J Appl Genetic 47(3), 2006, pp. 221–230

Development of an efficient retrotransposon-based fingerprinting method for rapid pea variety identification

Petr Smýkal

AGRITEC Plant Research Ltd., Plant Biotechnology Department, Šumperk, Czech Republic

Abstract. Fast and efficient DNA fingerprinting of crop cultivars and individuals is frequently used in both theoretical population genetics and in practical breeding. Numerous DNA marker technologies exist and the ratio of speed, cost and accuracy are of importance. Therefore even in species where highly accurate and polymorphic marker systems are available, such as microsatellite SSR (simple sequence repeats), also alternative methods may be of interest. Thanks to their high abundance and ubiquity, temporary mobile retrotransposable elements come into recent focus. Their properties, such as genome wide distribution and well-defined origin of individual insertions by descent, predetermine them for use as molecular markers. In this study, several *Ty3-gypsy* type retrotransposons have been developed and adopted for the inter-retrotransposon amplified polymorphism (IRAP) method, which is suitable for fast and efficient pea cultivar fingerprinting. The method can easily distinguish even between genetically closely related pea cultivars and provide high polymorphic information content (PIC) in a single PCR analysis.

Key words: fingerprinting, IRAP, pea, polymorphism, retrotransposon.

Introduction

The analysis of genetic diversity and relatedness among different individuals, species or populations is the central topic in genetics. Marker technology, based on polymorphism, started from analysis of seed storage proteins and isozymes. These markers have been superseded by DNA-based methods that generate fingerprints, i.e. distinctive patterns of DNA fragments (Weissing et al. 2005). In addition to estimation of the genetic uniqueness of each accession, relative to others, the redundant germplasm is removed in order to increase the efficiency of resource management. Passport data are often not fully reliable, so traditional morphological or biochemical characterizations are restricted due to limited variation and influence of the environment. Even though not currently implemented and legally recognized by UPOV (The International Union for the Protection of New Varieties of Plants), molecular markers are being used as supplementary descriptors (UPOV 2002).

Since in a majority of cultivated crops no or very limited genome sequence information is available, the RAPD (randomly amplified polymorphic DNA) method introduced by Williams et al. (1990) has become widely used and popular. Nowadays, microsatellites (tandem repeats of short DNA motifs) are frequently used. The initially used ISSR (inter-simple sequence repeats polymorphism, Zietkiewicz et al. 1994) a broadly applicable but error-prone technique was gradually substituted by locus-specific, robust and co-dominant SSR (simple sequence repeats, Morgante and Olivieri 1993). Both RAPD and ISSR techniques were applied to pea, but low polymorphism was obtained (Samec et al. 1998; Simioniuc et al. 2002). In 1995, a universal AFLP (amplified length fragment polymorphism) technique was invented and since then it has been

Received: April 7, 2006. Accepted: June 20, 2006.
Correspondence: P. Smýkal, AGRITEC Plant Research Ltd., Plant Biotechnology Department, Zemědělská 2520/16, CZ-787 01 Šumperk, Czech Republic; e-mail: smykal@agritec.cz

widely used (Vos et al. 1995). This method works on any organism without any knowledge of genomic sequences, but it is prone to contamination sensitivity and is labour intensive. Major drawbacks of AFLP, RAPD and ISSR are due to unspecific primers, producing multilocus banding patterns of dominantly inherited markers. Subsequently, abundant repeat motives forming microsatellites were isolated with flanking genomic DNA regions, leading to design of locus-specific and robust SSR (simple sequence repeats) markers, also for pea (Burstin et al. 2001; Ford et al. 2002; Loridon et al. 2005). Integration of several available marker systems for pea germplasm evaluation was used by Baranger et al. (2004) and Tar´an et al. (2004). The great advantage of microsatellite analysis is the accuracy, high polymorphism, multiallelism, genome coverage, and co-dominance (Becher et al. 2000). Microsatellites are commonly used for genome mapping, estimation of genetic diversity, gene tagging, and marker-assisted selection (McCouch et al. 1997). Therefore, the working group on biochemical and molecular techniques of UPOV has identified SSR markers as the most widely used marker system for plant variety characterization (UPOV, 2002).

Yet another marker system has emerged during the last few years, when it become evident that another class of highly abundant sequences dominate in genomes. Retrotransposons (RTNs), together with transposons, represent a ubiquitous class of repetitive elements in all eukaryotic genomes. They can constitute a majority of the genome, e.g. in grasses making up to 60–90% (Kumar and Bennetzen 1999). Both their ubiquitous nature and activity in creating genomic diversity (Kazazian 2004) by integrating stable DNA fragments into dispersed chromosomal loci make these elements ideal for use as molecular markers (Kumar and Hirochika 2001). Unlike transposons, the retrotransposons do not excise, but are transcribed and by reverse transcription inserted into DNA as copies of the mother element. Retrotransposons are divided into 2 separate groups (Kumar and Bennetzen 1999): those containing long terminal repeats (LTRs) and those lacking LTRs. Retrotransposons share many similarities with the retroviruses, both in their organization and the encoded gene products. LTR elements are further divided into 2 major groups: *Ty1-copia* and *Ty3-gypsy* types, on the basis of organisation of their coding domains. Any given element insertion site, which was shared between accessions, was likely to occur in a common ancestor, which retrotransposon markers currently the most informative for phylogenetic and pedigree studies (Batzer and Deininger 2002; Shedlock and Okada 2000). The developed retrotransposon-based marker methods rely on the principle that a joint between retrotransposon and genomic DNA is formed during the integration process. These joints can be detected by PCR amplification between a primer corresponding to the retrotransposon and a primer matching a nearby motif in the genome (Schulman et al. 2004). The SSAP (sequence-specific amplified polymorphism) method derives from AFLP (Waugh et al. 1997). In IRAP (inter-retrotransposon amplified polymorphism), segments between 2 nearby elements are amplified (Kalendar et al. 1999). Retrotransposon-based insertional polymorphism (RBIP) provides a locus-specific co-dominant marker system, suitable for high-throughput analysis (Flavell et al. 1998, 2003; Jing et al. 2005). The SSAP is now used in numerous species where terminal LTR sequences can be effectively cloned (Pierce et al. 1998; Schulman et al. 2004). However, proper use of the SSAP technique requires either radioactivity or fluorescent labelling of primers and product detection. Both these drawbacks could be overcome by the IRAP method.

