

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,200

Open access books available

128,000

International authors and editors

150M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



Capillary Electrophoresis as Useful Tool in Analysis of *Fagus sylvatica* L. Population Genetic Dynamics

Justyna A. Nowakowska and
Malgorzata K. Sulkowska

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/59197>

1. Introduction

1.1. What kind of tool is Capillary Electrophoresis (CE)

The Capillary Electrophoresis (CE) is one of the method widely used in modern molecular genetics, applied for fast and efficient DNA fragment separation in the sieving polymer in the electric field [1]. The use of this method has increased dramatically over the last fifteen years, due to high precision of small nucleic acid separation (even to the single nucleotide level) of the available material analysed in the field of analytical chemistry, physical chemistry, biochemistry, and biotechnology. Essentially, the CE technique has been used for genome sequencing projects, e.g. Human Genome Project [2] or many others assignments (for animals, plants, bacteria and fungi), published in NCBI database (www.ncbi.org).

Progress in the CE area in recent years relied more on increasing the number of samples analyzed at the same time, as well as the development of new gels (which allow multiple separation in the same capillary filling), and "chemistry" (a mixture of buffers, substrates, and the so-called polymerase enhancers) for analyzing a sequence of one fragment nearly 1000 base-pairs. Nowadays, the CE technique applied for sequencing analysis is overcome by pyrosequencing method relying on the luminometric detection of pyrophosphate that is released during primer-directed DNA polymerase catalyzed nucleotide incorporation [3].

The CE technique is commonly performed in automated sequencers, i.e. CEQ™ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA) composed by two main components: hardware (apparatus) and the CEQ System software. The named model is equipped by 8 capillary system which ensure 8 sample analysis in the same time.

Another apparatus used for CE, especially recommended in case of DNA sequencing, is e.g. 3500 Genetic Analyzer (Life Technologies) which includes the capillary electrophoresis instrument with the workstation and the 3500 Data Collection Software for instrument control, data collection, quality control and autoanalysis of sample files for basecalling and fragment sizing. Auto-analysis can also be performed in this model thanks to the GeneMapper® or GeneMapper® ID-X Softwares.

Both sequencers cited above are equipped with 8 capillaries, but 16-or even more capillary system are available [4]. All type of automated sequencers require appropriate chemistry used for the given goal (genotyping or sequencing), comprising polymer (separation gel), chemical buffers and washing solutions.

The current chapter describes application of DNA based genotyping of *Fagus sylvatica* population by capillary electrophoresis performed in CEQ™ 8000 (Beckman Coulter) sequencer.

1.2. Advantages of Capillary Electrophoresis (CE) and weak points

The CE method is widely used in modern molecular biology science in assessment of gene or allele presence (genotyping) in a given DNA sample, as well as in determining of the DNA nucleotide composition (sequencing) of the studied gene.

In general, the CE technique has several advantages:

- Good tool to support conservation and management of forest trees genetic resources, in order to:
 - Characterize the genetic structure of the forest tree stands
 - Assess the initial gene pool of the population
 - Detect the selection processes and to maintain high level of the natural diversity of forest stands
 - Reflect the history of the stand in relation to the post-glacial migration refugia in Europe and in the world (phylogeny study)
- Provide genetic characteristics of different forest tree species reproductive material in:
 - Mother and progeny stands gene flow analysis
 - Seed orchards and progeny plantation mating system
 - Assignment of populations selected for preservation in gene banks or in situ or ex situ measures
- Solving of problems from seed stand management point of view:
 - Clonal/pedigree identification/selection processes
 - Pollen contamination especially important in management of artificial tree stands like forest seed plantations

- Patterns of gene flow and mating system in natural and artificial stands
- Tracing of forest tree species with DNA markers as a support in the combat with illegal logging:
 - Thanks to the DNA profiles established on a basis of minimum 4 microsatellite nuclear DNA loci, and at least one cytoplasmic (mitochondrial or chloroplast) DNA marker
 - Strong proof to support the decision taken by several District Law Courts, as far as the identification of wood samples is proved with a high probability (approximately 98–99%)
 - Consistent with the assumptions of the European Parliament Directive on Timber Regulation (EUTR), which came into effect in 2013 to stop the circulation of illegally logged wood in the European Union.

Major weak points of the CE method are derived from:

- Can be avoid by appropriate programme i.e. GeneMapper® or GeneMapper® ID-X Softwares which analyse, process and report the data basing on bank of dataset
- Addition of at least 10% formamide improves the denaturing capacity of the sieving matrix and reduces the compressions
- Can be avoid by appropriate machine running respecting all routine maintenance and installation rules according to the manufacturer.

1.3. Methodical problems

Two main errors may occur during the genotyping procedure performed with CE technique. First of all, some alleles are not identified as peaks in chromatograms (so called null alleles), influencing the general allele distribution in the population. Appropriate software, e.g. GenALEx [5] or Micro-Checker (<http://www.microchecker.hull.ac.uk>) may help in calculation of the probability of the null allele occurrence in the studied group of trees.

Secondly, the homoplasy phenomenon may also occur during genotyping. This term is applied to the DNA fragments of the same size (in base-pairs, i.e. 324 bp) deriving from different microsatellite loci e.g. mcf-5 and mcf-11 in the case of *Fagus sylvatica* species. Such errors are avoided by different fluorochrome labelling of the primers during PCR prior to the CE detection in automated sequencer.

In sequencing data obtained from CE technique, the sample contamination by the DNA molecules from the other species, or errors in trimming of the 5' and 3' ends of the coding regions. The trimming errors are estimated to be very low, corresponding to the rate of 0.07, 0.06, 0.05, 0.03, 0.01 with medium default value of 5%.

Good quality of DNA sequences and fragments are obtained thanks to cautious application of the user guide advice provided by the sequencer manufacturer, e.g. CEQ™ 8000 Genetic Analysis System User's Guide (www.beckmancoulter.com/wsrportal/wsr/index.htm) or Applied Biosystems 3500/3500xL Genetic Analyzer User Guide: http://tools.lifetechnologies.com/content/sfs/manuals/cms_069856.pdf.

All those troubleshoots occurring within genotyping and sequencing procedure can be overcome by use of the control samples (individual of reference, with known DNA structure) analysed in the same run. The other possibility of solving problems related with CE analysis is repetition of the experiment from the beginning with new reagents and new PCR reactions (see Methods below).

2. Methods

2.1. The idea of Capillary Electrophoresis (CE) electrophoresis

The assessment of the genetic structure (both for genotyping and sequencing) rely on five basic steps of treatment applied to the plant material collected from the field.

Applied procedure of DNA analysis in CE consists on the following steps:

- Total genomic DNA isolation from the plant tissues (homogenization and extraction of nucleic acid molecules)
- Specific loci amplification via PCR technique using labeled primers
- Separation of the amplified fragments via CE in automated sequencer
- Allele scoring and / or sequence data processing
- Computing of the obtained data base with utilization of proper software.

Concerning the DNA isolation, there are several techniques of extraction, based on the lysis of cell walls to facilitate proper isolation of nucleic acids from the plant, fungal and animal tissues [6]. Good performance of extraction is guaranteed by kits ready for DNA isolation, e.g. DNeasy Plant Mini Kit (QIAGEN®) or NucleoSpin® Plant II (Machery-Nagel).

