

SEGREGATION OF AFLP MARKERS IN *BETULA PENDULA* (ROTH)Satu Åkerman^{1,2}, Jussi Tammissola³, Mikko Regina¹, Veli Kauppinen³ & Seppo Lapinjoki^{1*}¹ University of Kuopio, Department of Pharmaceutical Chemistry, P. O. BOX 1627, FIN-70211 Kuopio, Finland² Present address: Dept. Pharmaceutics³ Technical Research Centre of Finland, Biotechnology and Food Research, P. O. BOX 1505, FIN-02044 VTT, Finland

* Corresponding author

Received June 5, 1996; accepted September 18, 1996

ABSTRACT

The generation and segregation of AFLP (amplified fragment length polymorphism) markers were studied in European white birch. The recently published technique is based on PCR amplification of restriction fragments from digested total DNA by utilizing known oligomers (adapters) linked to the ends of the fragments. PCR primers partially complementary to the adapters and the restriction site are used to create the marker bands. In the present work *EcoRI* or *TaqI* were used for digestion. The primers had three arbitrary bases at the 3' ends for selective amplification of only a subset of the initial fragments. Each primer produced about 40 marker bands, of which about 10 in each case were found polymorphic between the parents of the full-sib progeny trees studied. When the segregation of alleles in the progeny was studied at 31 polymorphic loci, 29 of these were confirmed to express Mendelian ratios. The results indicate that the AFLP markers are very suitable for fingerprinting and genetic marker mapping in the birch. We found them superior to RAPD markers particularly in repeatability. The AFLP technique also produced substantially more polymorphic loci per PCR reaction. It was possible to use an automated DNA sequencer for marker band scanning after initial evaluation of the marker profiles on agarose gel.

Keywords: AFLP, DNA markers, PCR, *Betula pendula*, DNA fingerprinting

INTRODUCTION

New types of DNA markers have been developed in recent years. The old but still useful RFLPs (TANKSLEY *et al.* 1989) are being replaced by RAPDs (WILLIAMS *et al.* 1990), and microsatellites or SSRs (RAFALSKI & TINGEY 1993, MORGANTE & OLIVIERI 1993). The RAPDs can be produced very easily and from a small amount of template DNA by PCR without any sequence information, which has led to their wide application for fingerprinting purposes. However, their application to more demanding problems in population studies and breeding is restricted or disturbed by their dominant nature and also by inherent technical problems that may result in artefacts (LAMBOY 1994), unsatisfactory reproducibility (SCHIERWATER & ENDER 1993), and common occurrence of missing or non-parental fragments (RIEDY *et al.* 1992, SCOTT *et al.* 1992). Such problems were also encountered with the birch, although RAPDs were successfully applied in parentage analysis (ÅKERMAN *et al.* 1995).

To overcome the marker generation problems encountered with RAPDs and to increase the yield of useful markers per a PCR reaction, we decided to make a basic evaluation on the applicability of the AFLP

markers introduced recently (ZABEAU & VOS 1993, VOS *et al.* 1995). The technique is based on PCR amplification of restriction fragments from digested total DNA by utilizing known oligomers (adapters) linked to the ends of the fragments. PCR primers partially complementary to the adapters and the restriction site are used to create the marker bands. The primers have three arbitrary bases at the 3' ends for selective amplification of a subset of the total fragments. The method is regarded to combine the reliability of RFLPs with the power of the PCR technique. Compared to RAPDs, this novel technique has a desirable feature related to the full-length selective primers. They allow the use of stringent primer annealing conditions in PCR, which results in a specific and reproducible co-amplification of large arrays of selected restriction fragments. On the other hand, AFLP analysis involves the fragmentation and ligation steps, which may create trouble in repeatability when isolation of sufficiently pure DNA of steady quality is a particular problem. This is the case with birch and most other forest trees that have high content of phenolics in the leaves and other tissues that are readily available for DNA isolation (JOHN 1992). In addition to the more complex and expensive fragment generation chemistry,

polyacrylamide sequencing gels are recommended for fragment separation and scoring in AFLP (VOS *et al.* 1995), while the RAPDs are routinely run on agarose (WILLIAMS *et al.* 1990).