In garden pea (*Pisum sativum* L.), a large body of information on various retrotransposons exists. Special attention was paid to the *PDR-1* element, a representative of the *Ty1-copia* class (Ellis et al. 1998). This low-copy element was successfully used in numerous studies of pea phylogeny (Ellis et al. 1998; Pearce et al. 2000;). Lately, *PDR-1* individual insertion loci were sequence- characterized, some positionally mapped and converted into high-throughput RBIP markers ready in a DNA chip format (Flavell et al. 2003; Jing et al. 2005). However, these applications require many PCR analyses to gain sufficient information content.

The objective of the present study was to demonstrate the development and usefulness of a relatively simple IRAP protocol for efficient DNA fingerprinting in pea. The study focused on another group of retrotransposons, *Ty3-gypsy* types, which are far more abundant, being represented by thousands of copies and making up to 10–30% of the pea genome (Neumann et al. 2001). Our research team has so far investigated *Ogre* (Neumann et al. 2003), *Cyclop* (Chavanne et al. 1998), *Pigy* (Neumann et al. 2003), *PisTR-A* (Neumann et al. 2001) and a miniature in-

Table 1. Origin of pea varieties used in this study (all registered for cultivation in the Czech Republic)

No	Cultivar	Pedigree	Breeder
1	Adept	PSS 373/75 × Tolar	SELGEN CZ
2	Alan	NDR x (R27 × Danielle)	ELITA CZ
3	Baryton	Renata × Erbi	Sudwestsaat GbR
4	Bohatyr	(Kralicky Unicum × Pyram) Dick Trom	SELGEN CZ
5	Canis	Bohatyr × U 51041	Svalof Weibull AB
6	Carrera	Belinda × Cebeco 756-921	Cebeco Seeds
7	Catania	Bohatyr × Princess	Lochow Petkus GmbH
8	Garde	Cebeco 1140 × Cebeco 1441	Cebeco Seeds
9	Gotik	Consort × LU 28B	SELGEN CZ
10	Grana	linie A × B × C	Nordsaat Saatzucht
11	Hardy	Baccara × Duel	SERASEM
12	Harnas	Canis × S 610.1	Clovis MATTON N.V.
13	Herold	(Sum × LU0040) × Emerald	SELGEN CZ
14	Jackpot	Bohatyr × Solara	Axel Toft Grovvarer A/S
15	Janus	DE 719 × Finale	Slovak breeding station-
16	Kamelot	(Sum × LU0040) × Emerald	SELGEN CZ
17	Komet	(Tyrkys × Erygel) × Bohatyr	SELGEN CZ
18	Lantra	Maro × Ricardo	Cebeco Seeds
19	Madonna	Solara × (Bohatyr × U50004)	Norddeutsche Pflanzenzucht Hans
20	Menhir	Hamil × Tolar	SELGEN CZ
21	Merkur	FP-81-61 × Tolar	SELGEN CZ
22	Olivin	Luzany15 × Dippes Gelbe Victoria	Slovak breeding station
23	Pegas	LU-FK-Y × (Multipod × Finale) × 141	ELITA CZ
24	Power	Solara × LU 57	Saatzucht Hans
25	Primus	((Multipod × Finale) × 141) × ((Dick Trom × HP) × Ludik)	ELITA CZ
26	Profi	Bohatyr × FM 420062	danisco Seeds
27	Romeo	(Dick Trom) × HP) × Ludik	ELITA CZ
28	Sonet	FP-81-230 × Bohatyr	SELGEN CZ
29	Sponzor	Bohatyr × Trille	Axel Toft Grovvarer A/S
30	Tempra	91594 × Ceb 1141	Cebeco Seeds
31	Terno	LU-134 × Rustic	SELGEN CZ
32	Tyrkys	Luzansky Krl-2 × Dick Trom	SELGEN CZ
33	Zekon	(Sum × LU0040) × Emerald	SELGEN CZ

verted-repeat transposable element, i.e. MITE (Macas et al. 2005) for use in IRAP format suitable for fast and efficient pea genotype fingerprinting.