The first step in obtaining DNA molecules from plant tissues, is initial mechanical homogenization or grinding the material in liquid nitrogen. Liquid nitrogen damages mechanically cell walls, which allows easier access to DNA molecules, at the same time as their stability at low temperature -176°C , is maintained. Efficiency of DNA extraction is then analyzed by spectrophotometry (Fig. 1) or by electrophoresis in agarose gel, followed by staining with ethidium bromide (50 mg ml^{-1}). During sequencing, template impurity due to presence anions coming from unpurified PCR product mix or the presence of another DNA sequence can significantly influence the separation performance of the capillary. At the end, obtained DNA molecules can be stored in stabilizing buffer (pH 7.0) for a long term (even years) at -75°C .

The polymerase chain reaction (PCR) proceeds in DNA multiplication in a thermocycler programmed for multiple (average of 30 to 40) cycles (Fig. 2). Most of DNA techniques are based on amplification of the genomic DNA fragments thanks to the thermostable enzyme *Taq* polymerase deriving from thermophilic bacteria *Thermus aquaticus*.

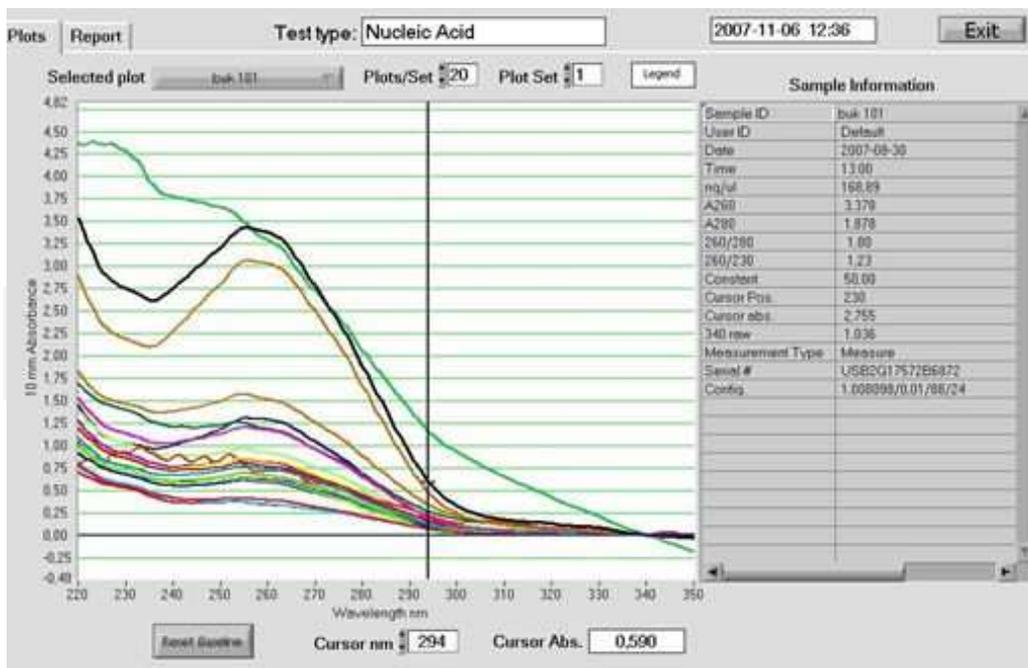


Figure 1. Spectrophotometrical assessment of the DNA extracts from the beech leaf samples in the spectrophotometer NanoDrop® ND-1000 (TK-Biotech, USA). All the DNA extracted from the plant tissues are examined for their purity and evaluation of the DNA concentration before the amplification of microsatellite loci in PCR reaction.

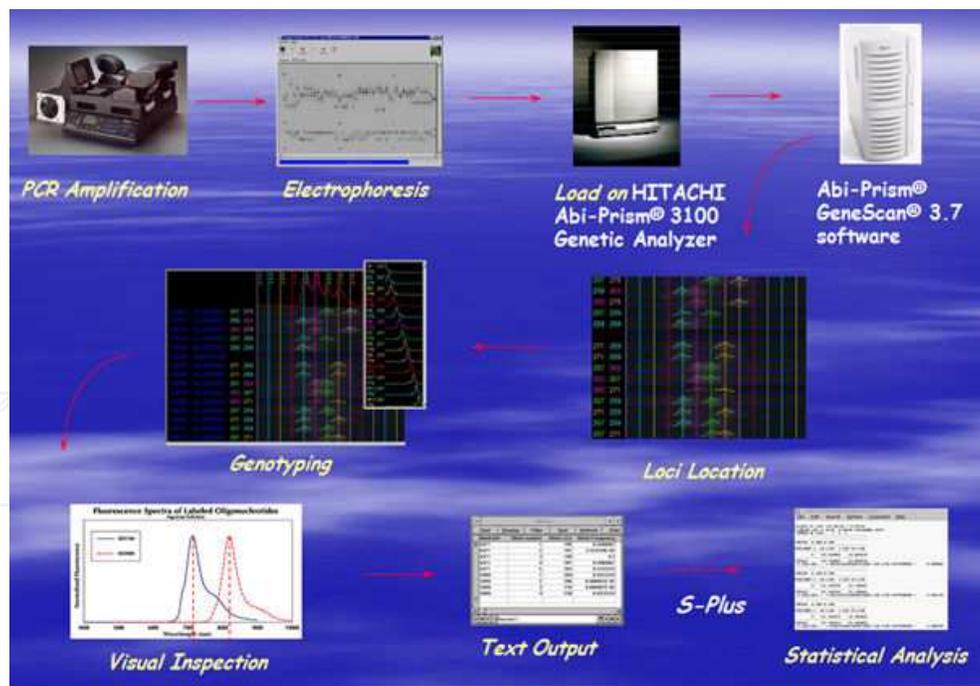


Figure 2. General scheme of the microsatellite loci analysis using the CE in automated sequencer. After PCR amplification followed by gel-electrophoresis, the DNA samples are loaded to the sequencer (e.g. HITACHI Abi-Prism 3100 Genetic Analyzer) for the CE running. After this, the separated DNA fragments are examined in software for precise loci location and genotyping comprising the visual inspection of the peaks (alleles), numeric data collection (list of alleles) and final statistical analysis.

The PCR reaction involves the DNA-matrix in the following reaction mixture: labeled oligonucleotide primers (from 10 to 24 base pairs length), four types of free nucleotides (dATP, dGTP, dCTP and dTTP), magnesium ions (Mg^{2+}), reaction buffer and *Taq* polymerase. The first stage of amplification takes about a few minutes and leads to double-stranded DNA template denaturation at $94^{\circ}C$. Then the stage of annealing at $32-42^{\circ}C$ comes, lasting from 30 sec. to a few minutes, when the formation of complementary DNA strand to the matrix (at $72^{\circ}C$) occurs. The temperature and duration of each stage depends on many factors, mainly on the G/C and A/T content in primers and the size of duplicated DNA fragments. Efficiency and precision of PCR is very high, and theoretically it allows reproducing the output-template DNA molecules present in the extract up to 10^9 copies [6].

Then, the PCR products, labeled with different fluorochromes (e.g. WellRed D2, D3 and D4 for CEQTM 8000 model; and FAM, JOE, ROX and TAMRA for Abi-Prism sequencer) are subjected to the CE run in automated genetic analyzer (Fig. 2 and 3). Several labeling strategies have been developed. With some exceptions, the DNA primer labeling is generally done at the 5' end of one primer. The DNA markers separated in the gel during electrophoresis can be detected by fluorescence of specific nucleotides in labeled primers.

During electrophoresis in polyacrylamide gel, negatively charged DNA molecules migrate in the direction to a voltage positive electrode of the gel at a speed proportional to their size and molecular weight. After completion of electrophoresis on capillary, the DNA fragments are detected by laser thanks to the appropriate fluorochrome labeling of the primers in PCR reaction (Fig. 3). The separated DNA fragments are then analyzed using computer software.

