COBISS Code 1.01

DOI: 10.2478/v10014-011-0011-y
Agrovoc descriptors: pisum sativum, transgenic plants, genotypes, peas, fungal diseases, plant diseases, ascochyta, rubus idaeus, resistance to injurious factors, agrobacterium, PCR

Agris category code: H20, F30

PCR identification of *rpgip1* transgene in *Pisum sativum* L.

Kornelia POLOK[1] and Hans-Jörg JACOBSEN[2]

Received July 30, 2011; accepted September 14, 2011.
Delo je prispelo 30. julija 2011, sprejeto 14. septembra 2011.

ABSTRACT

Recent efforts to increase Ascochyta blight resistance of pea have focused on the introduction of foreign genes by genetic engineering. The *rpgip1* gene from *Rubus idaeus* was introduced by *Agrobacterium*-mediated transformation into *Pisum sativum*, cv. Baroness with the aim to increase pea resistance to fungal diseases. Notwithstanding this success, practical applications have to be preceded by the development of analytical methods for screening. Singleplex and multiplex PCR assays were employed to test primer efficiency in identifying the *rpgip1* transgene in 11 pea genotypes. Five from ten primer combinations were effective in identifying transgene or insert sequences. PCR amplification using five other primer pairs revealed unspecific amplicons. According to *in silico* analyses, they arose from retrotransposons and pea genes including homologues of *rpgip1*. Two sets of primers were prepared with the aim of simultaneous amplification of different *rpgip1* fragments. Fingerprints were sums of bands observed from individual pairs so the utility of multiplex assays was demonstrated. An additional advantage of multiplex PCR was clear differentiation between the transgene and endogenous *pgip* genes present in the donor species, *R. idaeus*. Sequencing of two PCR products confirms that no substantial rearrangements at the *rpgip1* transgene arose during development of transgenic plants. However, a deletion occurred at 59 bp in the PGIP+VST line and a substitution at 392 bp in the PGIP line. The frequency of point mutations was not high (1.1×10^{-3}) and comparable with the frequency expected for host genes based on the neutral theory of molecular evolution.

Key words: Transgenic pea, fungal diseases, *Rubus idaeus*, *pgip* homologues, multiplex PCR

IZVLEČEK

PCR IDENTIFIKACIJA TRANSGENA *rpgip1* PRI GRAHU (*Pisum sativum* L.)

Novejši dosežki pri povečanju odpornosti graha na Ascochyta so povezani z uvajanjem tujih genov s pomočjo genskega inženiringa. Gen *rpgip1* iz malinjaka (*Rubus idaeus*) je vključen v grah, cv. Baroness, s transformacijo z bakterijo *Agrobacterium*, da bi se povečalo odpornost graha na to glivično bolezen. Pred praktično uporabo te metode je potrebno razviti načine za spremljanje dedovanja tega transgena. Enojna in multipleks PCR sta bili uporabljeni za testiranje učinkovitosti začetnikov in za identificiranje transgena *rpgip1* pri 11 genotipih graha. Pet od desetih začetnikov je bilo uporabnih za identifikacijo transgenov ali za vključevanje sekvenc. PCR namnoževanje z drugimi petimi začetniki je dalo nespecifične namnožke. Glede na *in silico* analize so ti nastali zaradi retrotranspozonov in grahovih genov, ki vključujejo homologe *rpgip1*. Dva seta začetnikov sta bila pripravljena za istočasno namnoževanje različnih odlomkov *rpgip1*. Elektroferogrami so bili vsote črt individualnih parov, tako je prikazana uporabnost multipleksnega poskusa. Dodatna prednost multipleksnega PCR je razločna diferenciacija med transgenom in genom *pgip* prisotnim v donorski vrsti *R. idaeus*. Sekvenciranje dveh PCR produktov potrjuje, da ni pri *rpgip1* bistvenega prerazporejanja tekom razvoja transgenih rastlin. Toda pojavila se je delecija pri 59 bp v liniji PGIP+VST in substitucija pri 392 bp v liniji PGIP. Relativna pogostnost točkovnih mutacij ni bila visoka (1.1×10^{-3}) in je bila primerljiva z pogostnostjo pri gostiteljivih genih, glede na nevtralno teorijo molekulske evolucije.