In the present study, generation of the markers was first surveyed using agarose electrophoresis to obtain initial information on performance of the AFLP chemistry with birch DNA, and to allow technical comparison with the RAPD method used earlier (ÅKERMAN *et al.* 1995). Final performance of the AFLP technique was evaluated in a segregation analysis, in which polyacrylamide sequencing gels and fragment scanning with an automated DNA sequencer were employed.

MATERIALS AND METHODS

Plant material

Two curly birch [*Betula pendula* var. *carelica* (Merklín) Hejtmánek] trees and their full-sib progeny from the collections of The Finnish Forest Research Institute in southern Finland were selected as the birch material in the study. The parent trees (registered as E 7881 and E 7882) grow in an old birch alley at the Punkaharju research station. Their full-sib progeny from a controlled cross in 1981 was grown at the Ruotsinkylä experimental area. Thirty-three progeny trees were included in the present study.

The natural origin of the parent trees is unknown, since it is not defined in the old planting documents. The curly birches are very rare in Finland and most of them are found in the same (Hauho) district. Hence, it is presumable that the two trees may originate from a single or nearby natural stands in this region, and may accordingly be genetically related. The seed parent is always given first, when crosses are described.

Arbitrary branches from the trees were cut and undamaged whole leaves were picked and rinsed by immersing in water. The wet leaves were packed in polyethylene bags and frozen between blocks of solid CO₂ within 1 h after the branches were cut. The samples were stored further at -75 °C until DNA extraction.

DNA isolation and fragmentation

The DNA was extracted according BOUSQUET *et al.* (1990). About 1.0 g samples of frozen birch leaves were ground in liquid nitrogen and homogenized in an extraction buffer (2% (w/v) N-cetyl-N,N,N-trimethylammoniumbromide (CTAB, Sigma Chemical Co., St. Louis, Mo.), 1.4 M NaCl, 0.2% (v/v) 2-mercaptoethanol, 20 mM EDTA, 100 mM Tris-HCl (pH 8.0), 1% (w/v) polyvinylpyrrolidone (PVP-40). The homogenate was incubated at 65 °C for 60 min, extracted with equal volume of chloroform-isoamyl alcohol (24:1,

v/v), and centrifuged at 1700 g for 10 min. The aqueous phase was transferred to a sterile tube and two-thirds volume of ice-cold isopropanol was added to precipitate the DNA. After centrifugation at 600 g the pellet was washed twice with 70% (v/v) ethanol in 10 mM ammonium acetate, dried and resuspended in 10 mM Tris-HCl, 1 mM EDTA, pH 7.5. An aliquot of 3 mg of the DNA was digested with restriction endonucleases *EcoRI* (37 °C 3h) or *TaqI* (65 °C 2h).

Adapter preparation and ligation

The adapters were constructed and synthesized as outlined by VOS *et al.* (1995). Equal amounts of partially complementary single strands were mixed and diluted to final concentration of 100 pmol adapter monomer/ml. The single strands were hybridized by incubating the mixture at 25 °C for at least 30 min. Structures of the final double-stranded adapters were as follows.

EcoRI: 5'-CTCGTAGACTGCGTACC
CTGACGCATGGTTAA-5'

TaqI: 5'-CGGTCCAGGACTCAT
CAGTCCTGAGTAGCAG-5'

For constructing the template fragments for PCR, 130 ng of the respective adapter and 3 mg of the digested DNA were incubated 3 h at 37 °C in a mixture containing 10 U T4-ligase (Pharmacia) and the commercial ligation buffer in a total volume of 50 ml. The template was precipitated by adding 80 ml 7.5 M NH₄-acetate and 100 ml of ice-cold 99% (v/v) ethanol and centrifuged 13 000 rpm 15 min in an Eppendorf-microfuge. The pellet was washed with ice-cold 99% (v/v) ethanol, dried and resuspended in 10 mM Tris-HCl, 1 mM EDTA, pH 7.5.

PCR conditions

PCR reaction mixtures (50 ml) contained 1 ml of the template DNA, 200 mM dNTP (Pharmacia), 1.2 mM FITC-dUTP (Boehringer-Mannheim), 150 ng of primer, 5U of *Taq* DNA-polymerase (Boehringer-Mannheim) and the commercial buffer for *Taq* DNA-polymerase. The PCR cycling was as described in VOS *et al.* (1995): The 1st cycle was 30s/94 °C, 30s/65 °C and 60s/72 °C, whereafter the annealing temperature was decreased by 0.7 °C for each of the following 11 cycles, and kept at the resultant 56 °C for the last 23 cycles. The total 34 cycles were run using a Perkin Elmer 9600 thermal cycler. The primer for *EcoRI* digested fragments was 5'-CTGCGTACCAATTC+ACA, and for the *TaqI* fragments 5'-TGAGTCCTGACCGA+CAA or

+CAC. The selective bases are separated by the "+" from the complementary regions, which are quoted by the enzyme names below (*cf.* VOS *et al.* 1995). In an initial technical evaluation the following primers were also used: *EcoRI* +CAG, +CAA, +AAC, and *TaqI*+ACC and +ACA.

