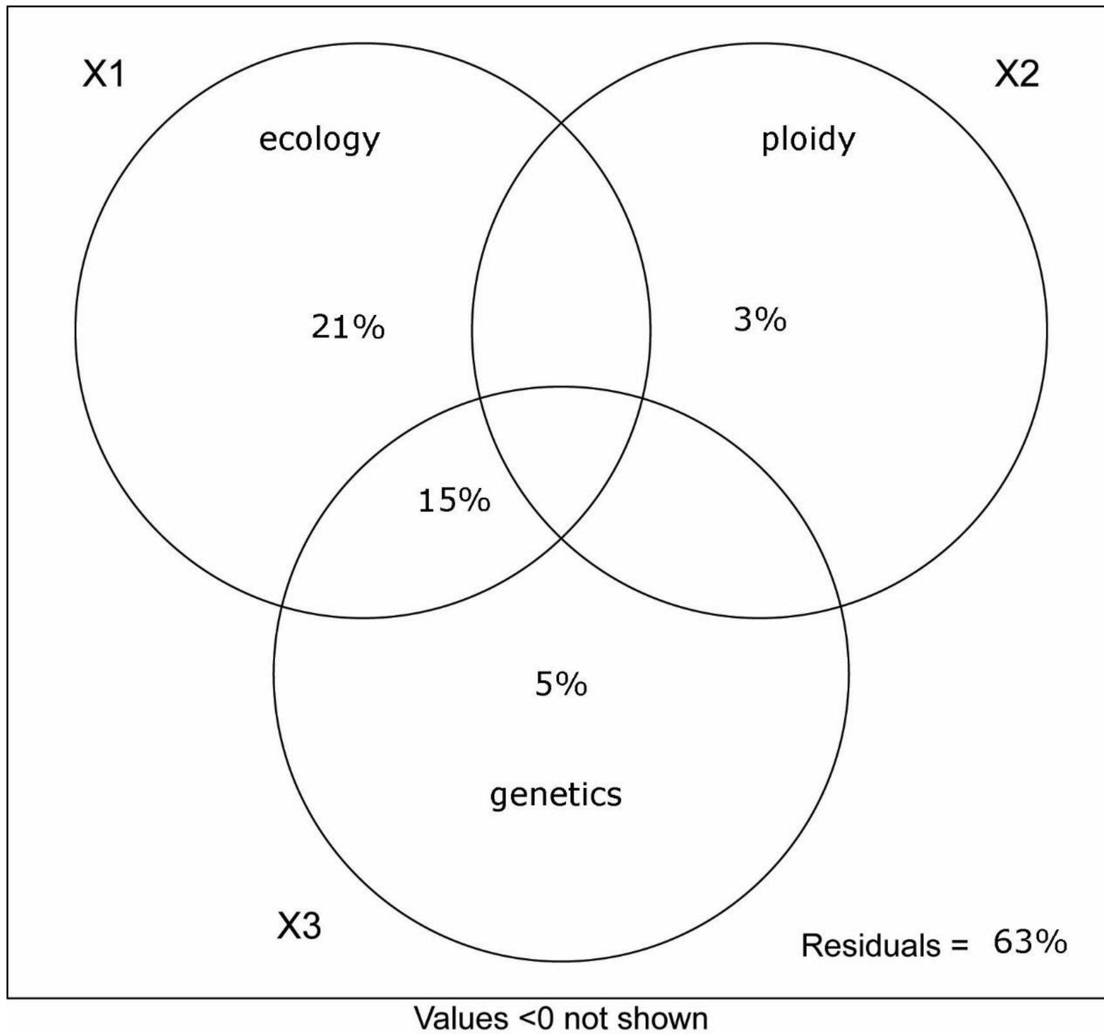

Fig. 32: Relative contribution of environmental parameters (bioclimatic, soil and indicator values derived from species composition data), ploidy level and genetic structure (SNP-based genetic distances) to the morphological variation of 23 *A. arenosa* populations inferred by variation partitioning through direct ordinations (redundancy analyses).

