ANALYSIS OF VARIABILITY OF EUROPEAN BLACK POPLAR ALONG THE DANUBE BASIN USING MOLECULAR MARKERS

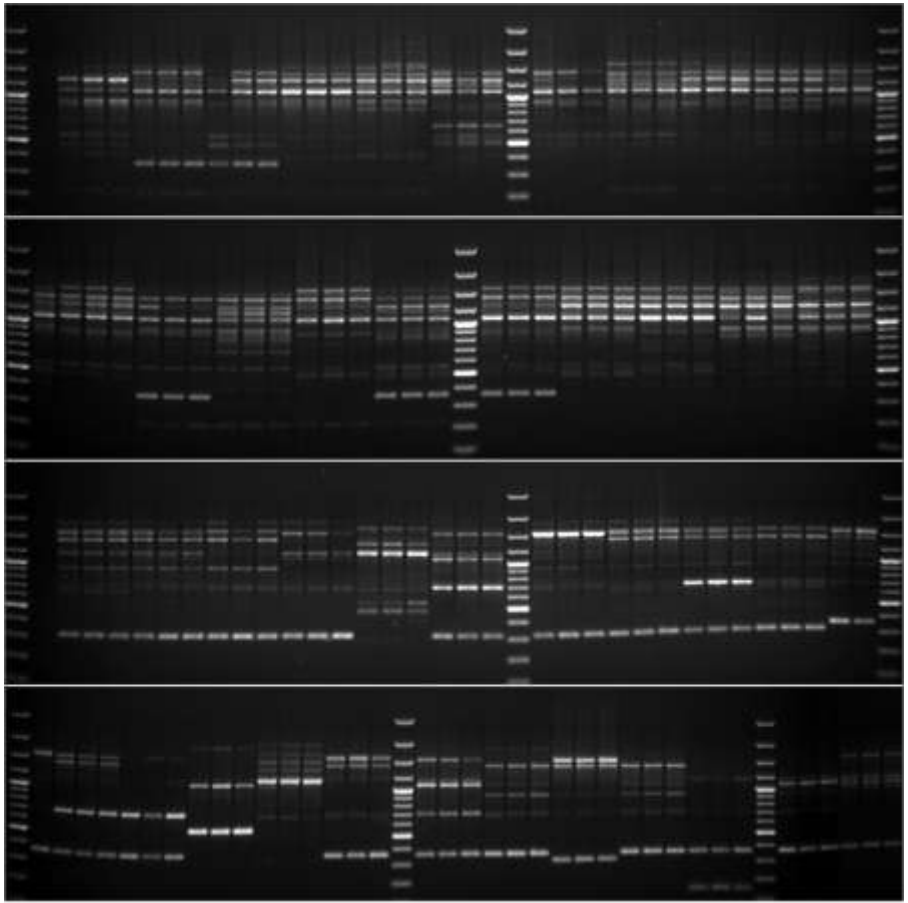
Why have we done it?

European black poplar

- Decline in number of individuals in natural populations is causing the loss of genetic variablility and further loss of ecological adaptiveness and plasticity.
- The knowledge of genetic diversity and population structure in remaining European black poplar populations is a prerequisite for the successful management of conservation

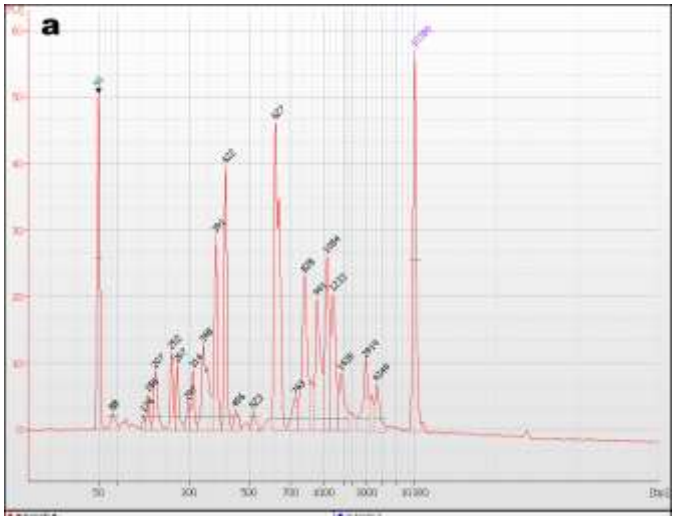
Molecular markers – a tool for exploring genetic diversity

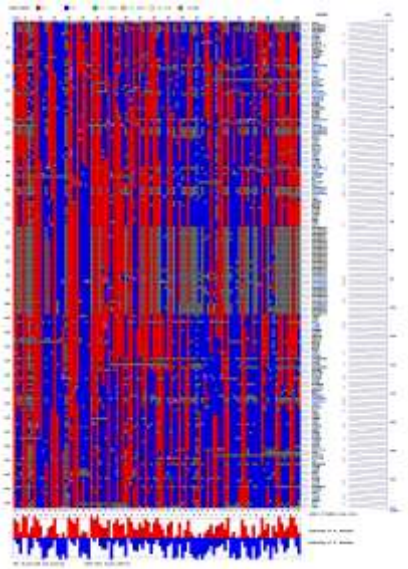
- Genetic variation is substantial and each individual of a species possesses the unique DNA sequence.