Fragment analysis

The marker fragment generation was initially evaluated by separating the PCR products on 2.5% Metaphor agarose (FMC, Denmark). Electrophoresis was run (2h, 180V) with 1X TBE buffer and stained with ethidium bromide. The DNA molecular weight marker kit VI (Boehringer GmbH, Germany) containing pBR 328 DNA fragments between 2176 and 154 base pairs was used to estimate the fragment length.

In the segregation analysis, the PCR products were separated using an automated DNA sequencer (A.L.F., Pharmacia) with fluorescence detector. For the detection, 1.2 mM FITC-dUTP was included in the PCR mix described above. Samples of 8 ml from the PCR were loaded in a ReadyMix Gel (A.L.F.TM grade) containing 7.0 M urea and 5.7% (w/v) acrylamide. The electrophoresis was run for 10 h (1500V, 38 mA and 34W) and keeping the gel temperature at 40 °C.

Data processing

For each of the polymorphic AFLP loci included in the segregation study, the marker allele distribution (band *versus* no band) observed in the birch progeny was compared statistically with the appropriate expected Mendelian ratios. Instead of using the approximate chi square tests, exact two-tailed binomial tests were applied throughout, *i.e.* the binomial probabilities were summed up for all possible outcomes at least as extreme as the one observed in the experiment. For practical calculations, a Q-BASIC program was constructed, in which the precision necessary in large calculations was achieved by utilizing integer arithmetic and cancellation.

RESULTS AND DISCUSSION

As outlined by VOS *et al.* (1995), clear and dense fingerprint patterns of AFLP markers are generated most efficiently by simultaneous use of a rare and a frequent cutter for the restriction. One of the primers is present in a limiting concentration and labeled with ³²P to obtain an even band intensity. The procedure aims at relatively short (< 400 bp) fragments, which are separated on a denaturing polyacrylamide sequencing gel. In the present work we employed the technically easy combination of agarose gels and ethidium bromide

staining for initial evaluation of marker signal production. Accordingly, only rare cutters were used to keep the fragments long enough for resolution on the agarose. This approach allowed a more direct comparison of the results with those obtained earlier with RAPDs (ÅKERMAN *et al.* 1995), and the fewer markers generated also facilitated scoring them in the segregation study.

An initial evaluation of marker generation from *EcoRI* and *TaqI* digests was performed with sets of 4 primers in each case as defined in Materials and Methods. The tests were run using DNA from the parent trees. The *EcoRI* digest gave relatively clear band patterns of about 10 to 20 fragments on the agarose, whereas the *TaqI* digest gave more than 20 fragments with every primer. Most of the bands accumulated within the size range of 100 to 300 bp, which caused resolution problems. Because the *EcoRI*+ACA primer produced clearest pattern of relatively few bands, it was used to evaluate the reproducibility of fragment generation from the full-sib progeny under study. Figure 1 shows the AFLP patterns from the parent trees and 16 individuals of the progeny. Although the amount and processing of template were kept as constant as possible, the initial results indicated that the agarose electrophoresis and ethidium bromide staining are not sufficient to produce marker profiles of repeatedly comparable quality. For example, the 4th sample lane of Fig. 1 shows numerous bands in the progeny, which are not detectable in the parents. Repetition of the analyses showed that the poor repeatability was not due to inconsistency in the marker fragment generation, but to the limited useful range of the agarose/ ethidium bromide detection, *i.e.* expected marker fragments were present, but it was practically impossible to produce a