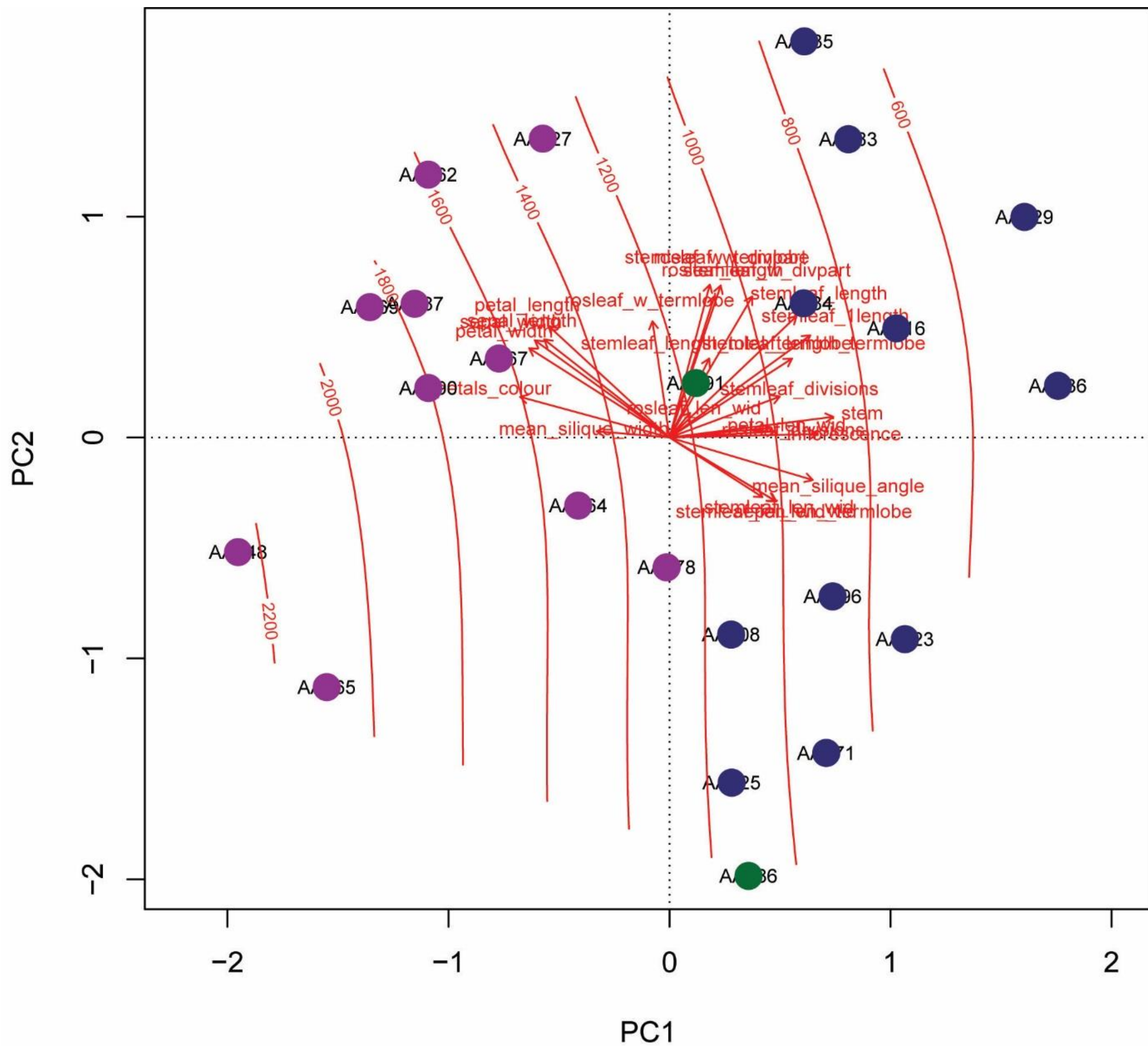

Fig. 33: Morphological differentiation of 23 *A. arenosa* populations related to the altitudinal variation. A Principal component analysis the populations (population means are displayed) coloured according to their assignment to the three altitudinal groups (blue – foothill, green – limestone (lower) alpine, purple – alpine); brown contour lines denote altitude that was passively fitted onto the ordination plot by thin plate spline generalized additive model.