Ključne besede: Transgeni grah, glivične bolezni, *Rubus idaeus*, homologi *pgip*, multipleksna PCR

[1] Department of Genetics, University of Warmia and Mazury in Olsztyn, Plac Lodzki 3, 10-967 Olsztyn, Poland, kpolok@moskit.uwm.edu.pl

[2] Plant Biotech Unit, Gottfried Wilhelm Leibniz Universität Hannover, Herrenhauserstr 2, D30-419 Hannover, Germany

1 INTRODUCTION

The pea (*Pisum sativum* L.) has been grown since the prehistoric ages providing forage (field peas), fresh vegetable (market peas) and the material for canning and freezing (vining peas). With approximately 26% - 33% proteins and a lower level of protein inhibitors than in soybean, peas are excellent protein supplements in human and animal diets (Cousin, 1997). Producing nearly two million tons of peas per year, Europe has recently provided one-fifth of world production. However, the yield is unstable with the average yield per hectare ranging from 1.3 tons to 4.8 tons (Eurostat, 2011). Consequently, many characters have to be improved and Ascochyta blight is the major factor limiting pea production up to 75% (McDonald and Peck, 2009). Symptoms include dark-brown lesions or flecks on all plant parts, seedling blight and foot rot. The disease is caused by related fungal pathogens of the Ascochyta complex (*Mycosphaerella pinodes, Ascochyta pisi, Phoma medicaginis*).

Pea genotypes differ in susceptibilities to Ascochyta pathogens but complete resistance to infection has not been observed (Timmerman-Vaughan *et al*., 2002). On the other hand, disease resistance is often encoded by major plus minor genes. Numerous genomic regions responsible for partial resistance have been located on linkage groups II, III, IV, V and VII (Timmerman-Vaughan *et al*., 2002; Prioul-Gervais *et al*., 2007). Unfortunately, the biological function of the underlying factors still remains unknown. The quantitative inheritance, the lack of information on the molecular mechanism underlying resistance and genotype-environment interaction emphasized by variation in disease development with both growing areas and climate conditions make the traditional resistant breeding difficult. Therefore, recent efforts to increase Ascochyta blight resistance of pea, directly correlated with the yield and quality of seeds, have focused on the introduction of foreign genes by genetic engineering.

A classic tactic for producing transgenic plants with increased resistance is to introduce *R* genes, products of which recognise pathogen *Avr* genes' determinants. When corresponding *R* and *Avr* genes are present, the disease resistance is brought in the classic gene-to gene manner. Good examples of this system are plant *pgip* genes encoding polygalacturonase inhibiting proteins (PGIPs) and fungal *PG* genes responsible for fungal endopolygalacturonases (PGs) - enzymes degrading polysaccharides of the cell wall. The PGIPs inhibit fungal endo-PGs, thus preventing the hydrolysis of the α-1,4 glycosidic bonds (Shanmugam, 2005; Di Matteo *et al*., 2006). A range of crops including tomato (Powell *et al*., 2000), apple (Szankowski *et al*., 2003), and wheat (Janni *et al*., 2008) were transformed with *pgip* genes to improve plant defence against fungal pathogens. In a case of pea, the *rpgip1* gene from red raspberry, *Rubus idaeus* (Accession N°AJ620336) was introduced by *Agrobacterium*-mediated transformation into cv. Baroness (Richter *et al*., 2006). Stable inheritance was confirmed and transgenic lines showed significant inhibitory effects on the polygalacturonases of *Colletotrichum lupini* and *Stenocarpella maydis* in a greenhouse experiment. Notwithstanding the usefulness of the transgenic approach to produce pea resistant to fungal diseases, practical applications have to be preceded by the development of analytical methods for screening. They are necessary to allow consumers a free choice between genetically modified (GM) and traditional crops. Controlling the structural integrity of the transgene during further manipulations is not less important. It is a norm that unexpected variation is found in transgenic lines during later experiments or commercial use. For instance, expression of *rpgip1* varied greatly among both transgenic pea individuals and subsequent generations (Richter *et al*., 2006). Among many reasons for this phenomenon, rearrangements at a transgene locus are frequently mentioned (Morino *et al*., 1999; Svitashev *et al*., 2002).