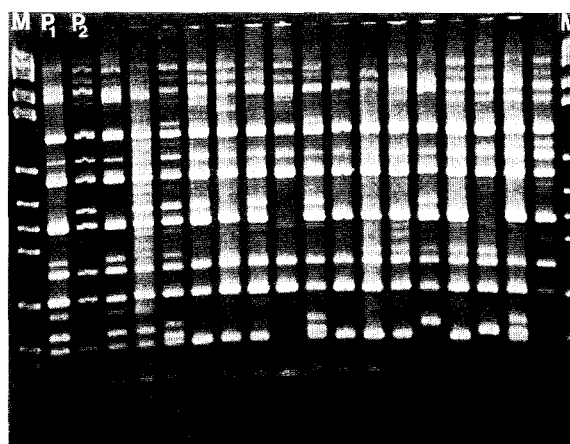


Figure 1. Ethidium bromide stained agarose gel with AFLP markers produced from the parent trees (P₁ and P₂) and 16 full-sibs (unmarked lanes) of a controlled cross. *EcoRI* was used for digesting and the *EcoRI* + ACA primer for PCR. M = fragment size marker lanes, with top band of 2176 and bottom band of 220 base pairs.

Table 1 Mendelian segregation of AFLP alleles in a F₁ progeny of white birch. *EcoRI* + ACA primer was used in the amplified reaction.

Fragment (bp)	Observed progeny phenotype ratio in F ₁ (+ : -)	Expected progeny phenotype ratio in F ₁ (+ : -)	Probability ¹
250	23 : 10	3 : 1	0.55
265	16 : 17	1 : 1	1.00
280	26 : 7	3 : 1	0.69
380	14 : 19	1 : 1	0.49
455	19 : 14	1 : 1	0.49
500	14 : 19	1 : 1	0.49
525	22 : 11	3 : 1	0.31
540	21 : 12	3 : 1	0.16
600	23 : 10	3 : 1	0.55
640	25 : 9	1 : 1	0.0046**

¹⁾ Binomial two-tailed risk probability of obtaining a ratio at least as extreme as the one observed. The conventional levels of statistical significance of deviations from Mendelian expectations are indicated as follows: * - (0.01 ≤ P < 0.05); ** - (0.001 ≤ P < 0.01); *** - (P < 0.001).

Table 2 Mendelian segregation of AFLP alleles in a F₁ progeny of white birch. *TaqI* + CAA primer was used in the amplification reaction.

Fragment (bp)	Observed progeny phenotype ratio in F ₁ (+ : -)	Expected progeny phenotype ratio in F ₁ (+ : -)	Probability ¹
250	15 : 13	1 : 1	0.85
285	20 : 8	3 : 1	0.66
300	22 : 6	3 : 1	0.82
305	24 : 4	3 : 1	0.27
320	25 : 3	3 : 1	0.12
340	18 : 10	1 : 1	0.18
350	17 : 11	1 : 1	0.34
375	26 : 2	3 : 1	0.028*
400	12 : 16	1 : 1	0.57
587	12 : 16	1 : 1	0.57
610	14 : 14	1 : 1	1.00

¹⁾ explanation as in Table 1

multisample gel, in which amounts of the different PCR products in the lanes are consistent enough to allow simultaneous detection of all the fragments.

For the final segregation analysis on polyacrylamide gel we chose three primers, which gave the clearest band patterns for the parents on agarose gel. When the gels were scanned in the DNA sequencer, a total of about 40, 40 and 30 different fragments were detected for the parent trees after PCR with *EcoRI*+ACA, *TaqI*+CAA, and *TaqI*+CAC, respectively.

In each case, ten or eleven of these AFLP loci (*i.e.* on average 25 % of the AFLP loci detected) showed polymorphism in the progeny. The remaining loci were monomorphic in the material. The polymorphism is of the plus-minus (*i.e.* band *versus* no band) type, since possible allelic variation in fragment length is not

detectable with the AFLPs (VOS *et al.* 1995). The segregation ratios observed in the progeny at the polymorphic AFLP loci are analyzed in detail in Tables 1 to 3. The fragment sizes in the Tables are defined as approximate mid-points of three bases wide scoring windows. These were used because the AFLPs were recorded as relatively broad and two-peaked fluorescence signals. The peak form follows from our technique to label both of the fragment strands, which have slightly different migration times in the denaturing gel. The approach did not complicate interpretation of the results, since the relatively few AFLPs generated now could be well identified and followed by using the three base window. As described by VOS *et al.* (1995), labelling of only one strand of the AFLP fragments and limiting the amount of the respective primer will results

Table 3 Mendelian segregation of AFLP alleles in a F₁ progeny of white birch. *TaqI* + CAC primer was used in the amplification reaction.