5. Discussion

The studied complex A. arenosa is unique, because of close related of model species A. thaliana, which provide a benefit in using various mainly molecular methods, and include naturally growing diploid and tetraploids lineages (Bomblies et Madlung 2014). Just a genetic closeness diploid and tetraploid cytotype allows to study direct causes of polyploidization in natural populations.

Ecological and morphological consequences of polyploidization are still not enough explored territory of plant evolution ecology (e.g. Ramsey et Ramsey 2014) just because of (i) focus most of studies to allopolyploids which combined consequences of polyploidization and hybridization and (ii) insufficient or even absent combination of genetical, morphological and ecological data for the same study system.

The A. arenosa populations in Tatry Mts. are unique within the entire model genus Arabidopsis due to common co-occurrence of diploid and tetraploid individuals in the alpine conditions within a small area (Kolář et al. 2015). In the rest of the distribution area of this species, it is only a single cytotype (tetraploid) that reach occur high altitudes. Common occurrence of diploid and tetraploid cytotyp in Tatry Mountains make an ideal model system for studying genome duplication and adaptation in context of alpine conditions.

The presented thesis explore causes of pronounced morphological differences alpine and foothill populations in complex A. arenosa in the Westhern Carpathians (Tatry Mountains, previous local studies by Měsíček 1970) and focuses to so far non applied connection of context of approach joining of ecological, genetetic and ploidy level data.

A. Morphological differentiation and its causes.

In Tatras *A. arenosa* populations we can find pronounced morphological differentiation between individuals from alpine and foothill areas. This was recorded already by Měsíček (1970) and he divided the morphotypes from Slovakia to several (sub)species, most of them, however, without a valid description.(Schmickl et al. 2012).

Because of this still unsolved taxonomy in the A. arenosa complex I treat this group as one widely defined species *A. arenosa*. For addressing my hypotheses I divided populations to three groups reflecting their ecological-geographical occurrence (see part 4.1.) and partly also genetic background (only within the diploid cytotype, though).

The field-collected individuals from alpine and foothill areas exhibit obvious morphological differentiation as was apparent in the PCA (Fig. 9 A). Importantly, while the tetraploid cytotype show no clear genetic separation, this morphological differentiation correlates with a major genetic split within the diploid cytotype. Specifically, the most strongly (both morphologically and genetically) differentiated diploid populations from the High Tatras alpine populations are separated from their lowland counterparts as well as from the populations occupying limestone areas in the higher elevations of Western and Belanske Tatry Mts (see Fig. 20, differentiation in the AMOVA analysis, Table 2).

Although the correspondence of genetic and morphological data (observed in diploids) may imply ecotypic differentiation of the alpine type(s) in diploids or even incipient speciation triggered by eco-spatial isolation and/or selection in the alpine environment , the observed morphological differentiation may still reflect only a plastic response of the “foothill-like” plants to the dramatically distinct environment (see also Fig. 19). Indeed, for the role of phenothypic plasticity speak that major drivers explaining the morphological variation of the field collected plants are the ecological factors with relative contribution 36% whereas the genetic structure has 5% relative contribution (Fig. 32).