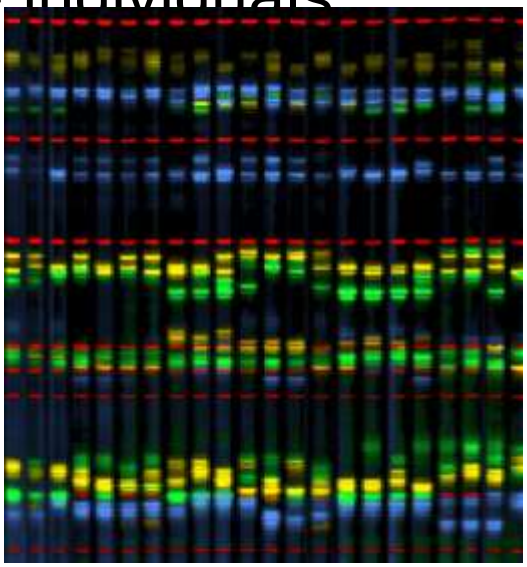


Molecular genetic markers

- Genetic markers – anything in the genome that is variable and can be used to compare individuals







- The most rapid and cost-effective measures of genetic diversity are obtained from the assay of polymorphisms using anonymous molecular genetic markers.

Molecular markers in poplars

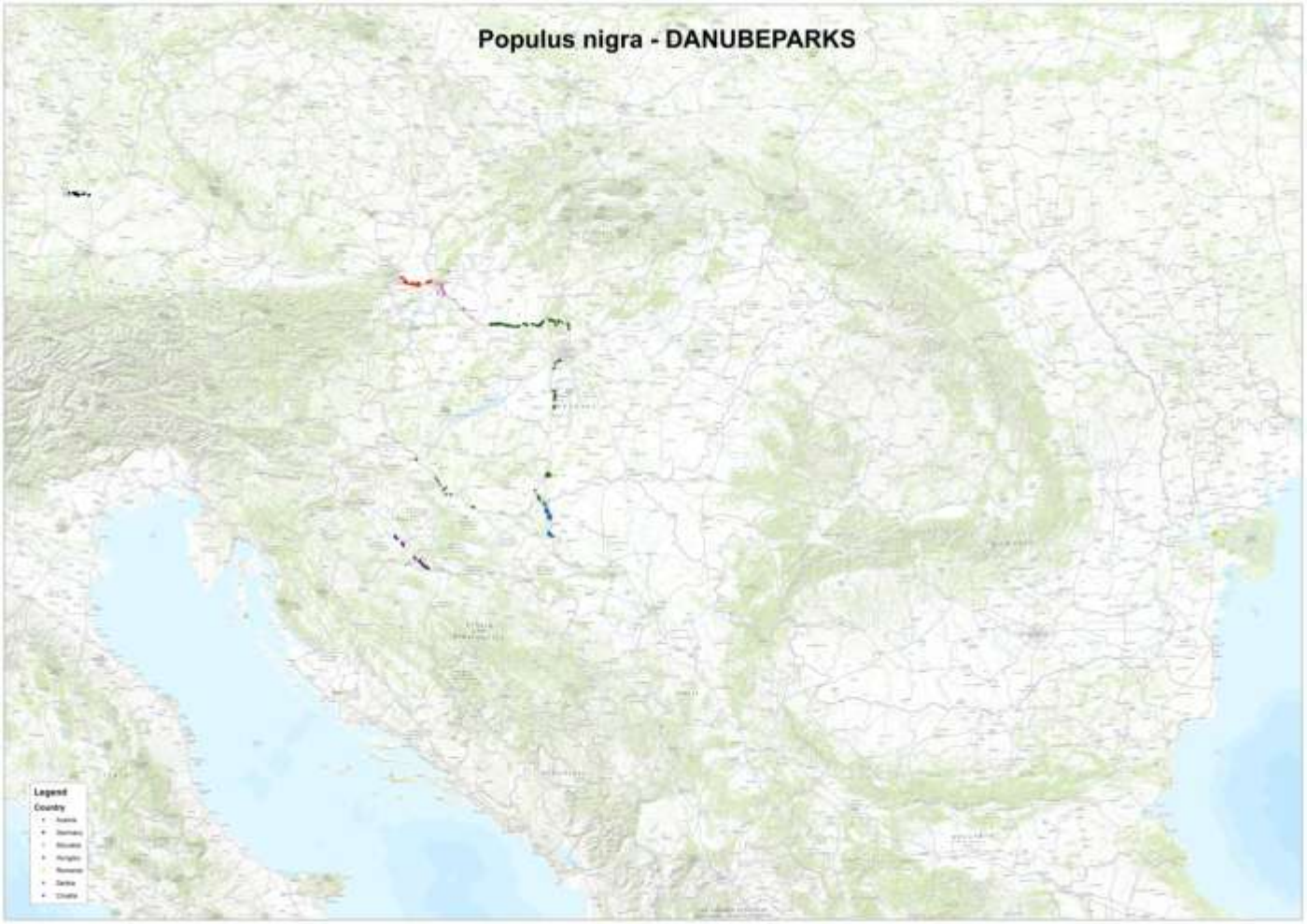
- Since 1990s various molecular markers are employed in genetic studies of the genus *Populus:*
  - Ribosomal DNA (D'Ovidio *et al.*, 1990, 1991; Faivre-Rampan *et al.*, 1992)
  - Mitochondrial DNA (Barrett *et al.,* 1993)
  - Chloroplast DNA (Smith and Sytsma, 1990)
  - RFLP (Keim *et al.*, 1989)
  - RAPD (Rajora, 1989; Castiglione *et al.,* 1993; Janssen, 1997)
  - AFLP (Arens *et al.*, 1998; Raamsdonk *et al.*, 2000)

Microsatellites in poplars

- Highly polymorphic, DNA-based and reproducible molecular markers, ideal for the assessment of intra- and inter-population studies of a species.
- In use for poplars since 2000 (van der Schoot).
- Many studies on *Populus nigra* published so far. They are dealing with population variability (Storme *et al.,* 2004; Smulders *et al.,* 2008), gene flow (Imbert & Lefèvre, 2003; Vanden Broeck *et al.,* 2004; Rathmacher *et al.,* 2010), genetic introgresion and hybridization (Fossati *et al.,* 2003; Vanden Broeck *et al.,* 2006), etc.