Fragment (bp)	Observed progeny phenotype ratio in F ₁ (+ : -)	Expected progeny phenotype ratio in F ₁ (+ : -)	Probability ¹
200	15 : 17	1 : 1	0.86
205	19 : 13	1 : 1	0.38
220	16 : 16	1 : 1	1.00
250	19 : 13	1 : 1	0.38
253	26 : 6	3 : 1	0.54
275	14 : 18	1 : 1	0.60
280	14 : 18	1 : 1	0.60
450	28 : 4	3 : 1	0.15
775	21 : 11	1 : 1	0.11
780	17 : 15	1 : 1	0.86

¹ explanation as in Table 1

in very sharp and even signal generation. This is essential if high amounts of AFLPs are intentionally generated per a PCR reaction.

When the probabilities of the observed marker phenotype segregation ratios were tested statistically, two of the 31 showed significant deviations from the expected ratios, *i.e.* the 640 bp band in Table 1 ($P = 0.0046$), and the 375 bp band in Table 2 ($P = 0.028$). For all the other segregations observed, the probabilities were between 0.11 and 1.0. When as many as 31 polymorphic loci are to be tested separately, about one statistically significant deviation ($P < 0.05$) is expected to occur merely due to chance, because $1/31 = 0.032$. In our study, two such deviations were observed. This slight excess may be due to a certain methodological inconsistency. Namely, after the original description of AFLP for the patent (ZABEAU & VOS 1993), the inventors have reported that primers with three selective bases tend to produce a low level of mismatch amplification (VOS *et al.* 1995). They have been able to solve the problem by using a two-step procedure, which includes a preamplification using a primer with only the first selective base at the 3-end. On the other hand, in applications where it is not necessary to generate high amounts of AFLPs per a PCR reaction, it may be more practical to utilize only such marker loci that are reliably detected by the one-step PCR and to avoid the more laborious and expensive two-step procedure.

Of the 31 loci segregating among the progeny, 12 were heterozygous in both parents and the remaining were heterozygous in one of the parents and homozygous for the negative (without band) allele in the other parent, whereas no loci with a parental genotype combination +/+ and -/- occurred in our material as determined from the segregation ratios (Tables 1-3). Are these numbers consistent with the expectations in a Mendelian population? Taken strictly, this question

remains premature, until adequate information about the allele frequencies at each individual AFLP locus will be accumulated from the basic population of European white birch. Estimated from the above numbers, the mean frequency of the plus allele over the polymorphic loci is 0.35. At present the actual allele frequencies are not yet known and may vary widely between loci. However, even if a fixed allele frequency \bar{p}_+ were postulated over all loci, the numbers observed for the genotype combinations (0:12:19) do not differ statistically significantly ($P < 0.05$) from their respective expectations for allele frequencies in the range $0.3 \leq \bar{p}_+ \leq 0.39$. The conditional probabilities required for the calculation of these expectations are given in the Appendix.

The combination of the most dissimilar two parental genotypes (+/+ and -/-) was underrepresented in the observed results, irrespective of the value postulated for \bar{p}_+ . Although not statistically significant, this shortage may indicate that the two curly birch parental trees are in fact genetically related, as commented above (see Plant material). If some genetic relatedness is allowed (*e.g.* the *coancestry* $f = 0.1$), the observed results can be statistically explained for a wide range ($0.18 \leq \bar{p}_+ \leq 0.46$) of possible allele frequencies. Relatedness of the parent trees may also be a reason for the observed distortion of the two markers (Tables 1 and 2).

In conclusion, the AFLP markers were found suitable for fingerprinting in the birch, and results of the segregation analysis suggest that they also may be utilized in construction of index maps. Although they are technically more demanding than RAPDs, AFLP markers still proved easy to use if facilities for using sequencing gels are available. They are very efficient in practice, because a great number of qualified loci can be produced per PCR reaction. When comparing our

present experience with AFLP to the earlier with RAPDs (ÅKERMAN *et al.* 1995), we feel that both of them may be used for any applications, where dominant markers are sufficient. However, the AFLPs seem to be the method of choice in studies in which good repeatability is essential, *e.g.* in research where marker data from different laboratories must be combined or compared.

ACKNOWLEDGEMENT

The work received financial support from the Scientific Council for Research of Environment and Natural Resources in Finland.