A cultivation in common conditions may provide important clues in separating the effect of heritable (i.e. (epi)genetic background) and non-heritable (i.e., plasticity) factors. From the results of a pilot cultivation of 8 populations we can see that individuals from alpine populations cultivated in conditions for foothill populations (see part 4.4.) retain their alpine-like morphology at least in the first generation of cultivation (Fig. 29). In sum, my study brings support for genetic background as a major driver in the origin of the alpine phenotype, although certain role of phenotypic plasticity could not be excluded. In addition, my experiment was based on progeny cultivated from filed-collected seeds and the maternal effect could not be excluded

B. Colonization of alpine environment.

From the genetic results we can see that alpine and foothill diploid populations are different (Fig. 21). A certain differentiation of diploid foothill and alpine population from Western Carpathians was already suggested by a previous study (Kolář et al. 2015), however, in the range-wide view, the populations from Western Carpathians regardless their altitude of origin are still much more related to each other than to diploids from other parts of the *A. arenosa*

range. This suggests that the foothill-alpine differentiation (accounting Table 2) is still relatively recent, probably reflecting the postglacial colonization of the formerly glaciated alpine areas in the Tatras.

Importantly, the observed genetic differentiation between the diploid and tetraploid populations (Figs. 20, 23) and within the diploid populations themselves suggest that the alpine populations might have originated several times independently. Firstly, two distinct diploid lineages (called here “alpine” and “limestone (lower)alpine”) colonized high altitudes over ca 1500 m. a.s.l., although each of them grow in very different environments. Secondly, the tetraploids, although their relationship to the diploids remain equivocal (see the next sections) exhibit clearly another group that was able to independently colonize the alpine environments in the Western (majority of the alpine pops in Fig. 22) and Lower Tatras (pop AA306 Dumbier). Unfortunately, the low divergence in tetraploid populations do not allow further distinguishing whether alpine populations from those two spatially isolated mountain ranges originated once or twice independently. Repeatedly originated plant ecotypes provide very promising systems for addressing molecular basis of plant adaptation to environmental challenges, yet it is a still rarely documented phenomenon in plants (but see e.g., Trucchi et al 2016). A. arenosa, due to availability of three closely realated species with annotated genomes (see. http://phytozome.jgi.doe.gov/) provides unique opportunity for follow-up study of genomic consequences of polyploid speciation and adaptation to extreme conditions in the alpine environment.

C. What is the role of genome duplication in colonization of alpine environments?

We did not confirm any strong differentiation neither in morphological traits nor in ecological preferences between the two cytotypes within the study area. Both diploid and tetraploid populations exhibit overall similar morphology (although slight differences were observed in common cultivation, Fig. 28), occupy entire span of habitats along the siliceous-calcareous gradient (Fig. 17) and are associated with the same accompanying species (Fig. 14, Fig. 18).

Although polyploidy is traditionally considered a key trait enabling further niche expansion within a species (e.g., Levin 2002), recent careful (meta)analyses show that no difference or even the opposite trend (diploid superiority) may be often present. Indeed, the results may strongly depend on the species studied (Glennon et al. 2014 EcolLEtt) or spatial grain

investigated (Kirchcheimer et al. 2015 JBiogeogr). My study, documenting no obvious ecological or phenotypic differences among the cytotypes along an altitudinal gradient, brings another piece of evidence supporting this recently emerged view.

D. Origin of alpine tetraploids.

In the overall view for results from genetic PCoA analysis for 23 Tatry Mts. populations (Fig. 20 B), we can observe a genetic differentiation between the diploid and tetraploid populations. The diploid populations groups are less similar to each other than to the tetraploid populations group. Which is saw for diploids in Kolář et al. 2015. From the treemix graph (Fig. 25) it seems that tetraploid populations originated by admixture between two lineages – alpine diploid and foothill diploid, although this should be tested explicitly, e.g. using coalescent simulations (Excoffier et al 2013 fastsimcoal)

We did not identify any clear structure within the tetraploid populations. There is no differentiation of alpine-foothill group in AMOVA test (Table 2) and in the K-means clustering there is shown just separation of the geographically and also genetically most distant population AA235 Kamenica (Fig. 23).

Based on these results we can reject the hypothesis that the alpine Tatry Mts. tetraploids represent a direct autopolyploid descendants of their alpine diploid counterparts. Although the involvement of alpine diploids (e.g. through introgression) in origin of alpine tetraploids could not be excluded completely. As next the very close genetic position within the tetraploid cytotypes as whole supports a recent origin of the alpine tetraploids from their foothill tetraploid counterparts.