Direct sequence analysis is the most reliable way to identify transgenes and study their structure. However, it is too time and cost consuming when a large number of samples have to be tested. For this reason, various types of polymerase chain reaction (PCR) are adopted to enable reliable and rapid assessments of GMOs. Event-specific methods, which are preferred in EU, are based on a sequence unique to a certain GMO. Such assays have to be developed for each new GMO, so do for transgenic peas expressing antifungal genes (e.g., *rpgip1*). A critical point is that PCR should differentiate between a transgene and homologous, endogenous sequences. At least two pea sequences homologous to *rpgip1* are deposited in the GenBank maintained by the National Centre for Biotechnology Information (NCBI). Both genes can be a putative template for unintended amplification driven by primers specific to *rpgip1*. One possibility to overcome the problem is to design a set of primers to track different gene fragments. A simpler but yet a reliable method to differentiate between a transgene and homologues can employ several primers at once in so called multiplex PCR.

In the present studies a set of primers was used to identify the *rpgip1* transgene in different genotypes of *P. sativum*. The primer efficiency in differentiating between the transgene and pea homologues was assessed as well as sets for multiplex PCR were proposed. We tried to understand why some primers preferentially amplified pea homologues instead of the *rpgip1* transgene. Results were discussed in the light of possible rearrangements of the transgene.

2 MATERIALS AND METHODS

2.1 Plant material

Two transgenic lines: PGIP – a homozygous line carrying two copies of the *rpgip1* gene from *Rubus idaeus*. PGIP+VST – a homozygous line with stacked antifungal genes, *rpgip1* and *Vst1* encoding stilbene synthase in *Vitis vinifera*. It was obtained by reciprocal crosses of the PGIP line expressing *rpgip1* and a line expressing the *Vst1* gene. Plants used in the analysis represented F_8 and further generations (Richter *et al*., 2006). In total 15 plants per both lines were used in the analyses.

Six *rpgip* hemizygotes: F_1 hybrids derived from crosses between the PGIP or PGIP+VST lines and the parent cultivar Baroness (Bar x PGIP and Bar x PGIP+VST), the *arthritic* mutant from Paloma (*arth* x PGIP), the *ramosus* mutant from Parvus (PGIP x *ram* and PGIP+VST x *ram*) and the Polish cultivar Sokolik (PGIP+VST x Sok). All hybrids were expected to carry at least one copy of *rpgip1*.

Negative controls: Baroness – parent cultivar, from which transgenic lines were derived (15 plants in total); Bar x VST – the F_1 between Baroness and the transgenic line with *Vst1* gene; *ram* x Bar – the F_1 between the *ramosus* mutant from Parvus and Baroness.

Positive controls: pSCP1 plasmid – a binary vector using for pea transformation and carrying a *nos* promoter-driven *bar* gene and a double 35S promoter-driven the *rpgip1* gene from *R. idaeus* (Richter *et al*., 2006). The *rpgip1* transgene sequence consists of 996 base pairs (bp) and corresponds to the 29 bp - 1024 bp of *rpgip1* in the NCBI file, accession N°AJ6200336. Moreover, a wild ecotype of red raspberry (*R. idaeus*) was used for comparisons between transgenic and donor species fingerprints.

2.2 DNA isolation

Plant genomic DNA was extracted from about 1 g of young leaves by the modified CTAB procedure (Polok, 2007). The quality of DNA was verified on 1% agarose gels while the purity was assessed spectrophotometrically and it ranged between 93% and 99%. The DNA content of the samples ranged from 177 µg to 568 µg. The DNA of pSCP1 plasmid was provided by the Plant Biotech Unit, Leibniz Universität Hannover.