Collection of genotypes

- 12 populations along the Danube river (*s.l.*).

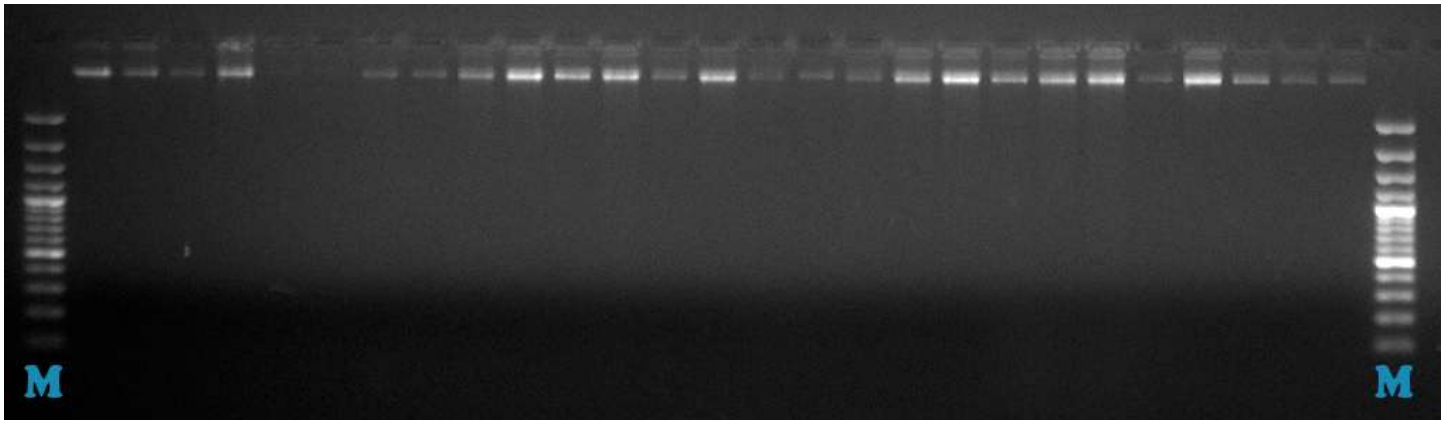


Collection of genotypes

- Thorough instructions for sampling and treatment of the plant material for the DNA extractions were given to the partners.
- Study aimed at estimating within and among population genetic variability of *Populus nigra.*
- At least 30 individuals per population.

DNA extraction

- It was possible to extract undamaged DNA from the most of the samples. However, some leaf samples have darken during drying, so the DNA was not extractable and these samples were omitted from the analysis.



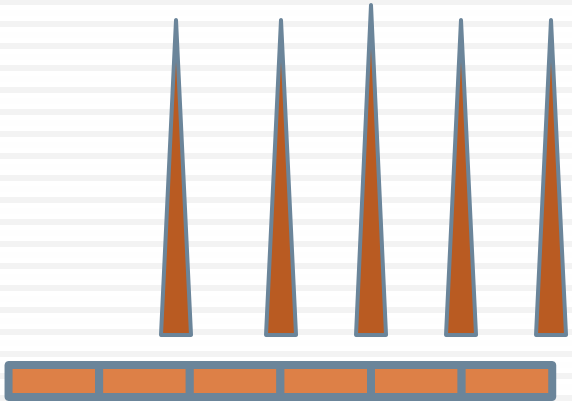
Sample size

- 12 populations from protected areas in the Danube Basin - Eight countries.
- 355 individuals genotyped in total.