6. Conclusions

In whole the thesis provide integrated view for evolutionary and ecological consequences of polyploidization in Tatras alpine populations and values in detail their morphological differences, too.

Based on morphological measurements and their appreciation it was possible to find out probably causes of obvious morphological differentiation between alpine and foothill populations in Tatry Mts. which was observed already by Měsíček 1970.

For diploids we can observe a strong differentiation in morphology and also strong genetical differentiation. While for tetraploids cytotype no clear genetic separation is evident and the morphological differentiation correlates with a major genetic split within the diploid cytotype.

In question about colonization of alpine environment it was researched that alpine populations are not closer altogether more than with foothill populations so they could colonized alpine environment repeatly.

For results from experimented planting is need planting next generations of plants to confirm hypothesis about genetic determination.

For the future prepared paper I will tested the appearing differences in PCA analyses.

Citations

Álvarez I, Wendel JF. **2003**. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29**: 417–434.

Clauss MJ, Koch MA. **2006**. Poorly known relatives of Arabidopsis thaliana. *Trends in Plant Science* **11**: 449–459.

Comai L. **2005**. The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* **6**: 836–846.

Elder JF Jr, Turner BJ. **1995**. Concerted evolution of repetitive DNA sequences in eukaryotes. *Quarterly Review of Biology* **70**: 297–320.

Hoffmann MH. **2005**. Evolution of the Realized Climatic Niche in the Genus: Arabidopsis (brassicaceae). *Evolution* **59**: 1425–1436.

Hollister JD, Arnold BJ, Svedin E, Xue KS, Dilkes BP, Bomblies K. **2012**. Genetic Adaptation Associated with Genome-Doubling in Autotetraploid Arabidopsis arenosa. *PLoS Genet* **8**: e1003093.

Hunter B, Bomblies K. **2010**. Progress and Promise in using Arabidopsis to Study Adaptation, Divergence, and Speciation. *The Arabidopsis Book / American Society of Plant Biologists* **8**.

Kolník M, Marhold K. **2006**. Distribution, chromosome numbers and nomenclature conspect of Arabidopsis halleri (Brassicaceae) in the Carpathians. *Biologia* **61**: 41–50.

Levin DA. **2002**. *The Role of Chromosomal Change in Plant Evolution*. Oxford University Press.

Mable BK, Beland J, Di Berardo C. **2004**. Inheritance and dominance of self-incompatibility alleles in polyploid Arabidopsis lyrata. *Heredity* **93**: 476–486.

Marhold K, Lihová J. **2006**. Polyploidy, hybridization and reticulate evolution: lessons from the Brassicaceae. *Plant Systematics and Evolution* **259**: 143–174.

Měsíček J. **1967**. The chromosome morphology of *Arabidopsis thaliana* (L.) Heynh. And some remarks on the problem of *Hylandra suecica* (Fr.) Löve. *Folia Geobotanica et Phytotaxonomica* **2**: 433–436.

Měsíček J. **1970**. Chromosome counts in *Cardaminopsis arenosa* agg. (Cruciferae). *Preslia* **42**: 225–248.

Měsíček J, Goliášová K. **2002**. *Cardaminopsis* (C. A. Mey.) Hayek. In: Goliašová K, Šípová H, eds. *Flóra Slovenska*. Bratislava: Veda, 388–415.

O'Kane SL Jr, Schaal BA, Al-Shehbaz IA. **1996**. The Origins of Arabidopsis suecica (Brassicaceae) as Indicated by Nuclear rDNA Sequences. *Systematic Botany* **21**: 559–566.

Parisod C, Holderegger R, Brochmann C. **2010**. Evolutionary consequences of autopolyploidy. *New Phytologist* **186**: 5–17.

Petit C, Bretagnolle F, Felber F. **1999**. Evolutionary consequences of diploid–polyploid hybrid zones in wild species. *Trends in Ecology & Evolution* **14**: 306–311.