Materials and methods

Plant material and DNA isolation

Leaves from 33 currently grown commercial pea cultivars registered in the Czech Republic (Table 1) were taken from greenhouse-grown plants. Genomic DNA was isolated according to a modified CTAB protocol (Doyle and Doyle 1987). Approximately 100 mg of leaves were manually homogenised with a pestle in 500 µL of CTAB extraction buffer (100 mM Tris-HCl pH 8.0, 1.4M NaCl, 20 mM EDTA, 2% CTAB, 1% PVP MW 360 000, 0.4% mercaptoethanol added prior to use) with addition of quartz sand, incubated at 65°C for at least 60 min with occasional mixing. Subsequently 500 µL of chloroform was added, vigorously vortexed and centrifuged for 15 min at 13 000 g. Upper aqueous phase was transferred into a new tube, supplemented with 0.5 × volume of 5M NaCl and 1× volume of iso-propanol and left at –20°C for at least 60 min. Then the sample was centrifuged for 30 min at 13 000 g, the pellet was washed with 75% ethanol and left to dry. Finally the pellet was resuspended in 400 µL of TE buffer containing 2 µL RNaseA (10 mg mL^{-1}) (Sigma, USA), left for at least 30 min at room temperature, extracted with chloroform, precipitated, and stored at 4°C for a short time or at –20°C. Optionally, genomic DNA was isolated by the salt-acidic protocol (Guillemaut and Marechal-Drouard 1992) or commercially available kits, such as DNeasy Plant Mini Kit (Qiagen, Germany) and Invisorb Spin Plant Mini Kit (Invitek, Germany), according to manufacturers' protocols.

IRAP analysis

Reverse (R) and forward (F) primers were designed to match close to the 5' and 3' termini of LTR sequences of the following retrotransposons: for *Ogre* (Y299398), R 5′-GTG GGC TGG GCT TTA GTG AGA TGC TTT CC-3' and F 5'-TCG CGA GAC CAT GTC TTT TCC CAG GTT TAC-3'; for *Cyclop* (J000640), R 5'-GCAAGG AAACGGAGTGAAAGATGC-3′ and F 5'-CGA

TAT CTC ACA ATC CCT GTG GAG AC-3'; for *Pigy* (AY299398), R 5'-ATC ATC AAA GTA TCA TCC GCC TTA GC-3' and F 5'-ACG CTC GTC ACA TGC CCG TGG CGG TC-3'; for *PisTR-A* (AF300839), R 5'-ACA TCC TGA AGA GAG CGA AGG-3' and F 5'-GTT ATG GAT AAT CTT CCA TGC GAG G-3'; and for *Stowaway*-MITE (AY833549), 5'-CTG TGA ATT TTT CCT TGC CTC CCT C-3'.

The IRAP protocol was developed by adaptation of an original method of Kalendar et al. (1999). PCR amplification was performed in a 20-μL reaction mixture containing: 20–50 ng of DNA, 1 × PCR buffer (75 mM Tris-HCl, pH 9.0, 50mM KCl, 20 mM $(NH_4)_2SO_4$), 2.5 mM $MgCl_2$, 4 pmol of each primer, 100 μM dNTP, and 1 U of *Taq* polymerase (BioTools, Spain). Amplification was performed in a Mastercycler (Eppendorf, Germany) in 0.2-mL tubes (TreffLab, Switzerland). After 4 min of initial denaturation at 94°C, amplification was carried out in 35 cycles at 94°C for 30 sec, 60°C for 30 sec and 72°C for 2.5 min, with final extension at 72°C for 10 min. Products were resolved on 8% non-denaturing Tris-Boric Acid-EDTA polyacrylamide electrophoresis (TBE-PAGE) on Protean II gel electrophoresis unit (Bio-Rad, Germany) according to product sizes (from 100 bp to 1000 bp) at 30 mA per 25 cm Protean II gel for 2–3 h. Gels were subsequently stained with ethidium bromide and visualized with UV. Alternatively, the products were resolved on horizontal 1.5% agarose gel (Serva, Germany).

Band scoring and data evaluation

IRAP data were recorded from the PAGE gel by using CrossChecker 2.91 software (Buntjer 1999) and visual inspection upon band size classification and scoring in the binary code. Monomorphic bands were removed from the raw scored data sets. The sets that contained missing values were also excluded at that stage. All DNA marker data were processed by NTSYS-pc 2.1 software (Rohlf 2000) by using a SIMQUAL module with the Nei-Li 1979 coefficient, also known as the Dice coefficient of similarity, $S = 2n_{ab}/(n_a + n_b)$, where n_a and n_b represent the number of bands present in lines *a* and *b*, and n_{ab} represents the number of shared bands. Polymorphic information content (PIC) was calculated for each marker by using a calculator (http://www.agri.huji.ac.il/~weller/Hayim/parent/PIC) based on the following formula:

$$PIC = 1 - \sum P_i^2 - \sum \sum P_i P_j^2 ,$$

where P_i is the frequency of an individual genotype.

Results and discussion

Selection of pea retrotransposons suitable for the IRAP technique

IRAP principally relies on amplification of DNA fragments between retrotransposons that are sufficiently close to be amplified by *Taq* polymerase (Kalendar et al. 1999). Terminal LTRs unique to each element are arranged as inverted repeats on