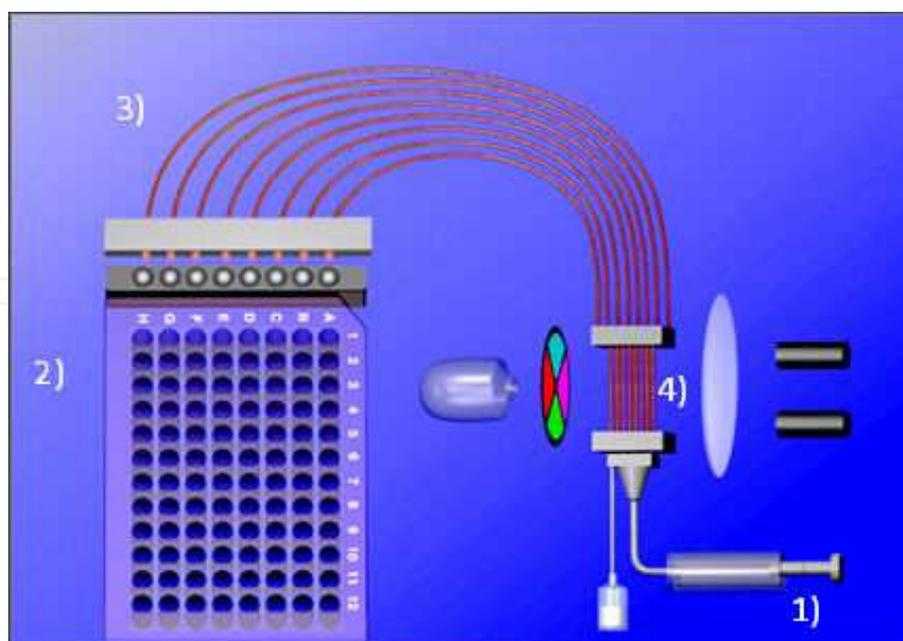


Figure 3. General scheme of the CE principle in the CEQTM 8000 (Beckman Coulter) sequencer. 1) Gel cartridge – containing the polymer used for separation, 2) 96-well plate with samples, 3) Eight capillaries in which the samples are automatically loaded and the CE is performed, 4) Optical system of windows for the peak detection.

2.2. Genetic differentiation evaluation process

The most crucial step in allele scoring is the exact allele-size determination. For this, a general scheme can be applied, helping to avoid the erroneous allele listing for examined population (Fig. 4). The most tricky allele assignment occurs into 2-base-pair repetition in the SSR fragment, especially when two adjacent alleles for the same locus differ only by 2-bp length (Fig. 4C). Good quality laboratory manipulation and broad experience of the scientist easily overcome such a discrepancy and depict the heterozygous loci in an individual. Otherwise, the false result would lead to the excess of homozygotes in the examined group of trees.

The nuclear microsatellite DNA sequences have so far been considered as the most informative markers, and have been used in the genetic diversity studies of many organisms. They are formed by short repeats of 1-6 base-pairs, so called Short Sequence Repeats (SSR or SSRs), and constitute the most powerful tool in modern population genetics and forensic studies.

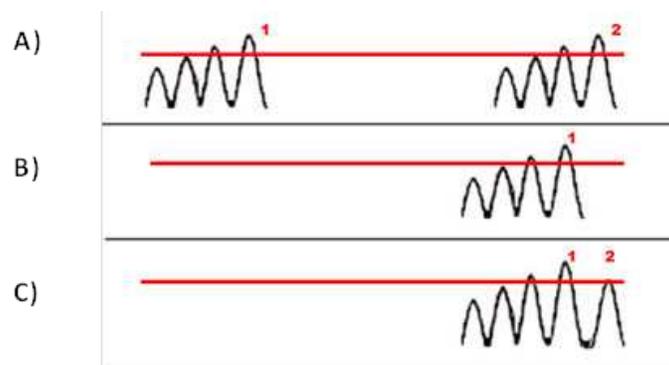


Figure 4. Microsatellite allele scoring, examined according to the appropriate threshold (red line) for peaks obtained from the CE performed in automated sequencer. The biggest peak (i.e. sharing the biggest peak-area) represent the correct fragment size following the rule: A) Well separated Heterozygote, B) Homozygote, C) 2 base-pair separated heterozygote.

The advantages of microsatellite sequences are numerous: they are uniformly distributed over entire genome, are present in high proportion in forest tree species, form discrete loci and co-dominant alleles. The observed mutation rates for the SSR markers vary from 10^{-3} to 10^{-6} . The SSR fragments obtained after the CE technique applied to the European beech populations in Poland illustrate the precision of the detection of different alleles in four nuclear SSR loci investigated (Fig. 5)

The sequencing methodology based on CE relies on nucleotide extension product growing from 5' to 3' direction by forming a phosphodiester bridge between the 3'-hydroxyl group at the growing end of the primer and the 5'phosphate group of the incoming deoxynucleotide. The principle of DNA replication follows Sanger dideoxy sequencing procedure [7]. The DNA sequence is copied with high fidelity at each base on the DNA template, as far as the DNA polymerase incorporates only one complementary nucleotide. The resulting nucleotide alignment is registered in chromatogramme and four letter code corresponding to the studied gene fragment (Fig. 6)