Ramsey J, Schemske DW. **1998**. Pathways, Mechanisms, and Rates of Polyploid Formation in Flowering Plants. *Annual Review of Ecology and Systematics* **29**: 467–501.

Schmickl R, Koch MA. **2011**. Arabidopsis hybrid speciation processes. *Proceedings of the National Academy of Sciences* **108**: 14192–14197.

Schmickl R, Paule J, Klein J, Marhold K, Koch MA. **2012**. The Evolutionary History of the Arabidopsis arenosa Complex: Diverse Tetraploids Mask the Western Carpathian Center of Species and Genetic Diversity. *PLoS ONE* **7**: e42691.

Al-Shehbaz IA, O'Kane SL. **2002**. Taxonomy and Phylogeny of Arabidopsis (Brassicaceae). *The Arabidopsis Book / American Society of Plant Biologists* **1**.

Shimizu-Inatsugi R, Lihová J, Iwanaga H, Kudoh H, Marhold K, Savolainen O, Watanabe K, Yakubov VV, Shimizu KK. **2009**. The allopolyploid *Arabidopsis kamchatica* originated from multiple individuals of *Arabidopsis lyrata* and *Arabidopsis halleri*. *Molecular Ecology* **18**: 4024–4048.

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS. **2009**. Polyploidy and angiosperm diversification. *American Journal of Botany* **96**: 336–348.

Suda J, Krahulcová A, Trávnícek P, Krahulec F. 2006. Polyploidy Level versus DNA Ploidy Level