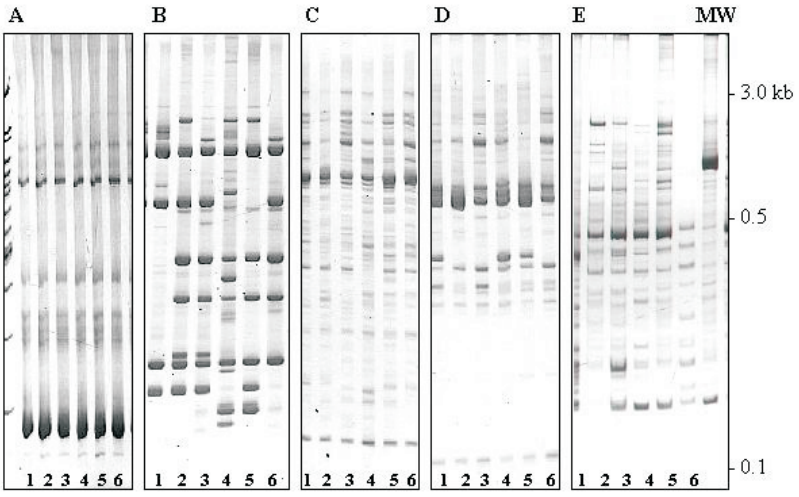

Figure 1. IRAP analysis with various pea *Ty3-gypsy* type retrotransposons–*Pis-TRA* (A), *Cyclop* (B), *MITE* (C), *Pigy* (D) and *Ogre* (E) – performed with F and R primers. Genomic DNA (50 ng) from cvs. Alan (1), Bohatýr (2), Merkur (3), Menhir (4), Komet (5), or Adept (6) was used per PCR reaction. Negative image of EtBr stained and UV visualized on 20-cm 8% PAGE gel is shown. MW = 100-bp DNA molecular weight marker (Fermentas).

both ends of a given element (Kumar and Bennetzen 1999). Therefore, 2 outward-facing primers can be designed. Theoretically, annealing of even a single outward-facing primer sufficiently close to an LTR end can be used. The IRAP method works due to frequent nested element insertions and a tendency of retrotransposons to cluster together. Five representatives of *Ty3-gypsy* type retrotransposons and one MITE have been selected to design primers sufficiently close to both ends of unique LTR regions (within 50–150 bp) and next applied in F/R combinations on pea genomic DNA of 6 varieties. The primer annealing temperature was set up sufficiently high (58°C) in all cases to minimize unspecific amplification. Analyses of *Cyclop* and *Ogre* elements performed with both F and R primers have produced consistently the largest number of clearly separated DNA fragments (Figure 1). Depending on pea accession tested, band number varied from 10 to 21, of which a high number proved to be polymorphic (average PIC from our set of 33 accessions was 0.299 for *Cyclop* and 0.236 for *Ogre*). Most information is gained when both F and R primers are used, but additional DNA fragments can be amplified with single primers. Combination of *Ogre* and *Cyclop* primers (*Ogre*-F *Cyclop*-F etc., 12 combinations in total) gave a pattern that was not simply additive (data not shown). Other elements (*PisTR-A, Pigy*) gave either only a small number (15) and/or poorly resolved smearing bands (*MITE*, *PisTR-A*) (Figure 1). The low number of fragments indicates that there are fewer elements of the respective retrotransposon in the pea genome, and/or that they are too distantly spaced for efficient IRAP. Conversely, too numerous bands suggest the opposite situation.

Optimization of the IRAP method

DNA concentration and isolation

In order to develop a robust and user-friendly method, a broad spectrum of DNA concentrations was tested. Genomic DNA isolated by CTAB and Qiagen Plant DNeasy protocol was used in concentrations of 10, 20, 40, 80, 100, and 200 ng per PCR reaction performed with either Ogre F/R or Cyclop F/R primer combinations in 3 different pea varieties. No differences in banding pattern were observed in this 20-fold concentration scale, indicating the robustness of the method (Figure 2). Additionally, dependence on the DNA isolation method was tested for 2 commercially available kits (Invitek and Qiagen) and 2 commonly used

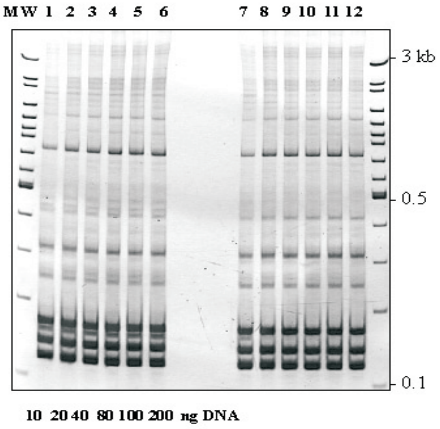

Figure 2. Testing of IRAP dependence on template DNA concentration. IRAP analysis with *Cyclop* F and R primers. After isolation by CTAB protocol, genomic DNA (10, 20, 40, 80, 100, 200 ng in 2 μL) was added per 20 μL of PCR reaction volume. Results for cvs. Tolar (1–6) and Gotik (7–12) are shown. MW = 100-bp DNA molecular weight marker (Fermentas).

classical isolation protocols. CTAB/chloroform by Doyle and Doyle (1990) and salt-acidic by Guillemaut and Marechal-Drouard (1992) were used. Isolated genomic DNA was diluted to an identical concentration of 50 ng μL^{-1}, of which 2 μL were used per PCR reaction performed with *Cyclop* and *Ogre* F/R primer combinations. Some differences were found, especially for DNA fragments of larger size and minor intensity (Figure 3). However, the dominant strong bands used for

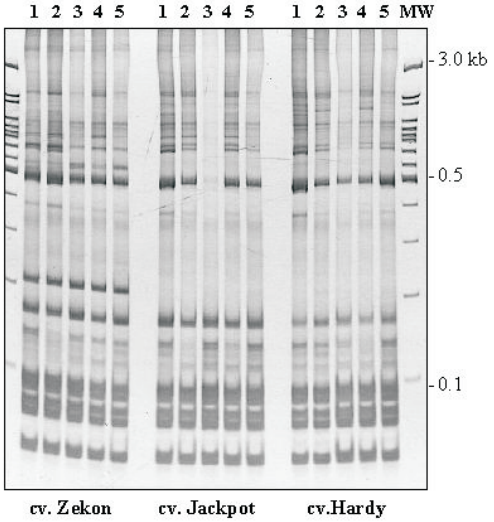

Figure 3. Effect of DNA isolation method – Qiagen DNeasy kit (1), Invitek (2), salt-acidic (3) and CTAB protocol (4, 5) – on performance of IRAP *Ogre* F and R primers. In each case, 50 ng of DNA was taken in 2 μL for the PCR reaction. MW = 100-bp DNA molecular weight marker (Fermentas).

polymorphism scoring remained identical in all four methods.