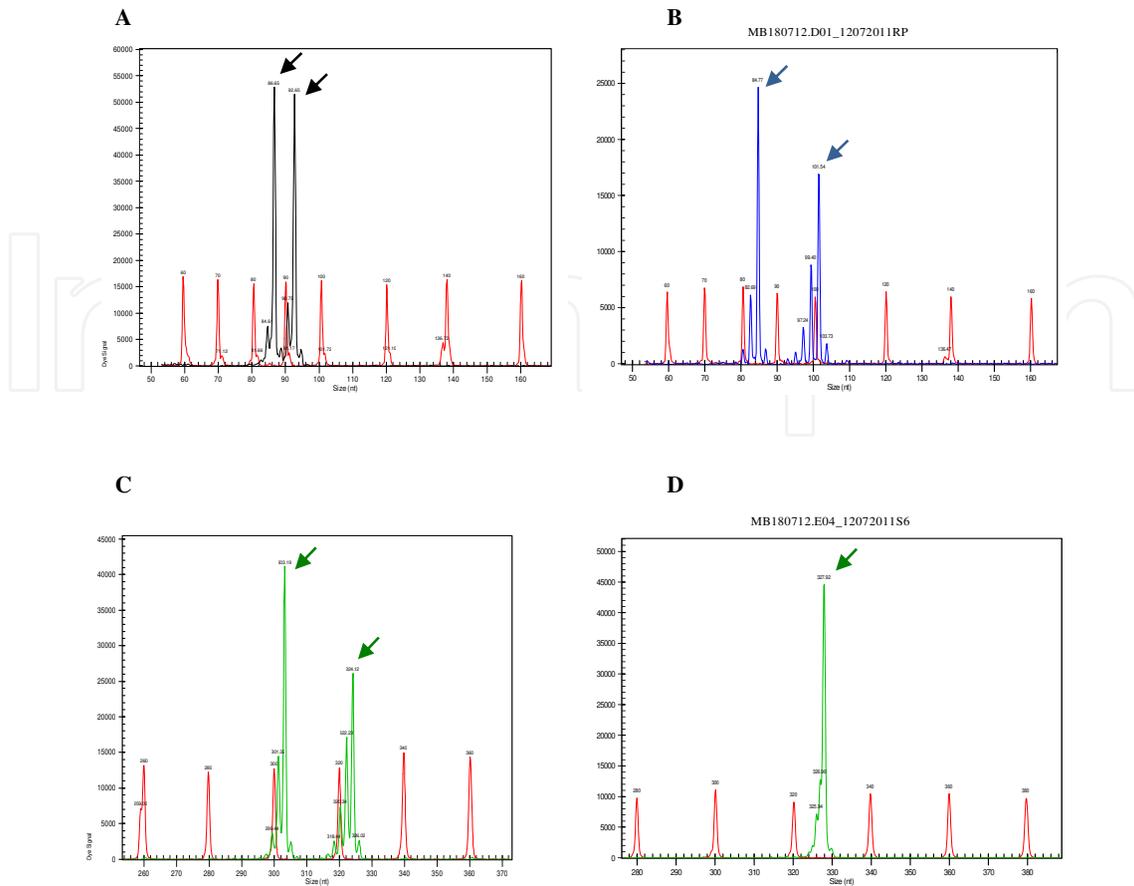


Figure 5. Example of microsatellite nuclear DNA analysis in *Fagus sylvatica* genotyping in population from Bieszczadzki National Park in Poland: two alleles 87 and 93 base-pairs in locus FS1-03 (A), two alleles 85 and 101 bp in locus FS1-25 (B), two alleles 303 and 324 bp in locus mcf-5 (C), one allele 328 bp in locus mcf-11 (D). Obtained from DNA capillary electrophoresis after Beckman Coulter® software CEQ™ 8000 Genetic Analysis System v 9.0 (Fullerton, USA).

2.2.1. Analysed parameters to describe population genetic variation and differentiation

The genetic diversity is defined as the probability of occurrence of the identical genotype among randomly chosen trees in a forest stand. The picture of electrophoretic separation of DNA fragments is converted to numerical data, using software such as CEQ System software (Beckman Coulter sequencer) or GeneMapper® or GeneMapper® ID-X Softwares (Abi-Prism sequencer).

The allele size obtained from sequencing data can be checked with the use of S-Plus software version 3.4 release 1 for SPARC (Statistical Sciences, Math Soft Inc., Seattle, WA).

In fact, the population genetic variation and differentiation are based on the heterozygosity parameter (1) calculated as values from zero (no heterozygosity) up to nearly 1.0 (when we observe a large number of almost equally frequent alleles). Instead of average number of alleles per locus more precise measure of effective number of alleles per locus (n_e) – Crow & Kimura [8] can be used (2).

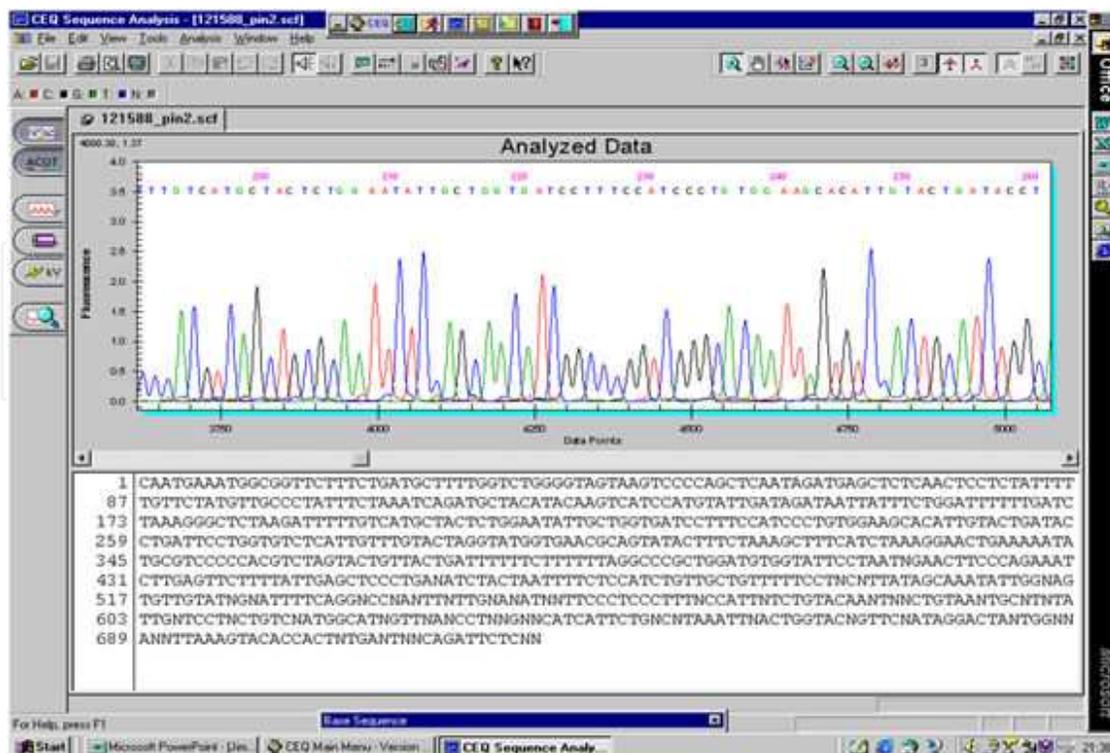


Figure 6. General scheme of capillary electrophoresis data processing for gene sequencing in automated sequencer CEQ™ 8000 (Beckman Coulter). All single nucleotide acids are visible thanks to the different fluorochrome labeling, Adenine (A) in red, Thymine (T) in blue, Guanine (G) in green and Cytosine (C) in black.

$$H_e = 1 - \sum_{i=1}^n p_i^2 \quad (1)$$

$$n_e = \frac{1}{\sum_{i=1}^n p_i^2} \quad (2)$$

p_i – frequency of n allele occurrence in population

Heterozygosity is often one of the most important parameter when we describe the genetic data. Using this measure we explain the general trend in the structure of analysed populations – even is their history and future genetic structure is concerned. Low values of heterozygosity is influenced by small population size and processes of genetic drift e.g. bottlenecks effect. A lot of heterozygotes in population signify high genetic variability. When we compare the level of the observed and expected heterozygosity in balanced populations concerning random and open mating system (i.e. under Hardy-Weinberg equilibrium) and the observed heterozygosity level is higher than the expected one, we can presume the gene flow via alien pollen outside of population. If the observed heterozygosity is lower than expected one we can assume that some inbreeding processes may occur in the population.

The interpopulational variation is described very often as G_{ST} [9] used as equivalent of F_{ST} in G_{ST} statistics [10, 11], and it enables to assess the distance for each population from other populations (3).

$$G_{ST} = \frac{H_T - H_S}{H_T} \quad (3)$$

H_T -interpopulation heterozygosity

H_S -intrapopulation heterozygosity

The F_{ST} parameter called as fixation index is the measure of proportion of the total genetic variance within subpopulations in relation to the total genetic variance (4). The values of this parameter can range from 0 to 1. High F_{ST} implies a considerable degree of differentiation among populations.

F_{IS} (inbreeding coefficient) is the proportion of the variance in the subpopulation. High F_{IS} implies a considerable degree of inbreeding (5). Values can range from -1 (outbred) to +1 (inbred).

$$F_{ST} = (H_T - H_S) / H_T \quad (4)$$

$$F_{IS} = (H_S - H_I) / H_S \quad (5)$$

H_T -total heterozygosity for a population

H_S -heterozygosity within a subpopulation

H_I -heterozygosity of an individual

2.2.2. Software

One of the oldest programs enabling computing of DNA markers analysis data is BIOSYS-2. This program was elaborated to help biochemical population geneticists to describe the analysis of electrophoretically detectable allelic variation. It can be utilized to study allele frequencies and genetic variability measures, to test deviation of genotype frequencies from Hardy-Weinberg, expectations, to calculate F-statistics, to perform heterogeneity of chi-square analysis, to calculate a variety of similarity and distance coefficients, and finally to construct dendrograms using among others cluster analysis procedures. The program, documentation, and test data are available from the authors [12].