Microsatellite alleles

Different number of repeats

Different positions of peaks on sequencer

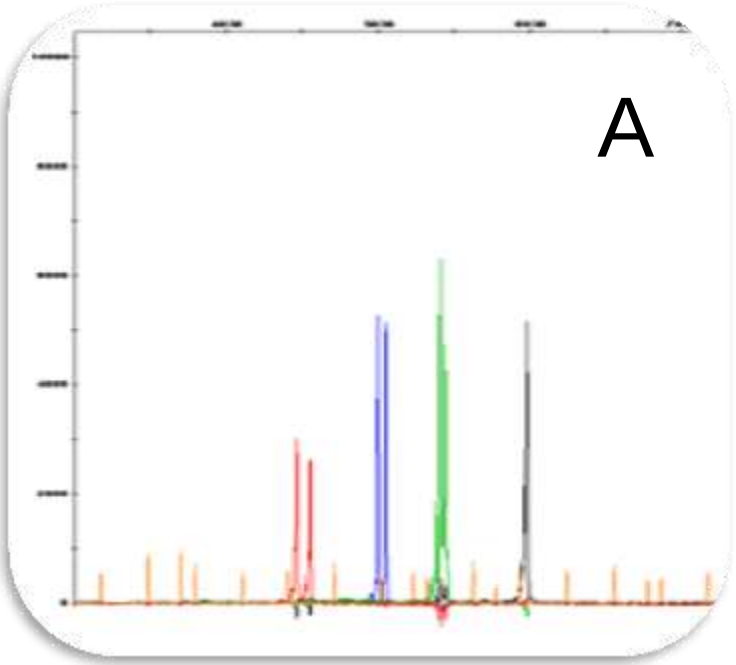


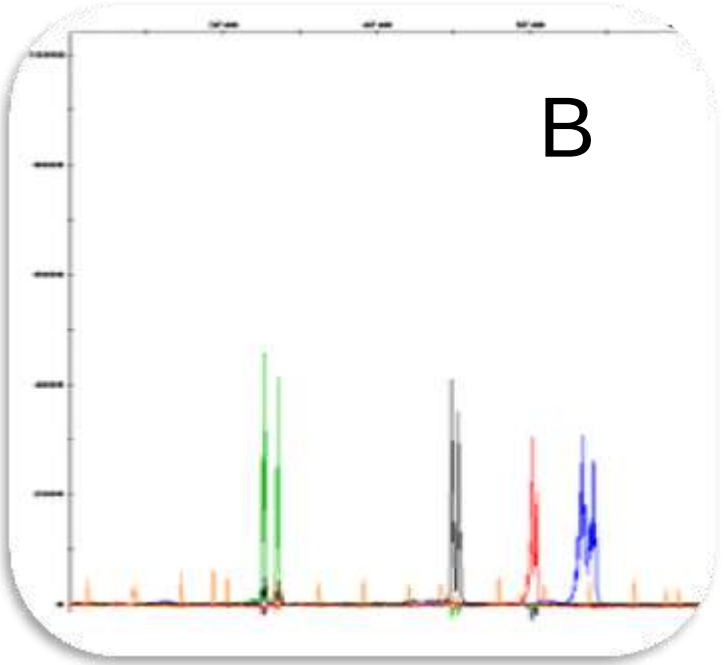
Microsatellite loci

- Eight microsatellite loci were used:

A: WPMS12, PMGC14, WPMS08, WPMS03

B: WPMS16, WPMS18, WPMS09, WPMS05





Pattern and amount of genetic variation

Allelic richness
Allelic size range
Observed heterozigosity (Ho)
Expected heterozigosity (He)

Pattern and amount of genetic variation

- Low level of clonality. In two population two individuals shared the same genotype H (DDNP) and RO (DDBR).
- Number of alleles ranged from 9-24 depending on a locus. Similar values in all populations (9.24±2.55).
- Allelic size range from 9.88-16.25 (12.88±1.83).

locus population	Mean	s.d.
D (DNI)	9.36	3.32
A (NPDA)	8.20	3.43
SK (PLADL)	9.22	2.40
H (FHNP)	8.82	2.09
H (DINP)	9.54	2.56
H (DDNP)	8.62	2.04
SRB (SNRGP)	8.81	1.51
HR (NPKR)	8.58	1.61
HR (NPLP)	7.13	1.80
BG (PNP)	8.51	2.40
BG (NPRL)	9.84	2.29
RO (DDBR)	9.15	2.60
Mean	9.71	2.44
Total	9.36	3.32

locus population	Mean	s.d.
D (DNI)	13.88	7.38
A (NPDA)	10.38	4.90
SK (PLADL)	11.50	4.24
H (FHNP)	13.13	5.79
H (DINP)	16.25	8.65
H (DDNP)	13.25	7.98
SRB (SNRGP)	13.63	7.86
HR (NPKR)	11.75	5.31
HR (NPLP)	9.88	3.83
BG (PNP)	13.50	7.48
BG (NPRL)	15.00	8.07
RO (DDBR)	12.38	6.87
Mean	12.88	5.55
s.d.	1.83	1.64
Total range	20.75	10.375

Pattern and amount of genetic variation

The mean H_O and H_E were high with similar values in all populations.

Ho

locus population	Mean	s.d.
D (DNI)	**0.66**	0.12
A (NPDA)	**0.71**	0.13
SK (PLADL)	**0.71**	0.10
H (FHNP)	**0.68**	0.18
H (DINP)	**0.74**	0.16
H (DDNP)	**0.68**	0.14
SRB (SNRGP)	**0.71**	0.16
HR (NPKR)	**0.72**	0.14
HR (NPLP)	**0.70**	0.16
BG (PNP)	**0.68**	0.09
BG (NPRL)	**0.67**	0.17
RO (DDBR)	**0.66**	0.12
Mean	**0.69**	

He

locus population	Mean	s.d.
D (DNI)	**0.80**	0.06
A (NPDA)	**0.78**	0.08
SK (PLADL)	**0.79**	0.05
H (FHNP)	**0.76**	0.05
H (DINP)	**0.82**	0.04
H (DDNP)	**0.80**	0.06
SRB (SNRGP)	**0.81**	0.06
HR (NPKR)	**0.79**	0.05
HR (NPLP)	**0.79**	0.06
BG (PNP)	**0.79**	0.07
BG (NPRL)	**0.80**	0.09
RO (DDBR)	**0.79**	0.06
Mean	**0.79**	0.06
Total	**0.81**	0.05