Taq polymerase source dependence

In studies of multiplex/multi-loci PCR techniques, some variation of the banding profile in relation to *Taq* polymerase source was observed (Schulman et al. 2004; Weissing et al. 2005). Therefore, in the process of method setup, several *Taq* polymerase sources for performance in the IRAP method have been tested. Appropriate buffers provided by the supplier were used, as well as identical primer and *Taq* polymerase concentrations and PCR amplification conditions, the Mg concentration was set at 2.5 mM. Four different *Taq* polymerases (TAKARA, NEB, Fermentas, Dynazine, BioTools) were tested on four pea genotypes. Significant differences were reproducibly found. In general, Dynazine and Fermentas enzymes performed well over smaller-size DNA fragments, while fragments longer than 500 bp were under-represented (Figure 4). The opposite situation was found for BioTools and TAKARA polymerases. Considering unit costs and performance, BioTools *Taq* polymerase was selected for further analyses. This demonstrates that in order to be able to compare properly different independent analyses, care has to be taken for *Taq* polymerase selection.

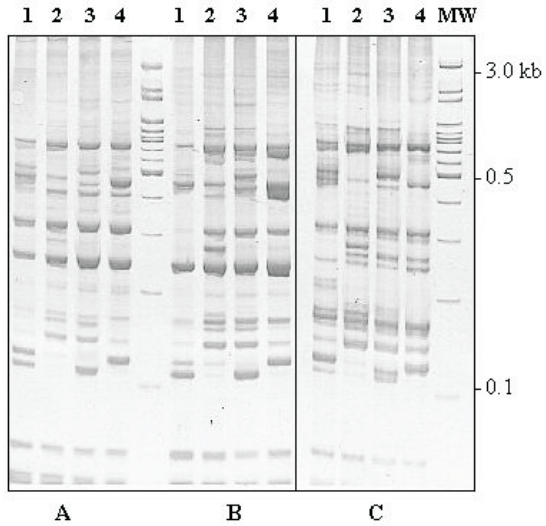

Figure 4. Effect of *Taq* polymerase source – TAKARA polymerase (A), BioTools (B) and Dynazine II, Finzymes (C) – on performance of IRAP *Cyclop* F and R primers. In each case, 50 ng of DNA and 2.5 mM of Mg ions were used for the PCR reaction with cvs. Merkur (1), Garde (2), Tempra (3) or Kamelot (4). MW = 100-bp DNA molecular weight marker (Fermentas).

Mg concentration

Another issue of great importance in all multiplex PCR amplifications is the concentration of magnesium ions, an important cofactor for *Taq* polymerase (Innis et al. 1995). Since genomic DNA is usually kept for better stability in TE buffer (containing EDTA, a chelator of Mg), sufficient magnesium has to be provided. In addition, a high Mg concentration influenced primer annealing, allowing undesirable mismatches. The concentration range of 1.0–1.5–2.0–2.5–3.0 mM $MgCl_2$ was tested with BioTools, Fermentas and TAKARA polymerases. In all 3 cases repeatedly the concentration of 2.5 mM was found optimal in terms of DNA fragment profile and reproducibility (Figure 5).

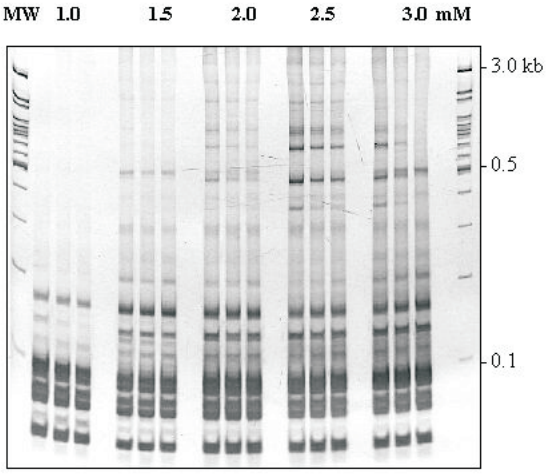

Figure 5. Optimization of Mg concentration tested on *Taq* polymerase (Fermentas). Three independent DNA isolations of 50 ng per PCR reaction from cv. Bohatyr were tested on IRAP with *Ogre* F/R primer combinations with indicated $MgCl_2$ concentrations. MW = 100-bp DNA molecular weight marker (Fermentas).

Reproducibility and robustness

Any DNA fingerprinting method used for genotype identification has to fulfil sufficiently high reproducibility criteria. This issue was addressed by analysis of 3 pea varieties, each done in 6 independent CTAB/chloroform DNA preparations from the same individual plant. IRAP *Cyclop* and *Ogre* F/R analyses were performed in triplicate within 1 day or consecutively. PCR amplification products were analysed simultaneously on PAGE. In all cases no differences were found in major dominant bands used for polymorphism scoring. The only variation was encountered in 100–200 bp and in larger DNA fragments (over 800 bp), parallel to intra-accession variation (Figure 6). Even though a proper round test was not carried out, the method proved to be transferable and is currently employed for rapid pea fingerprinting by 2 additional labs.