population	locality name	altitude	ploidy level	molecular anal	cultivating exp	morpho analysi	shape analysis	soil analysis	releve analysis	region	locality descripti	habitat	collected	latit	longit	altit
AA016	Sucha Bela	F	2x	x	x	x	x		x	Košický kraj	Podlesok, rocks	limestone rocks	E. Záveská, J. K	48.96030556	20.38327778	600
AA021	Tatranska Kotlin	F	2x			x	x	x	x	Prešovský kraj	Tatranská kotlin	slopes and lime	E. Záveská, F.	49.22891667	20.31169444	900
AA023	Besenova	F	2x	x		x	x	x	x	Žilinský kraj	Bešeňová, trave	travertine rock	E. Záveská, F.	49.10725	19.43466667	574
AA025	Strecno	F	2x					x	x	Žilinský kraj	Strečno, limesto	limestone rocks	E. Záveská, F.	49.17408333	18.8617	425
AA027	Sulov	F	4x					x		Žilinský kraj	Súľov, slopes b	foest on rocky sl	F. Kolář	49.17551944	18.58361944	600
AA084	Velicka dolina	A	2x		x	x	x	x	x	Prešovský kraj	Vysoké Tatry, w	alluvial gravel,	F. Kolář, E. Záv	49.162	20.15419444	1823
AA086	Osobita	LA	2x	x		x	x	x	x	Žilinský kraj	Zuberec, Zvero	exposed rocks,	F. Kolář, E. Záv	49.25925	19.72183333	1552
AA087	Placlivy Rohac	A	4x	x	x	x		x	x	Žilinský kraj	Zuberec, Zvero	exposed rocks	F. Kolář, E. Záv	49.19702778	19.74477778	2031
AA090	Zelene pleso	A	2x	x	x	x	x	x	x	Prešovský kraj	Vysoké Tatry, s	scree, wet rocks	F. Kolář, E. Záv	49.20652778	20.21505556	1625
AA091	Kopske sedlo	LA	2x	x		x	x	x	x	Prešovský kraj	Ždiar, Belianské	exposed rocks,	F. Kolář, E. Záv	49.22991667	20.21902778	1751
AA093	Krakova hola	LA	4x					x	x	Žilinský kraj	Nízke Tatry, pat	eroded slope ab	F. Kolář, S. Špa	48.97680556	19.63369444	1630
AA094	Demanovska ja	F	4x			x	x			Žilinský kraj	Nízke Tatry, roc	shady rocks	F. Kolář, S. Špa	48.99839444	19.58376667	850
AA096	Liptovsky Jan	F	4x	x		x	x	x	x	Žilinský kraj	Nízke Tatry, Lipt	alluvial gravel, e	F. Kolář, S. Špa	49.01075	19.67327778	744
AA119	Drevenik	F	4x			x	x	x	x	Košický kraj	Dreveník hill, 1	rock on the egd	M. Lučanová, E.	48.98296667	20.77745	552
AA137	Velka Stozka	LA	2x					x	x	Banskobystrický	Muráň, eastern	half-shaded roc	F. Kolář	48.785498	19.977113	1339
AA162	Skok	A	2x	x		x	x	x	x	Prešovský kraj	Vysoké Tatry, M	wet rocks, alluvi	F. Kolář, M. Luč	49.15341	20.04584	1750
AA164	Giewont	LA	4x	x		x	x	x		Województwo	Zakopane, rock	exposed rocks,	F. Kolář, J. Sma	49.250383	19.934022	1838
AA167	Rackove plesa	A	4x	x		x	x		x	Žilinský kraj	Pribylina, Račko	open places in	M. Lučanová, K.	49.200048	19.804658	1690
AA168	Tri Kopy	A	4x		x		x	x	x	Žilinský kraj	Zuberec, Zvero	wet rocks, scree	F. Kolář, K. Mar	49.204509	19.735202	1783
AA169	Banikov	A	4x	x		x	x	x	x	Žilinský kraj	Vysoké Tatry, B	wet rocks	M. Lučanová, J.	49.201166	19.708372	1864
AA170	Tristarska Dolin	LA	mix			x	x	x	x	Prešovský kraj	Žiar, Tristárska	open gravely sit	F. Kolář, M. Luč	49.250216	20.205255	1380
AA171	Hranovnica	F	4x	x	x	x	x	x	x	Prešovský kraj	Hranovnica, slo	rocky eroded sl	F. Kolář, M. Luč	49.00716	20.286407	720
AA172	Primovce	F	4x			x	x	x	x	Prešovský kraj	Primovce, Prim	shady rocks	F. Kolář, M. Luč	49.015874	20.382482	605
AA173	Ganovce	F	2x			x	x	x	x	Prešovský kraj	Gánovce, traver	exposed rocks	F. Kolář, M. Luč	49.029877	20.320654	664
AA176	Velke Hincovo p	A	2x			x	x	x	x	Prešovský kraj	Vysoké Tatry, M	open moist grav	F. Kolář	49.175629	20.060517	1950
AA178	Sivy Vrch	LA	4x	x		x	x	x	x	Žilinský kraj	Zuberec, rocks	exposed rocks	F. Kolář	49.2124	19.63682	1700