Pattern and amount of genetic variation

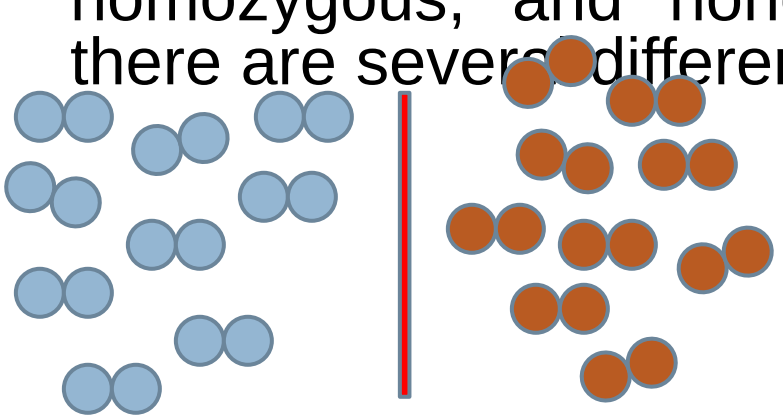
			locus							
		para-meter	WPMS 03	WPMS 05	WPMS 08	WPMS 09	WPMS 12	PMGC 14	WPMS 16	WPMS 18
S	our study	n	**24**	**19**	**22**	**23**	**12**	**11**	**9**	**14**
D	12 populations the Danube River, 8 countries	H_O	**0.60**	**0.73**	**0.49**	**0.78**	**0.64**	**0.81**	**0.76**	**0.73**
No	264 samples	H_E	**0.84**	**0.83**	**0.80**	**0.83**	**0.69**	**0.81**	**0.77**	**0.77**
S	van der Schoot *et al*., 2000	n	15	13	12	11	10	/	/	/
D	EUFORGEN Core Collection	H_O	/	/	/	/	/	/	/	/
No	23 samples	H_E	/	/	/	/	/	/	/	/
S	Smulders *et al*., 2001	n	/	/	/	/	/	/	7	7
D	EUFORGEN Core Collection genotypes,West and Middle Europe	H_O	/	/	/	/	/	/	/	/
No	23 samples	H_E	/	/	/	/	/	/	/	/
S	Smulders *et al*., 2008	n	/	/	/	25	15	12	10	17
D	17 populations, 11 river valleys, 7 catchments (Danube, Ebro, Elbe, Po, Rhine, Rhone, and Usk)	H_O	/	/	/	0.82	0.76	0.74	0.59	0.73
No	1069 samples	H_E	/	/	/	0.83	0.72	0.74	0.65	0.74
S	Imbert and Lefèvre, 2003	n	/	/	/	13	15	13	11	12
D	22 populations the Drôme River, the Netherlands	H_O	/	/	/	0.79	0.78	0.71	0.82	0.85
No	652 samples	H_E	/	/	/	0.79	0.8	0.72	0.65	0.73
S	Rathmacher *et al*., 2010	n	/	10	/	17	/	9	/	8
D	Hesse, central Germany, at the Eder River	H_O	/	0.73	/	0.67	/	0.78	/	0.62
No	~290 samples	H_E	/	0.76	/	0.74	/	0.76	/	0.58

Comparisons to previous studies

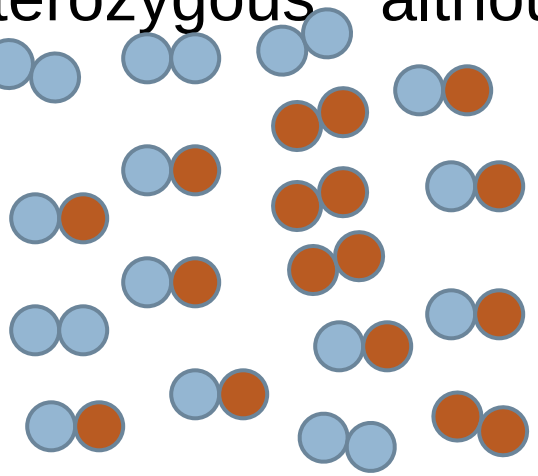
- Consistency with other studies considering sample size, number of populations and different set of loci used.
- Allelic richness detected in our study was high, and did not vary much among populations.
- Allelic richness was slightly lower compared to studies where other catchments were analyzed, and higher compared to studies with lower sample size.
- The results of H_O and H_E are high and similar among analyzed populations, and were close to values reported by other studies.
- Genetic variability is preserved in all populations.

Wright's F-statistics

- If populations are divided with limited gene flow, and are genetically different, in a way that allele frequencies of particular genetic loci are different, there are fewer H_E.
- Extreme examples are populations each with only one, but different allele. All individuals are thus homozygous, and none is heterozygous, although there are several different alleles.



Observed



Expected

Wright's F-statistics

- The main effect that a population subdivision has on genetic diversity is the reduction in Ho compared to He.
- The extent of this reduction can be used to quantify the level of genetic differentiation between populations.
- Formalization in Wright's F-statistics.
- Very useful tool in elucidating the pattern and extent of genetic variation residing within and among natural population of different species.
- F_{ST}=0 no differentiation between populations. Populations have the same allele frequencies.
- F_{ST}=1complete differentiation –populations are fixed for different alleles.

F_{ST} between pairs of populations

- The majority of pairwise comparisons showed a **small** but statistically significant genetic differentiation (F_{ST}<0.05).
- Two comparisons showed a **moderate** and statistically significant genetic differentiation (0.05<F_{ST}<0.15): D (DNI) vs. RO (DDBR), H (FHNP) vs. RO (DDBR).