The statistical analysis of the alleles consists in calculation of the genetic parameters estimated by Nei [9, 13], i.e. expected heterozygosity (H_E), observed heterozygosity (H_O), observed number of alleles (A_O) and population differentiation parameters (H_S , H_T , F_{ST}). Those parameters may be calculated with the programs like GENEPOP software 3.2a [14], GenALEX [5] or

ARLEQUIN (<http://lgb.unige.ch/arlequin/>). The spatial correlations may be evaluated with SPAGeDi v.1.2 [15] and the genetic distances estimated according to Nei [13]. Another one interesting software enabling the analysis of DNA data markers is POPGENE [16]. The current version of POPGENE is designed specifically for the analysis of co-dominant and dominant markers using haploid and diploid data. The software performs most types of data analysis encountered in population genetics and related fields. It can be used to compute summary statistics, including: allele frequency: estimates of gene frequencies at each locus from raw data, effective number of alleles per locus, percentage of all polymorphic loci, observed and expected homozygosity, Shannon Index, gene diversity Nei's [9], F-Statistics, gene flow from the estimates of G_{ST} or F_{ST} and many others parameters.

All those programs represent good tool for population genetics analysis and simulations, including: Hardy-Weinberg Equilibrium (HWE), multiple allele and loci inheritance, natural selection, genetic drift, migration, mutation and inbreeding.

3. Results presentation Genetic variation characteristics of *Fagus sylvatica* L. as an example of utilization of capillary electrophoresis method on the basis of nuclear and chloroplast DNA markers

3.1. Object of the study

European Beech (*Fagus sylvatica* L.) is one of the most important forest tree species in Poland. Beech forests cover about 5.6 % of forest area [17]. The most typical beech forest tree populations are formed at the lower forest belt in Carpathians and Sudety Mountains on the South and at the moraine landscape of Pomeranian Lake District of the North of the country. In Poland beech attains its north-eastern limit of natural range [18, 19]. Varying environmental conditions have resulted in a great number of ecotypes and populations which are characterized by various ecological requirements [20, 21]. The growth of beech stands outside the natural beech limit indicates that species possess potentially much wider range [22, 23].

Present genetic structure of beech populations in Poland was formed by many different factors, not only environmental and genetic ones but also anthropogenic [24, 25, 26, 27]. Recent investigations of beech variation in Poland performed with isoenzyme study [28, 29, 30, 31], showed high genetic diversity, similar to other neighboring European populations, slight decrease of average number of alleles per locus and lower level of differentiation towards the North of the natural range limit, which generally confirm the migration paths after glaciations. The present paper describes the genetic structure within one generation, i.e. mother and progeny beech stands in Poland assessed with chloroplast and nuclear DNA markers.

There were investigated six beech populations representing natural beech range in Poland. The populations were classified according to phytosociological characteristics to the following plant associations: *Galio-odorati-Fagetum* (Gryfino and Kwidzyn), *Dentario glandulosae-Fagetum* (Bieszczadzki National Park), *Luzulo-luzuloides-Fagetum* (Suchedniów, Tomaszów),

Dentario enneaphyllidis-Fagetum (Zdroje) - Fig. 7. The genetic structure of these populations was analysed.

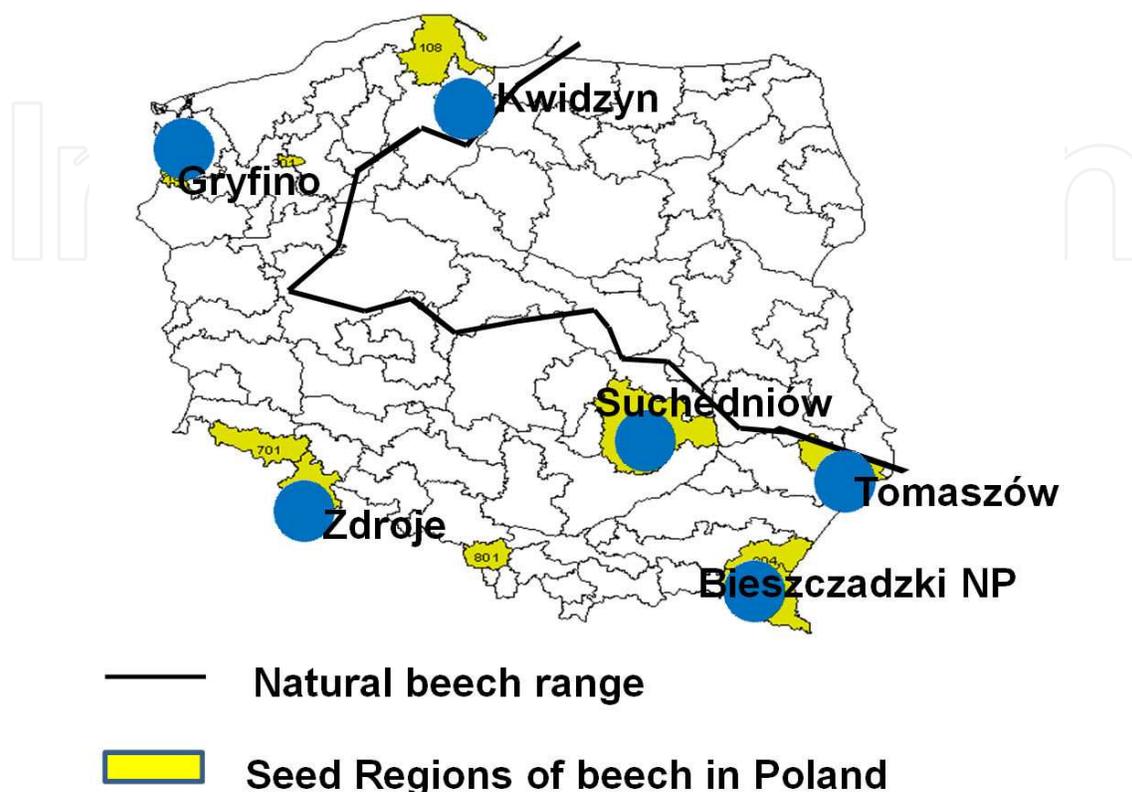


Figure 7. Localisation of investigated Polish European beech populations.

3.2. Methods

The genetic variation and differentiation of mother stands and their open-pollinated progeny were characterised on the basis of nuclear microsatellite markers, i.e. FS1-03, FS1-25, FCM5, mcf5, mcf11 [32, 33], as well as chloroplast DNA markers: ccmp4, ccmp7, ccmp10, according to reference [34, 35]. Thirty individuals per one generation (mother, progeny stands) in every provenance were investigated.

The extraction of total DNA from the leaves was performed using Qiagen DNeasy™ Plant Minikit according to the manufacturer instruction (Qiagen). The quality and purity of DNA were analyzed on 1% agarose gel electrophoresis and via absorption in 230, 260 and 280 nm in NanoDrop® spectrophotometer (Wilmington, USA). DNA samples were analyzed with DNA capillary electrophoresis in Beckman Coulter® sequencer, and analyzed using the software CEQ™8000 Genetic Analysis System v 9.0 (Fullerton, USA).

Parameters of genetic diversity (H_S and H_T) and differentiation (G_{ST}) were counted and compared between mother and progeny generation according to Nei [36, 11] in PopGene 1.32 software [16].