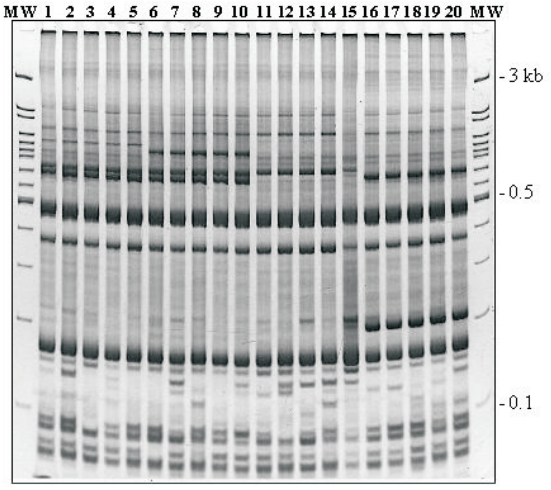

Figure 6. Intra-accession variation/stability of IRAP performed with F and R primers of the *Ogre* retrotransposon. Individual plant DNA isolation and 8% PAGE analysis, cvs. Garde (1–5), Grana (6–10), Menhir (11–15) and Zekon (16–20). MW = 100 bp DNA marker (Fermentas).

Intra-accession variation

Eight pea varieties were selected for close investigation of intra-accession variation. Five individual plants per variety were sampled, and further cultivated to enable seed setting. The subsequent generation was sown in controlled greenhouse conditions. After genomic DNA isolation, IRAP with *Cyclop* and *Ogre* F/R primer combinations was performed on all samples. No alteration of the major scorable band pattern was observed among 5 individual plants of the given variety (Figure 6) and in 2 subsequent generations (in total, 10 progeny plants were assessed, e.g. 50 plants per accession). The only detected variation was in minor low-intensity fragments of less than 150 bp and more than 800 bp, corresponding to the reproducibility level of the IRAP technique. This demonstrates both the robustness and suitability of IRAP for genotype fingerprinting together with minimal intra-accession variation in the largely self-pollinating and therefore highly homozygous plants.

DNA fingerprinting of pea varieties

To demonstrate the suitability of IRAP for efficient DNA fingerprinting in pea, 33 commercially grown varieties were selected from the entire AGRITEC collection of pea germplasm (consisting of 1300 accessions) and were analysed with *Cyclop* and *Ogre* forward and reverse LTR primers. The PCR composition and amplification conditions were set at previously determined optimal values (e.g. 50 ng template DNA, 2.5 mM Mg concentration, 1 U of BioTools *Taq* polymerase and 35 cycles). The products were resolved on 8% TBE-PAGE and run at 30 mA for 3 h. After binary code scoring, genetic similarity matrices were calculated by using Nei-Li coefficient and UPGMA analyses performed on NTSYS-pc software were used to generate dendrograms (Figure **7**). Even

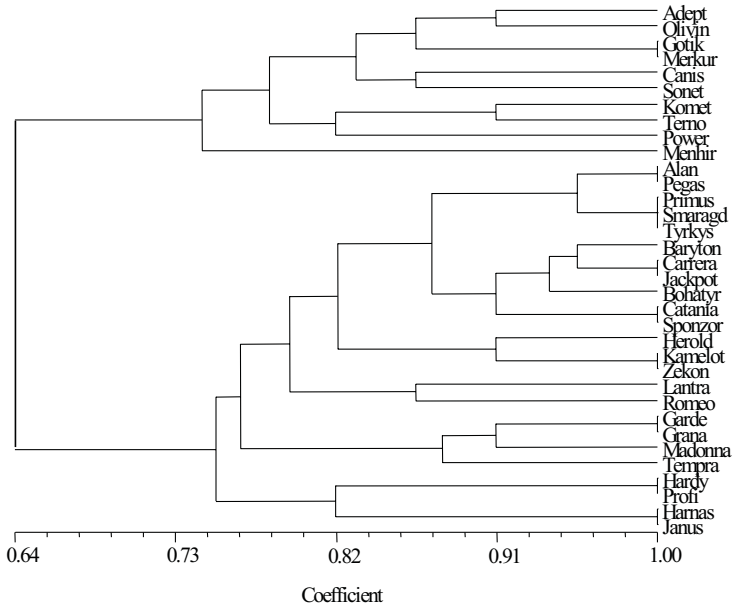

Figure 7. Cluster analysis of 33 pea cultivars currently registered in the Czech Republic, based on genetic similarity coefficient (Dice or Nei-Li) computed by NTSYS-pc software (UPGMA = unweighted pair-group method, arithmetic average) with summary data from IRAP with *Ogre* and *Cyclop* F/R primer combinations.

though the pea varieties used were of relatively close genetic relationship (6 have cv. Bohatyr as parent), a single *Cyclop* F/R IRAP analysis has provided clear resolution, with 13 varieties being uniquely identified and the others being clustered in 4 groups with 2–4 members each. Clusters reflected the pedigree relationship, as demonstrated by cvs. Herold, Kamelot and Zekon with the same parents. When data for *Ogre* IRAP were added, resolution improved even further, with 15 varieties unequivocally identified and the others being clustered in 9 groups with 2–3 members each. These two PCR analyses have produced altogether 11 scorable major bands for *Cyclop* and 8 for *Ogre*, of which 7 for *Cyclop* (64%) and 5 for *Ogre* (62%) were polymorphic and informative. A comparable level of resolution is achieved when 11 individual SSR loci are used.