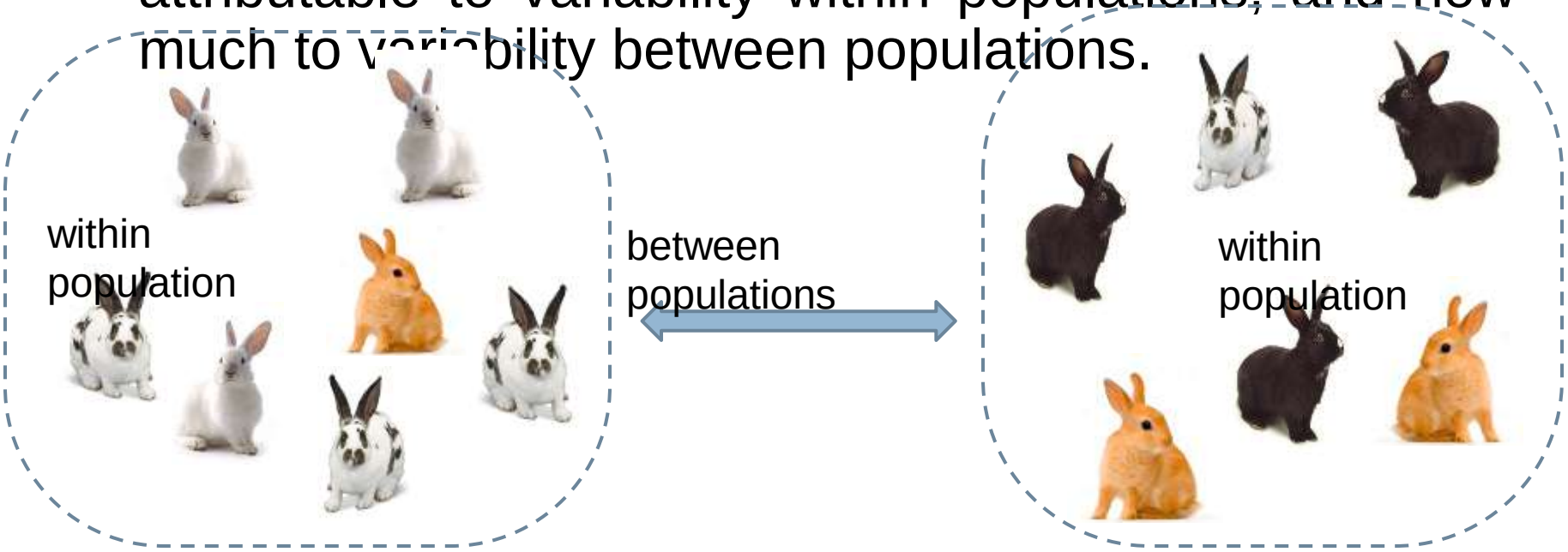
population	D (DNI)	A (NPDA)	SK (PLADL)	H (FHNP)	H (DINP)	H (DDNP)	SRB (SNRGP)	HR (NPKR)	HR (NPLP)	BG (PNP)	BG (NPRL)
A (NPDA)	0,0339 ***										
SK (PLADL)	0,0305 ***	-0.0018									
H (FHNP)	0,0340 ***	0,0225 ***	0,0216 ***								
H (DINP)	0,0295 ***	0,0111 *	0,0123 **	0.0073							
H (DDNP)	0,0309 ***	0.0075	0,0125 *	0,0234 ***	0.0098						
SRB (SNRGP)	0,0390 ***	0,0242 ***	0,0256 ***	0,0385 ***	0,0167 ***	0.0072					
HR (NPKR)	0,0349 ***	0,0213 ***	0,0243 ***	0,0159 **	0,0123 *	0.0003	0.0065				
HR (NPLP)	0,0377 ***	0,0349 ***	0,0284 ***	0,0374 ***	0,0154 **	0,0245 ***	0,0218 ***	0,0295 ***			
BG (PNP)	0,0486 ***	0,0363 ***	0,0453 ***	0,0388 ***	0,0295 ***	0,0366 ***	0,0395 ***	0,0371 ***	0,0254 ***		
BG (NPRL)	0,0366 ***	0,0264 ***	0,0323 ***	0,0363 ***	0,0212 ***	0,0342 ***	0,0307 ***	0,0487 ***	0,0214 ***	0,0120 *	
RO (DDBR)	0,0694 ***	0,0263 ***	0,0334 ***	0,0526 ***	0,0279 ***	0,0328 ***	0,0324 ***	0,0553 ***	0,0371 ***	0,0373 ***	0,0148 *

F-statistics between pairs of populations

- The highest F_{ST} recorded between two most distant populations.
- Populations in the middle of the Danube flow show milder levels of differentiation.
- Furthermore, populations H (DDNP), SRB (SNRGP), and HR (NPKR), which are geographically very close, show no significant differentiation, and very low F_{ST} levels.
- As expected, populations at the lower flow of the Danube (populations BG (PNP) and BG (NPRL) from Bulgaria) are also very similar.
- Interestingly, population HR (NPLP) from Croatia, (Danube tributary), evenly and significantly differs from all other populations.
- Pattern of F_{ST} values indicates isolation by distance.
- Small values are the result of high gene flow between populations.

AMOVA

- Analysis of Molecular Variance (AMOVA) is a methodology that partitions genetic variability of a population to different sources, i.e. different levels of organization.
- It tells how much of the genetic variability is attributable to variability within populations, and how much to variability between populations.