3.3. Results

3.3.1. Quality and quantity of the analyzed DNA

The very high quality and purity of DNA were assessed on the basis of the ratio of absorbance at 260 and 280 nm. A ratio about ~1.8 was typical for most of the samples (Fig. 2), and proved a high purity of the extracted DNA. A ratio ~2.0 is generally accepted as RNA-free. If the ratio is lower in either case, it indicates the presence of contaminants. The good quality of isolated DNA from samples was confirmed also by measurement of absorbance at 230 nm wavelength. The quantity of the genomic DNA samples balanced between 35 up to 160 ng.µl⁻¹ and was fully appreciated to perform next steps of the DNA analysis procedure.

3.3.2. Genetic structure based on nuclear DNA markers

As far as nuclear microsatellite markers of Polish beech stands were concerned, different fragment-variants of haplotypes, i.e. from 73 to 348 base-pair size were distinguished. In investigated trees both heterozygous and homozygous alleles were found (Fig. 5).

The FS1-03 locus was the most polymorphic, i.e. exhibited 27 allele variants: 73 up to 154 bp DNA fragments, while the smallest number of variants was observed in case of mcf-11 locus – 11 alleles variants with DNA fragment length from 315 up to 348 bp. Usually the observed number of alleles per locus was higher for mother stands comparing to progeny stands except practically only population Tomaszów. For instance, mother trees presented mean observed number of alleles per locus 9.4, comparatively to 7.4 found in progeny from Tomaszów population (Fig. 8).

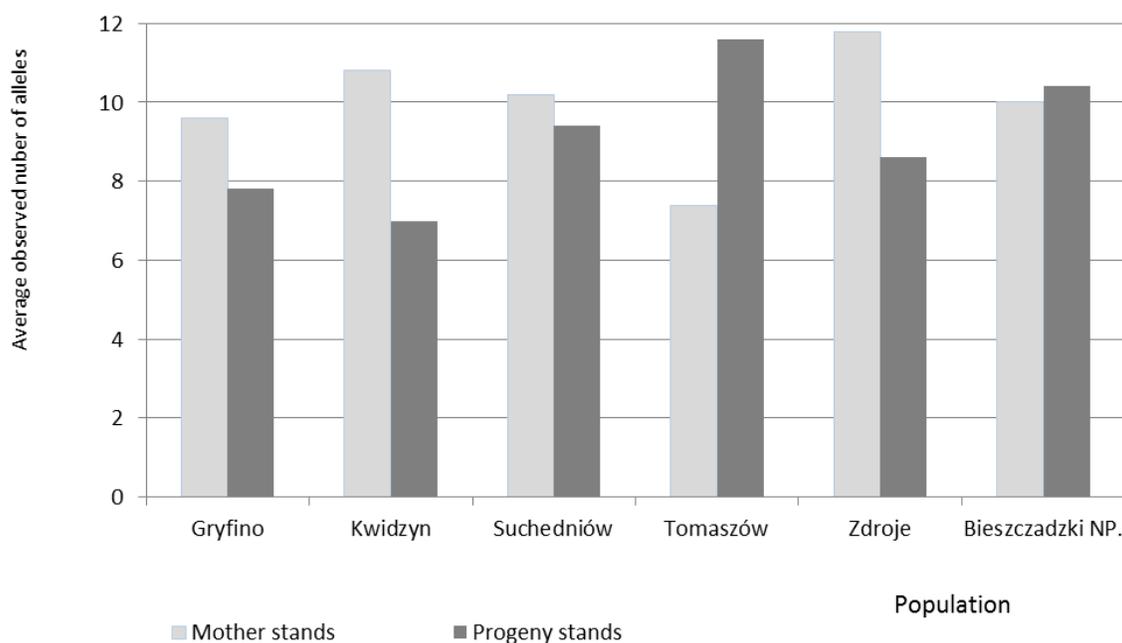


Figure 8. Average number of alleles in populations of investigated mother and progeny *Fagus sylvatica* L. stands.

Mean gene diversity among and within all studied mother stands (Tab. 1) were more or less at the same level ($H_T=0.8291$, $H_S=0.7693$, respectively) than mean genetic diversity in the progeny stands ($H_T=0.8257$, $H_S=0.7672$), as well as the gene diversity level of all mother and progeny stands were almost at the same level: $F_{ST}=0.0587$ and $F_{ST}=0.0576$ respectively (Tab. 1). The differentiation among studied Polish *Fagus sylvatica* population stands and their progeny can be explain that most genetic diversity resides within the stands.

Differentiation based on Nei [13] genetic distances was independent to geographical location of populations (Fig. 9).

Locus	Mother Stands			Progeny Stands		
	H_T	H_S	F_{ST}	H_T	H_S	F_{ST}
FS1-03	0.7165	0.6994	0.0404	0.7730	0.7284	0.0433
FS1-25	0.8855	0.8371	0.0489	0.8861	0.8354	0.0496
FCM-5	0.9072	0.8370	0.0590	0.8862	0.8275	0.0591
mcf-5	0.9127	0.8370	0.0683	0.8984	0.8309	0.0708
mcf-11	0.7236	0.6357	0.0766	0.6847	0.6140	0.0652
Mean	0.8291	0.7693	0.0587	0.8257	0.7672	0.0576
St. Deviation	0.1001	0.0955	-	0.0939	0.0966	-

Table 1. Genetic variation parameters of nuclear DNA markers microsatellite loci analyzed in *Fagus sylvatica* L. stands

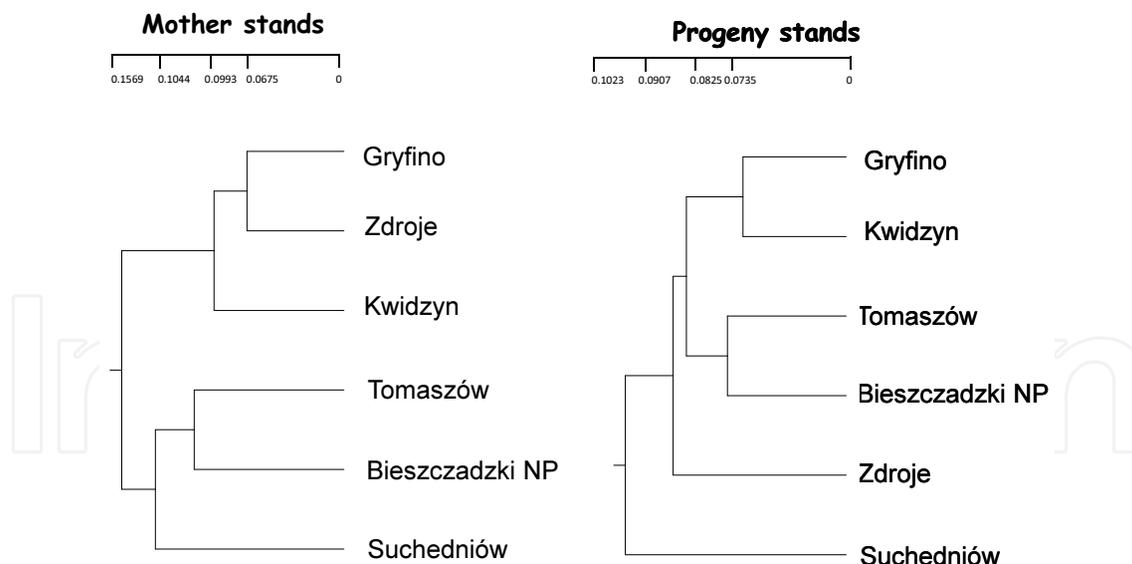


Figure 9. Genetic distances on the base of five SSR nuclear microsatellite markers