Conclusions

The IRAP method has recently been exploited to study biodiversity and phylogeny in the genera *Brassica* (Tatout et al. 1999), *Hordeum* (Waugh et al. 1997; Kalendar et al. 1999), *Oryza* (Iwamoto et al. 1999), and *Spartina* (Baumel et al. 2002). Additionally, it proved to be suitable for studies of genome evolution (Kalendar et al. 2000; Vicient et al. 2001), and in gene mapping of barley (Manninen et al. 2000) and wheat (Boyke et al. 2002). In the latter 2 studies, good coverage of the whole genome was demonstrated. At this moment microsatellite SSR markers are the method of first choice to complement the DUS (distinctness, uniformity and stability) testing procedure (McCouch et al. 1997; Becher et al. 2000). The advantages of SSR markers are: high polymorphism, reproducibility, co-dominance, multiallelic type of variation, and possibility to map. Especially the possibility to map is very important for mapping of the genome region carrying the desired loci.

However, the development of a SSR marker system for new crops is highly time- and money-consuming, since specific sequence information is needed. Additionally, testing of numerous SSR loci is required in order to gain enough information. When speed and simplicity are of priority, then multilocus techniques (such as AFLP) can be chosen. This paper demonstrates the development and utility of the relatively simple but still highly specific and accurate IRAP fingerprinting to assess pea cultivar fidelity. The major advantage of IRAP over the other, more accurate methods readily applicable in pea (SSR, AFLP and retrotransposon-based SSAP and RBIP) is the high information content gained per single PCR analysis, which substantially cuts time and cost. The most efficient was IRAP performed with *Cyclop* and *Ogre* retrotransposons, but there is still a potential to improve and use other elements in both IRAP and the more demanding SSAP format. Certain improvement could be achieved by manipulating LTR primer position (Kalendar et al. 1999, and data not shown). Especially application of inter-MITE IRAP would be of practical interest since target size preference near or within genes makes them a good tool for marker-assisted breeding (Chang et al. 2001). Since in our approach a consensus-sequence-derived primer was used, the complex problematically scorable banding pattern could be reduced by a more specific primer design. Resolution can be further manipulated by acrylamide gel concentration and electrophoresis running time. Analysis can be also performed on simple horizontal agarose gel, allowing higher sample through-put but with lower resolution (data not shown).

Due to dominance and possibly non- representative genome coverage, the IRAP method would not be suitable for applications in marker-assisted selection, even though it is used in cereals together with other methods. The major application field would be for breeders to test quickly fidelity of seeds, crosses and breeding material, even though not currently supported by UPOV. The simplicity and robustness guarantee good application potential.

Acknowledgements. This work was supported by the Czech Ministry of Education (project MSMT 2678424601). Dr. J. Macas' sharing of pre-publication data and excellent technical support of Ms. L. Vítámvásová and Ms. E. Fialová are gratefully acknowledged.

REFERENCES

Baranger A, Aubert G, Arnau G, Lainé AL, Deniot G, Potier J, et al. 2004. Genetic diversity within *Pisum sativum* using protein- and PCR-based markers. Theor Appl Genet 108: 1309–1321.

Batzer MA, Deininger PL, 2002. *Alu* repeats and human genomic diversity. Nature Rev Genet 3: 370–379.

Baumel A, Ainouche M, Kalendar R, Schulman AH, 2002. Inter-retrotransposon amplified polymorphism (IRAP), and retrotransposon-microsatellite amplified polymorphism (REMAP) in populations of the young allopolyploid species *Spartina angel-*

ica Hubbard (Poaceae). Mol Biol Evol 19: 1218–1227.

Becher SA, Steinmetz K, Weising K, Boury S, Peltier D, 2000. Microsatellites for cultivar identification in *Pelargonium*. Theor Appl Genet 101: 643–651.

Boyko E, Kalendar R, Korzun V, Gill B, Schulman AH, 2002. Combined mapping of *Aegilops tauschii* by retrotransposon, microsatellite, and gene markers. Plant Mol Biol 48: 767–790.

Burstin J, Deniot G, Potier J, Weinachter C, Aubert G, Baranger A, 2001. Microsatellite polymorphism in *Pisum sativum*. Plant Breeding 120: 311–317.

Cooke RJ, 1999. Modern methods for cultivar verification and the transgenic plant challenge. Seed Sci and Technol 27: 669–680.

Doyle JJ, Doyle JL, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19: 11–15.

Ellis THN, Poyser SJ, Knox MR, Vershinin AV, Ambrose MJ, 1998. Polymorphism of insertion sites of *Ty1-copia* class retrotransposons and its use for linkage and diversity analysis in pea. Mol Gen Genet 260: 9–19.

Flavell AJ, Bolshakov VN, Booth A, Jing R, Russell J, Ellis THN, Isaac P, 2003. A microarray-based high throughput molecular marker genotyping method: the tagged microarray marker (TAM) approach. Nucleic Acids Res 31: e115.

Flavell AJ, Knox MR, Pearce SR, Ellis THN, 1998. Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. The Plant J 16: 643–650.

Ford R, Le Roux K, Itman C, Brouwer JB, Taylor PWJ, 2002. Diversity analysis and genotyping in *Pisum* with sequence tagged microsatellite site (STMS) primers. Euphytica 124: 397–405.

Guillemaut P, Marechal-Drouard L, 1992. Isolation of plant DNA: A fast, inexpensive and reliable method. Plant Mol Biol Rep 10: 60–65.