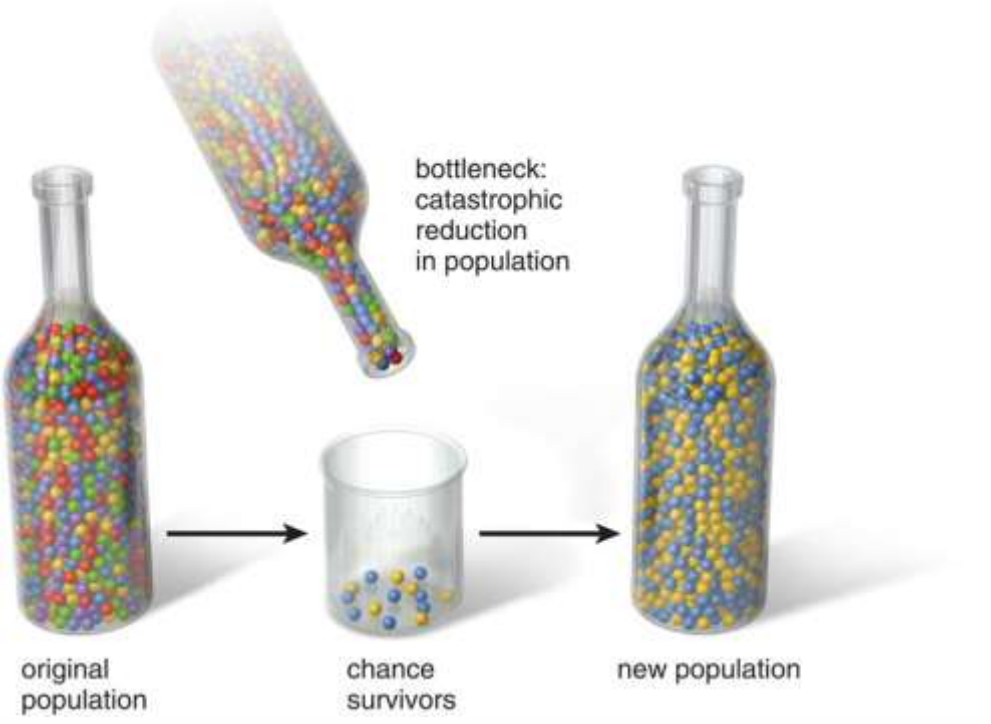
AMOVA

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	11	89.873	0.08734	2.83
Within populations	698	2096.658	3.00381	97.17
Total	709	2186.531	3.09114	
Fixation index	F_{ST}: 0.02825			

- The greatest amount of genetic variance occurred within populations.
- Smulders *et al*. (2008) reported a similar proportion of genetic variation within the catchment.

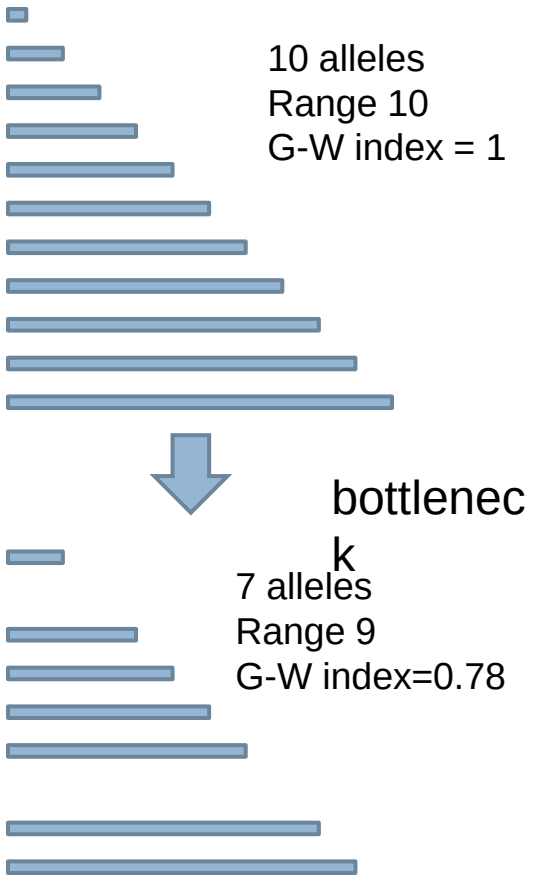
Reduction of population size

Bottleneck effect



Testing reduction of population size

- Populations that have experienced a recent reduction of their size exhibit a correlative reduction of the allele numbers as well as heterozygosities at polymorphic loci.
- However, the allelic diversity is actually reduced faster than the heterozygosity, i.e. H_O is larger than H_E from the observed allele number where the locus is at mutation-drift equilibrium.
- Similarly, Allelic richens is lost faster than allelic size range.



Population bottleneck

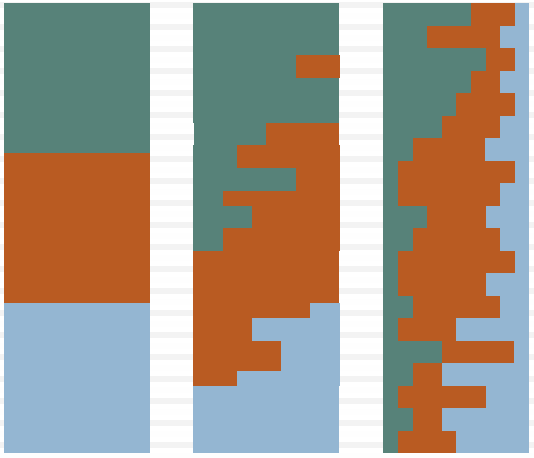
- The majority of populations of the European black poplar in the Danube Basin have not experienced a recent reduction of population size.
- The exception is population HR (NPLP) (heterozigosity excess, p< 0.05)

	D (DNI)	A NPDA)	SK (PLADL)	H (FHNP)	H (DINP)	H (DDNP)	SRB (SNRGP)	HR (NPKR)	HR (NPLP)	BG (PNP)	BG (NPRL)	RO (DDBR)
P	0.844	0.629	0.875	0.990	0.844	0.629	0.629	0.963	**0.0137***	0.726	0.875	0.844

Metapopulation structure

How many different genetic clusters/groups are there in our sample?