3.3.3. Genetic structure based on chloroplast DNA markers

As far as ccmp4 and ccmp10 microsatellite markers of Polish beech stands were concerned, different fragment-variants of haplotypes, i.e. from 116 to 152 base-pair size were distinguish-

ed. The *ccmp7* locus was the most polymorphic, i.e. exhibited 8 allele variants: 144, 145, 147, 148, 149, 150, 151 and 152 bp. Nevertheless, some populations e.g. Bieszczadzki NP showed less polymorphism in the locus *ccmp7*, with only one 147 bp variant present, similarly to the previous study performed on Polish beech [38]. Another two loci, *ccmp4* and *ccmp10*, shared the same range of 5 allele-size variants of 116, 117, 118, 119 and 120 bp. Generally, mother beech trees had more variable allele, than the progeny trees of the same provenance. For instance, mother trees presented alleles from 117 to 119 bp, comparatively to only one allele of 119 bp found in progeny from Bieszczadzki NP population. Mean gene diversity among and within all studied mother stands (Tab. 2) were slightly higher ($H_T=0.4916$, $H_S=0.3606$ respectively) than mean genetic diversity in the progeny stands ($H_T=0.4600$, $H_S=0.3375$). The gene diversity level of all mother and progeny stands were almost at the same level: $G_{ST}=0.2666$ and $G_{ST}=0.2663$ respectively (Tab. 2). The overall haplotypic differentiation among studied Polish *Fagus sylvatica* populations was quite low ($G_{ST}=0.016$), which means that most genetic diversity resides within the stands.

Differentiation based on Nei [13] genetic distances was independent to geographical location of populations (Fig. 10).

Locus	Mother Stands			Progeny Stands		
	H_T	H_S	G_{ST}	H_T	H_S	G_{ST}
<i>ccmp4</i>	0.5053	0.3058	0.3947	0.3957	0.3042	0.2313
<i>ccmp7</i>	0.3228	0.2667	0.1738	0.4650	0.2767	0.4050
<i>ccmp10</i>	0.6468	0.5092	0.2128	0.5193	0.4317	0.1688
Mean	0.4916	0.3606	0.2666	0.4600	0.3375	0.2663
St. Deviation	0.0264	0.0169	-	0.0038	0.0068	-

Table 2. Genetic variation parameters of chloroplast DNA markers microsatellite loci analyzed in *Fagus sylvatica* L. stands

3.4. Discussion

The DNA markers constitute powerful tool when gene variability of forest trees is assessed. The complicity of genome of these organisms makes impossible in most cases to obtain information about particular gene structure or variability. Nowadays, microsatellite markers analysed via CE technology overcome this difficulty, enabling the genetic variation study of the given organism, but only at the level of non-coding DNA regions.

The great advantage of the CE method is that all steps of allele fragment genotyping and DNA sequencing analysis are performed automatically in one integrated system. The laboratory tasks are then concentrated on high purity DNA molecules isolated from plant tissue and on appropriate PCR amplification procedures.

The data of CE application to study the SSR markers in European beech stands in Poland revealed: high genetic diversity of beech, similar like in other neighboring European popula-

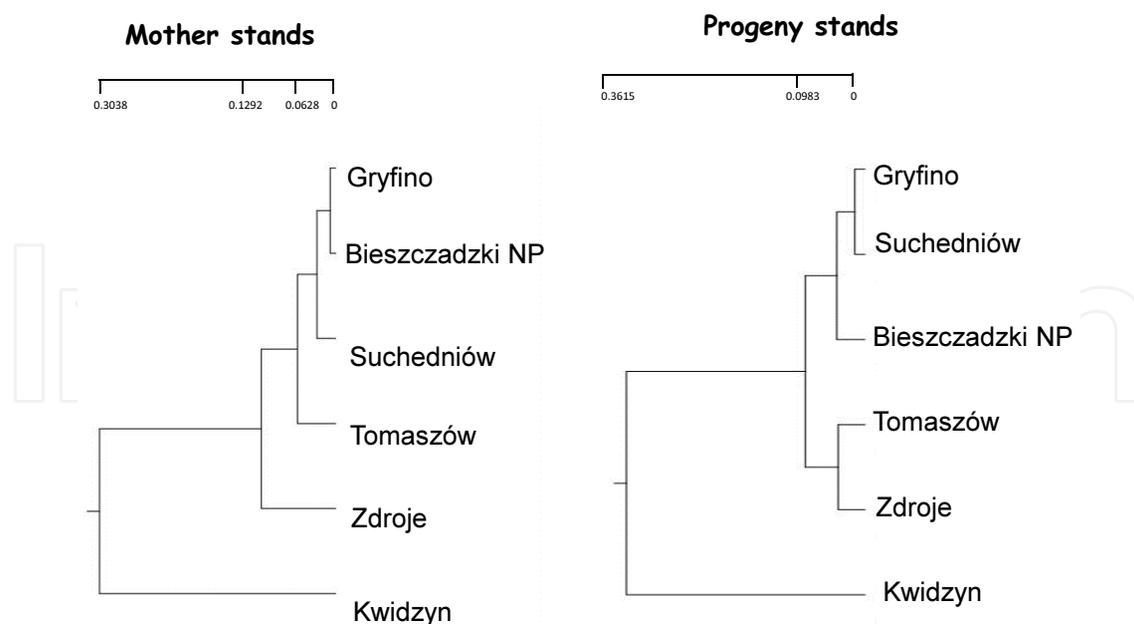


Figure 10. Genetic distances on the base of three chloroplast ccmp markers

tions, slight decrease of average number of alleles per locus and level of differentiation towards the North of the natural range limit, which generally confirm the migration paths after glaciations but it is not the basis to distinguish geographic regions. More cpDNA variation in the chloroplast ccmp4, ccmp7 and ccmp10 loci ($G_{ST}=0.810$) was reported for other 400 European beech populations [37]. Nevertheless, some populations e.g. Bieszczadzki NP showed less polymorphism in the locus ccmp7, with only one 147 bp variant present, similarly to the previous study performed on Polish beech [38].

In most cases, higher level of genetic variation was found within the investigated beech populations, predispose them to higher genetic tolerance against harmful environmental factors-reference [39].

4. Conclusions

In recent years the development and dissemination of new molecular methods based on CE method has remarkably increased. The use of automatic machines (sequencers) enables the investigation not only of DNA and RNA structure, but also gene bank creation for the forest tree species in order to achieve the appropriate forest stand management level. The CE gives very good resolution in separation of DNA molecules at the level of single nucleotide base-pair. Moreover it helps to overcome the problem of homoplasy in population genetics and offer cheap and fast results in sequencing of a small amount of samples. Concurrent development of software and computerisation offered the possibility of complete automation of the sample processing data, including the transfer of data and results.

Application of CE electrophoresis method is useful tool in forest trees genetic diversity assessment, despite of their limits. It is possible to use them on wild scale, especially in:

- Genetic characteristics of different forest reproductive material of natural and artificial stands e.g.
 - Mother stands and progeny stands gene flow
 - Seed orchards gene flow
 - Gene banks representation of populations assessment
- Solving of particular problems in case of:
 - Clonal/pedigree identification/selection process
 - Pollen contamination
 - Mating system
 - Patterns of gene flow
- Good tools to support conservation and management of forest genetic resources e.g. to support following activities:
 - Selection and protection of ecotypes
 - To asses initial gene pool for needs of effective gene conservation measures
 - To present selection processes of forest stands to maintain rich natural diversity
 - Prevention of illegal logging procedure thanks to the polymorphic SSR markers assessed with CE technique.

Acknowledgements

Special thanks are addressed to Jolanta Bieniek and Malgorzata Borys for their laboratory assistance in CE technique.