AA182	Kralova Lehota	F	2x			x	x	x	x	Žilinský kraj	Kráľova Lehota,	exposed dry roc	K. Marhold, J. S	49.015286	19.811092	670
AA183	Maluzina	F	2x	x		x	x	x	x	Žilinský kraj	Malužiná, slope	eroded slope	K. Marhold, J. S	48.9845	19.7573	735
AA184	Liptovsky Hrado	F	2x			x	x	x	x	Žilinský kraj	Liptovský Hrádo	eroded slope	K. Marhold, J. S	49.03862	19.7005	624
AA208	Baba	F	2x	x	x	x	x	x	x	Prešovský kraj	Svit, Baba hill, S	forest clearing i	J. Bayerová, J.	49.043514	20.180772	844
AA225	Bukovinka - Ru	F	4x	x		x	x	x	x	Žilinský kraj	Ružomberok - J	rocks and spars	J. Bayerová	49.017836	19.283829	576
AA226	Prosiecka dolin	F	4x			x	x	x	x	Žilinský kraj	Prosiek, transec	semi-shady roc	J. Bayerová	49.160034	19.496156	656
AA227	Velka Studena	A	2x	x		x	x	x	x	Prešovský kraj	Vysoké Tatry, V	alpine scree	J. Bayerová	49.176168	20.149984	2070
AA228	Pod Polskym hr	A	2x			x	x	x	x	Prešovský kraj	Vysoké Tatry, n	rocks	J. Bayerová	49.173809	20.139095	2147
AA229	Kvacianska doli	F	4x	x	x	x	x	x	x	Žilinský kraj	Kvačany, rocks	shady rocks, dis	J. Bayerová	49.183171	19.541024	673
AA234	Pieniny	F	4x	x		x	x	x	x	Prešovský kraj	Lesnica, N facin	shady rocks in b	F. Kolář, G. Fux	49.411733	20.448837	437
AA235	Kamenica	F	4x	x		x	x	x	x	Prešovský kraj	Kamenica, rock	rocks and scree	F. Kolář, G. Fux	49.210747	20.928184	633
AA236	Kamenna Baba	F	2x	x		x	x	x	x	Prešovský kraj	Lipovce, at the l	shady limestone	F. Kolář, G. Fux	49.05893	20.930605	659
AA240	Bela	F	2x				x			Žilinský kraj	Vysoké Tatry, al	aluvial deposits	F. Kolář, A. Kno	49.12157	19.866136	853
AA241	Branisko	F	2x			x	x	x	x	Prešovský kraj	Branisko, Rajto	disturbed sites	J. Bayerová, K.	49.003297	20.861206	990
AA242	Cingov	F	4x			x	x	x	x	Košiský kraj	Spišká Nová Ve	shady north-faci	J. Bayerová, K.	48.940175	20.477409	530
AA243	Babia Gora	LA	2x					x	x	Województwo	Babia Gora, roc	rock crevices	F. Kolář, A. Kno	49.57483	19.538132	1660
AA244	Sutovo	F	2x			x	x	x	x	Žilinský kraj	Šútovo, small li	rocky and grave	F. Kolář, A. Kno	49.15197	19.085165	480
AA247	Rysy - chata	A	2x			x	x	x	x	Prešovský kraj	Rysy - just unde	rock along turist	M. Holcová	49.17465	20.08639	
AA248	Rysy - pod vrch	A	2x	x		x	x	x	x	Prešovský kraj	Rysy- just under	top of the Rysy	M. Holcová	49.179549	20.08806	2488
AA306	Chabenec	A	4x			x	x	x	x	Nizke Tatry	Nizke Tatry - Ch	vlhka sluchta	J. Bayerová	48.943407	19.493051	1852
AA307	Tri vody	A	4x						x	Nizke Tatry	Nizke Tatry - na	mokry chodnik z	J. Bayerová	48.94905	19.55399	1495
AA308	Zadna voda	F	mix			x	x			Nizke Tatry	Nizke Tatry - ce	prikopa okolo a	J. Bayerová	48.96848	19.57413	1152
AA309	Teply zlab	LA	2x			x	x			Zapadne Tatry	Zapadne Tatry -	kameny na turis	J. Bayerová	49.258917	19.706397	1149
AA310	Trencin	F	4x			x	x	x	x	Trenčianský kra	Trenčín, Trenči	rocks in castle	F. Kolář	48.894439	18.04673	
AA321	Dumbier	A	4x					x	x	Nizke Tatry	Nizke Tatry - po	uzsi, vlhci, zapa	J. Bayerová	48.93825	19.632066	1893
AA322	Krakova hola 2	LA	4x			x	x	x	x	Nizke Tatry	Nizke Tatry - m	vykaceny svah	J. Bayerová	48.991868	19.623027	1472
AA323	Smutna dolina	A	mix					x	x	Zapadne Tatry	Zapadne Tatry-	cca dva roky sta	J. Bayerová	49.202365	19.750004	1596
AA324	Choc	LA	4x					x	x	Dolný Liptov	Chočské vrchy -	forest and dwarf	M. Holcová	49.148497	19.348828	1259