Chang RY, O´Donoughue LS, Bureau TE, 2001. Inter-MITE polymorphism (IMP): a high throughput transposon-based genome mapping and fingerprinting approach. Theor Appl Genet 102: 773–781.

Chavanne F, Zhang DX, Liaud MF, Cerff R, 1998. Structure and evolution of *Cyclops*: a novel giant retrotransposon of the *Ty3/Gypsy* family highly amplified in pea and other legume species. Plant Mol Biol 37: 363–375.

Innis MA, Gelfand DH, Sninsky JJ, 1995. PCR Strategies. Academic Press.

Iwamoto M, Nagashima H, Nagamine T, Higo T, Higo K, 1999. p-SINE-like intron of the CatA catalase homologs and phylogenetic relationships among *Aagenome, Oryza* and related species. Theor Appl Genet 98: 853–861.

Jing RC, Knox MR, Lee JM, Vershinin AV, Ambrose M, Ellis THN, Flavell AJ, 2005. Insertional polymorphism and antiquity of PDR1 retrotransposon insertions in *Pisum* species. Genetics 171: 741–752.

Kalendar R, Grob T, Regina M, Suoniemi A, Schulam A, 1999. IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. Theor Appl Genet 98: 704–711.

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH, 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proc Natl Acad Sci USA 97: 6603–6607.

Kazazian HH, 2004. Mobile elements: drivers of genome evolution. Science 303: 1626–1632.

Kumar A, Bennetzen JL, 1999. Plant retrotransposons. Annu Rev Genet 33: 479–532.

Loridon K, McPhee K, Morin J, Dubreuil P, Pilet-Nayel ML, Aubert G, et al. 2005. Microsatellite marker polymorphism and mapping in pea (*Pisum sativum* L.). Theor Appl Genet 111: 1022–1031.

Macas J, Koblizkova A, Neumann P, 2005. Characterization of Stowaway MITEs in pea (*Pisum sativum* L.) and identification of their potential master elements. Genome 48: 831–839.

Manninen O, Kalendar R, Robinson J, Schulman AH, 2000. Application of BAR/M retrotransposon markers to the mapping of a major resistance gene for net blotch in barley. Mol Gen Genet 264: 325–334.

McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, 1997. Microsatellite marker development, mapping and application in rice genetics and breeding. Plant Mol Biol 35: 89–99.

Morgante M, Olivieri AM, 1993. PCR-amplified microsatellites as markers in plant genetics. The Plant J 3: 175–182.

Nei M, Li WH, 1979. Mathematical model for studing genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76: 5269–5273.

Neumann P, Nouzova M, Macas J, 2001. Molecular and cytogenetic analysis of repetitive DNA in pea (*Pisum sativum* L.). Genome 44: 716–728.

Neumann P, Pozarkova D, Koblizkova A, Macas J, 2005. PIGY, a new plant envelope-class LTR retrotransposon. Mol Genet Genomics 273: 43–53.

Neumann P, Pozarkova D, Macas J, 2003. Highly abundant pea LTR retrotransposon *Ogre* is constitutively transcribed and partially spliced. Plant Mol Biol 53: 399–410.

Pearce SR, Knox M, Ellis THN, Flavell AJ, Kumar A, 2000. Pea *Ty1-copia* group retrotransposons: transpositional activity and use as markers to study genetic diversity in *Pisum*. Mol Gen Genet 263: 898–907.

Rohlf M, 2000. NTSYS-pc: Numerical taxonomy and multivariate analysis system. Version 2.1, Department of Ecology and Evolution, State University of New York, USA.

Samec P, Posvec J, Stejskal V, Nasinec V, Griga M, 1998. Cultivar identification and relationships in *Pisum sativum* L. based on RAPD and isozymes. Biol Plant 41: 39–48.

Shedlock AM, Okada N, 2000. SINE insertions: powerful tools for molecular systematics. BioEssays 22: 148–160.

Schulman AH, Flavell AJ, Ellis THN, 2004. The application of LTR retrotransposons as molecular markers in plants. In: Method in Molecular Biology 260: 145–175.

Simioniuc D, Uptmoor R, Friedt W, Ordon F, 2002. Genetic diversity and relationships among pea cultivars revealed by RAPDs and AFLPs. Plant Breeding 121: 429–435.

Tar´an B, Zhang C, Warkentin T, Tullu A, Vandenberg A, 2005. Genetic diversity among varieties and wild species accessions of pea (*Pisum sativum* L.) based on molecular markers, and morphological and physiological characters. Genome 48: 257–272.

Tatout C, Warwick S, Lenoir A, Deragon JM, 1999. SINE insertions as clade markers for wild crucifer species. Mol Biol Evol 16: 1614–1621.

UPOV-BMT: BMT/36/10. 2002. Progress report of the 36th session of the technical committee, the technical working parties and working group on biochemical and molecular techniques and DNA-profiling in particular. Geneva.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407–4414.

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas WT, Powell W, 1997. Genetic distribution of Bare-1 like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphism (S-SAP). Mol Gen Genet 253: 687–694.

Weising K, Nybom H, Wolf K, Kahl G, eds. 2005. DNA fingerprinting in plants: Principles, methods and applications 2nd edn. CRC Press Taylor and Francis Group, Boca Raton, USA.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV, 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18: 6531–6535.

Zietkiewitz E, Rafalski A, Labuda D, 1989. Genome fingerprinting by simple sequence repeat (SSR) anchored polymerase chain reaction amplification. Genomics 20: 176–173.