Author details

Justyna A. Nowakowska and Malgorzata K. Sulkowska*

*Address all correspondence to: m.sulkowska@ibles.waw.pl

Forest Research Institute, Sekocin Stary, Braci Lesnej, Raszyn, Poland

References

- [1] Camilleri P. Capillary Electrophoresis: Theory and Practice. R. Weinberger Ed., CRC Press; 1997.
- [2] Hawkins T.L., McKernan K.J., Jacotot L.B., MacKenzie B., Richardson P.M., Lander E.S. DNA sequencing – a magnetic attraction to high-throughput genomics. *Science* 1997; 276: 1887-1889.
- [3] Fakruddin M., Mannan K.S., Chowdhury A., Mazumdar R.M., Hossain M.N., Islam S., Chowdhury M.A. Nucleic acid amplification: Alternative methods of polymerase chain reaction. *Journal of Pharmacy and Bioallied Sciences* 2013; 5: 245-252.
- [4] Dolník V. DNA sequencing by capillary electrophoresis (review). *Journal of Biochemical and Biophysical Methods* 1999; 41: 103-119.
- [5] Peakall, R., Smouse P. E. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 2012; 28: 2537–2539.
- [6] Sambrook J., Russell D.W. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA. 2001; Vol. 1 and 2.
- [7] Sanger F., Nicklen S., Coulson A.R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* 1977; 74(12): 5463-5467.
- [8] Crow, J.F., Kimura, M. Introduction to Population Genetics Theory. Harper and Row, New-York; 1970.
- [9] Nei, M. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA*. 1973; 70(12, Pt. 1), 3321-3323.
- [10] Wright, S. Evolution and the Genetics of Populations, Vol. II. The Theory of Gene Frequencies. University of Chicago Press, Chicago; 1969.
- [11] Wright, S. Evolution and the Genetics of Populations, Vol. IV. Variability Within and Among Natural Populations. University of Chicago Press, Chicago; 1978.
- [12] Swofford, D.L., Selander, R.B. (1997). Biosys-2. User Manual. University of Illinois; 1997. Available from <http://fortinbras.us/biosys2/> (accessed 5 February 2009).
- [13] Nei, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 1978; 89: 583-590.
- [14] Raymond, M., Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenism. *Journal of Heredity* 1995; 86: 248-249.
- [15] Hardy O.J., Vekemans X. Spagedi: a versatile computer program to analyse spatial-genetic structure at the individual or population levels. *Molecular Ecology Notes* 2003; 4: 618-620.

- [16] Yeh, F.C., Boyle, T.B.J. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* 1997; 129: 157.
- [17] Forests in Poland 2013. Warszawa, Centrum informacyjne Lasów Państwowych 2013; <http://www.lasy.gov.pl/publikacje/in-english/forests-in-poland-2013/view> (access 11 March 2014).
- [18] Szafer W., Pawłowski W. Szata roślinna Polski. [The look–description of vegetation.] Warszawa, PWN; 1972.
- [19] Boratyńska K., Boratyński A. Systematyka i geograficzne rozmieszczenie. [Systematic and geographic distribution.] In: Białobok S. (ed.): Buk zwyczajny *Fagus sylvatica*. [European beech *Fagus sylvatica*] Warszawa – Poznań: PWN; 1990. p 27-73
- [20] Dzwonko, Z. Ekologia. [Ecology] In: Białobok S. (ed.): Buk zwyczajny *Fagus sylvatica*. [European beech *Fagus sylvatica*] Warszawa – Poznań: PWN; 1990. p237-328.
- [21] Giertych M. Genetyka. [Genetics] In: Białobok S. (ed.): Buk zwyczajny *Fagus sylvatica*. [European beech *Fagus sylvatica*] Warszawa – Poznań: PWN; 1990. p193-237.
- [22] Brzeziecki B. Skale nominalne wymagań klimatycznych gatunków leśnych. [Nominal climatic requirements forest species range.] *Sylwan* 1995; 139(3): 53-65.
- [23] Tarasiuk S. Buk zwyczajny (*Fagus sylvatica* L.) na obrzeżach zasięgu w Polsce. Warunki wzrostu i problemy hodowlane. [European Beech (*Fagus sylvatica* L.) in its marginal sites in Poland. Growth conditions and silvicultural problems.] Warszawa, Fundacja „Rozwój SGGW”; 1999.
- [24] Szafer, W. The significance of isopollen lines for the investigation of the geographical distribution of trees in the Post-Glacial period. *Bulletin de l'Academie Polonaise des Sciences et des Lettres, Serie B* 1935; 1: 235-239.
- [25] Huntley, B., Birks, H.J.B. An atlas of past and present pollen maps for Europe: 0–13000 years ago. Cambridge Univ. Press; 1983
- [26] Ralska-Jasiewiczowa, M. Isopollen maps for Poland: 0–11,000 years B.P. *New Phytology* 1983; 94: 133-175.
- [27] Hazler, K., Comps, B., Šugar, I., Melovski, L., Tashev, A., Gračan, J. Genetic structure of *Fagus sylvatica* L. populations in Southeastern Europe. *Silvae Genetica* 1997; 46(4): 229-236.
- [28] Sułkowska M. Isoenzyme analysis chosen provenances of beech (*Fagus sylvatica* L.) on the provenance trial in Bystrzyca Kłodzka. *Sylwan* 2002; 146 (2): 129-137.
- [29] Gömöry, D., Paule, L., Schvadchak, M., Popescu, F., Sułkowska, M., Hynek, V., Longauer, R. Spatial patterns of the genetic differentiation in European beech (*Fagus sylvatica* L.) at allozyme loci in the Carpathians and adjacent regions. *Silvae Genetica* 2003; 52(2): 78-83.

- [30] Sulkowska, M. Genetic and ecotype characterization of European beech (*Fagus sylvatica* L.) in Poland. *Acta Silvatica et Lignaria Hungarica* 2010; 6: 115-122.
- [31] Sułkowska, M., Gömöry, D., Paule, L. Genetic diversity of European beech in Poland estimated on the basis of isoenzyme analysis. *Folia Forestalia Polonica – series A Forestry* 2012; 54(1): 48-55.
- [32] Pastorelli R., Smulders M. J. M., Westende Van't W. P. C., Vosman B., Giannini R., Vettori C. Vendramin G. G. Characterization of microsatellite markers in *Fagus sylvatica* L., *Fagus orientalis* Lipsky. *Molecular Ecology Notes* 2003; 3, 76-78.
- [33] Vornam B., Decarli N., Gailing O. Spatial distribution of genetic variation in natural beech stands (*Fagus sylvatica* L.) based on microsatellite markers. *Conservation Genetics* 2004; 5: 561-570.
- [34] Grivet, D., Heinze, B., Vendramin, G. G., and Petit, R. J. Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molecular Ecology Notes* 2001; 1(4), 345-349.
- [35] Weising K, Gardner RC. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* 1999; 42(1):9-19
- [36] Nei, M. Genetic distance between populations. *American Nature* 1972; 106: 283-292.
- [37] Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latalowa, M., Litt, T., Paule, L., Roure, J. M., Tantau, I., van der Knaap, W. O., Petit, R., Beaulieu, J. L. A new scenario for the Quaternary history of European beech populations: paleobotanical evidence and genetic consequences. *New Phytology* 2006; 171(1): 199-221.
- [38] Nowakowska J.A., Oszako T. Health condition and genetic differentiation level of beech in the Siewierz Forest District assessed with cpDNA markers. *Sylwan* 2008; 152 (9): 11-20.
- [39] Starke, R., Ziehe, M., Müller-Starck, G. Viability selection in juvenile populations of European beech (*Fagus sylvatica* L.). *Forest Genetics* 1996; 3(4): 217-255.