Which population draws genetic constitution from which cluster?



STRUCTURE software

Popul. 1

Popul. 2

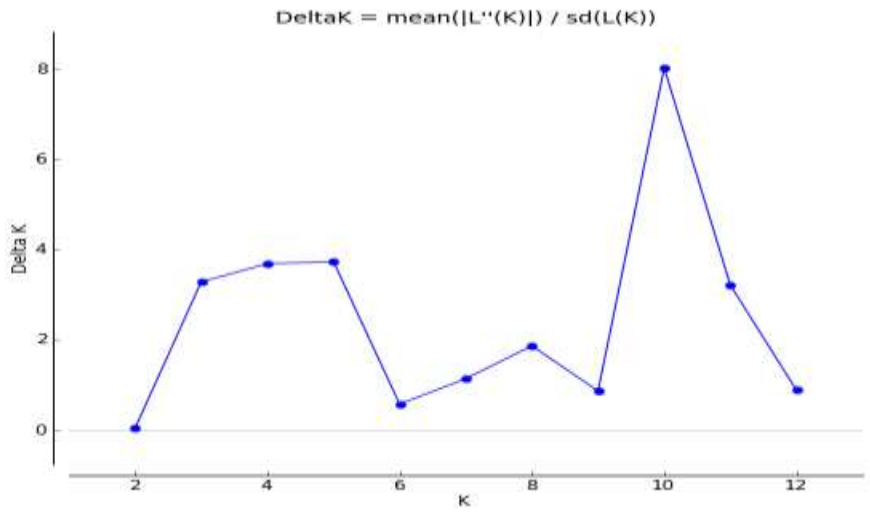
Popul. 3

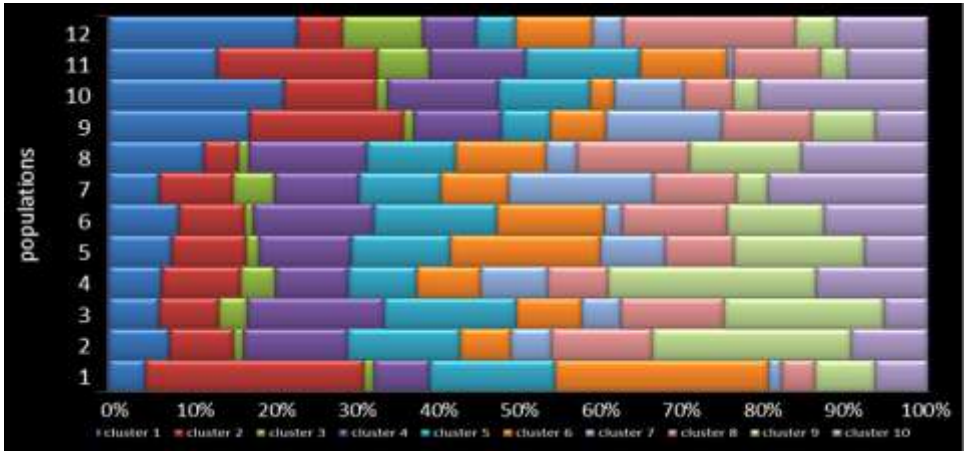
- Based on allele frequencies in each population, each individual can be assigned (with certain probability) to a genetic clusters where it originated. Each population can be assigned to clusters that it is made of.
- It can show that a specific population has the alleles that originate only from one of the assumed groups, or it can be said that its genetic constitution is made of e.g. 70% of the first group, 20% of the second group, and 10% of the third group.
- If populations are differentiated and isolated some genetic groups/clusters would be seen preferentially in some populations and not in the others.
- If populations are not differentiated all will have similar proportion of different genetic clusters.

STRUCTURE software

- In the STRUCTURE software one can assume any number of supposed groups, and the program will assign populations to each of the groups.
- Than it can be said which number of different genetic clusters is the most probable.

STRUCTURE





- The most probable number of different genetic clusters in our sample is 10.
- All populations of the European black poplar in the Danube Basin recruited plants from all clusters
- Geographically closer populations tended to draw evenly from the same clusters.

THANK YOU FOR YOUR ATTENTION!

Are there alleles of *P. deltoides* in our samples?

- Finding true diagnostic alleles for each species would require extensive sampling from both species natural ranges, as alleles common in one species may be rare, but still present, in the other.
- If introgression is expected, than it is best to analyze population of invasive species in the near by of the threatened population.
- While it is possible to differentiate between pure parent species and hybrid plants of the first generation with singular marker, further generations of hybrids and backcrosses produce individuals that cannot be as easily assessed. Individuals will arise that carry *P. nigra* pattern at one marker but contain many *P. deltoides* genes in the rest of their genome.

Are there alleles of *P. deltoides* in our samples

- WPMS09, WPMS18, PMGC14 that we used can discriminate
- WPMS09 - 234 allele specific for *P. deltoides* found only in three individuals.
- WPMS16 - 220 allele specific for *P. deltoides is* present in our populations, but our sample size and number of populations is quite larger than in studies that defined this alleles as diagnostic and it is possible that this allele exists in *P. nigra* too.
- PMGC14 - alleles 193 and 199 specific for *P. deltoides*, we have some 199, but none 193.
- There may be only few individuals, but they are no F1 generation hybrids!
- Promising markers SSR: win3, PTR6, PMGC2163

(photo: PE Vojvodinašume, Serbia)