REFERENCES

- ÅKERMAN, S., TAMMISOLA, J., LAPINJOKI, S. P., SÖDERLUND, H., KAUPPINEN, V., VIHÄRÄ-AARNIO, A., REGINA, M. & HAGQVIST, R. 1995: RAPD markers in parentage confirmation of a valuable breeding progeny of European white birch. *Can. J. For. Res.* **25**:1070–1076.
- BOUSQUET, J., SIMON, L. & LALONDE, M. 1990: DNA amplification from vegetative and sexual tissues using polymerase chain reaction. *Can. J. For. Res.* **20**:254–257.
- FALCONER, D. S. & MACKAY, T. F. C. 1996: Introduction to Quantitative Genetics. 4th ed. Longman, London, 464 pp.
- JOHN, M. E. 1992: An effective method for isolation of RNA and DNA from plants containing polyphenolics. *Nucl. Acids Res.* **20**:2381.
- LAMBOY, W. F. 1994: Computing genetic similarity coefficients from RAPD data: the effects of PCR artifacts. *PCR Methods Applic.* **4**:38–37.
- MORGANTE, M. & OLIVIERI, A. M. 1993: PCR-amplified microsatellites as markers in plant genetics. *Plant J.* **3**: 175–182.
- RAFALSKI, J. A. & TINGEY, S. V. 1993: Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends Genet.* **9**: 275–280.
- RIEDY, M. F., HAMILTON, W. J. & AQUADRO, C. F. 1992: Excess of non-parental bands in offspring from known primate pedigrees assayed using RAPD PCR. *Nucleic Acids Res.* **20**:918.
- SCHIERWATER, B. & ENDER, A. 1993: Different thermostable DNA polymerases may amplify different RAPD products. *Nucleic Acids Res.* **21**:4647–4648.
- SCOTT, M. P., HAYMES, M. & WILLIAMS, S. M. 1992: Parentage analysis using RAPD PCR. *Nucleic Acids Res.* **20**:5493.
- TANKSLEY, S. D., YOUNG, N. D., PATERSON, A. H. & BONIERBALE M. W. 1989: RFLP mapping in plant breeding: new tools for an old science. *Bio/Technology* **7**:257–264.
- WILLIAMS, J. G. K., KUBELIK, A. R., LIVAK, K. J., RAFALSKI, J. A. & TINGEY, S. V. 1990: DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* **18**:6531–6535.
- VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., VAN DE LEE, T., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M. & ZABEAU, M. 1995: AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**:4407–4414.
- ZABEAU, M. & VOS, P. 1993: Selective restriction fragment amplification: a general method for DNA fingerprinting. Eur. Pat. Appl. EP 0534858A1, 43 p.

APPENDIX

A large array of biallelic loci are studied in two trees taken from a randomly mating population. The genetic relatedness between the trees is expressed in terms of the *coancestry* f , which measures the probability that a randomly chosen allele from each will be identical by descent (FALCONER & MACKAY 1996). Postulating that the frequency of the plus allele, p_+ , is identical in each locus, the following probabilities can be derived for the different parental genotype combinations over loci. The formulas are derived for each "polymorphic combination", *i.e.* situations in which phenotypic differences will occur either between the parents or their cross progeny. As presented in the Discussion, there are three such cases, namely

$$P\{A\} \equiv P\{+/- \otimes +/- \text{ (order disregarded)}\} \\ = 4 \cdot (1-f)^2 \cdot [f + (1-f) \cdot p_+] \cdot [f + (1-f) \cdot p_-] \cdot p_+ p_- ,$$

$$P\{B\} \equiv P\{+/- \otimes -/- \text{ (order disregarded)}\} \\ = 4 \cdot (1-f)^2 \cdot [1 - (1-f)^2 \cdot p_+] \cdot p_+ p_-^2 ,$$

$$P\{C\} \equiv P\{+/+ \otimes -/- \text{ (order disregarded)}\} \\ = 2 \cdot (1-f)^4 \cdot p_+^2 p_-^2 .$$

The respective conditional probabilities, *i.e.* the relative frequencies of these combinations, assuming a polymorphic situation, are obtained by dividing each of these probabilities by their sum, *e.g.*

$$P\{A | A \text{ or } B \text{ or } C\} = P\{A\} / [P\{A\} + P\{B\} + P\{C\}] .$$