2.3 Primers

Eight primers distributed over the entire transgene were designed on the *rpgip1* sequence used in the pSCP1 transformation vector. Numbers in primers' abbreviations identify their 5'position on the *rpgip1* template. **Forward primers:** 1F: 5'atgatggacttcaagctctt3'; 6F: 5'ggacttcaagctcttctccc3'; 108F: 5'caagacagccttcaacaaccc3'; 421F: 5'cagctcaagaacctcacatt3'. **Reverse primers:** 366R: 5' cttgagatgtttaagcttgg3'; 733R: 5'ccacaatctgggtggtcttgt3'; 958R: 5'ggttatggaaatacgacgtg3', 971R: 5' gcaacttgggaggggagcac3'. Ten pairwise combinations of forward and reverse primers and two multiplex sets with four and three primers were applied (Table 1).

Table 1. Primer combinations and PCR conditions.

	Primer combination	Annealing temperature [°C]	Number of cycles [n]	Predicted product length [bp]
Singleplex	1F-366R	58	30	366
	1F-733R	58	30	733
	1F-958R	58	30	958
	108F-366R	Touchdown: 48-0.8 then 40	10+20	259
	108F-733R	48	30	626
	108F-958R	58	30	851
	421F-366R	40	30	942*
	421F-733R	50	30	313
	421F-958R	58	35	538
	6F-971R	41	35	958, 1700*
Multiplex	**Set 1:**			
	1F-366R + 733R + 958R	58	35	366, 733, 958
	Set 2:			
	108F + 421F-958R	58	35	538, 851

*Product predicted for the plasmid and only when circular matrix is assumed.

2.4 PCR conditions

Singleplex PCR: PCR was performed in a 20 µl volume containing 20 mM $(NH_4)_2SO_4$, 50 mM Tris-HCl, pH 9.0 at 25°C, 1.5 mM $MgCl_2$, 2 µl of Enhancer with betaine (Epicentre Technology), 200 µM dNTPs, 1.0 µM primers, 0.75 U of *Tfl* polymerase (Epicentre Technology) and 80 ng of template DNA. The standard thermal conditions were: 94°C for 3 min, followed by 30 cycles of 94°C for 1 min, 58°C for 1 min, 72°C for 1 min. The final extension was at 72°C for 5 min. However, conditions for five pairs were optimized and finally a number of cycles were increased in two cases while annealing temperature was modified in three ones (Table 1). PCR products were loaded on 1.5% (w/v) agarose gels containing 0.5 µg/ml ethidium bromide, separated in 1 x TBE buffer (Tris-Borate-EDTA) at 100 V constant power, visualized under UV light (312 nm), photographed with Olympus Camera and stored as .jpg files. **Multiplex PCR:** The same reagent concentrations and standard thermal conditions were applied as for singleplex PCR with the exception of a number of cycles that was increased to 35.

2.5 Sequencing

PCR products revealed by primers 1F-733R and 1F-958R in transgenic plants and the pSCP1 plasmid were sequenced. The PCR products were separated on 1.5% agarose gels, excised and purified using the gel-out system (AKOR Laboratories). The identities of the products were confirmed by nested PCR and restriction analyses. Sequencing was performed using the ABI3730 and BigDye Terminator Ready Cycle Sequencing Kit (Applied Biosystem) with the 1F primer by oligo.pl, Institute of Biochemistry and Biophisics PAS.

2.6 *In silico* and sequence analyses

The specificity of primers was tested by Primer-BLAST against *Pisum* sequences deposited in the non-redundant database at NCBI. Possibilities of unintended amplification were studied by performing *in silico* PCR on the *rpgip1* template and two pea homologues deposited at NCBI (AB0877839, AJ749705) using FastPCR software. Multiple alignment algorithms in CLUSTALX2 were applied for sequence comparisons. Sequences were viewed in Jalview 2.6.1 editor.

3 RESULTS

3.1 Amplification of the *rpgip1* transgene on the pSCP1 template

The majority of primer pairs amplified a product of a size corresponding to *rpgip1* while using the pSCP1 template (Fig. 1, 2). Thus, their usefulness in identification of the transgene was confirmed. Interestingly, only five pairs amplified a single, expected band (1F-366R, 1F-958R, 108F-366R, 108F-733R, 421F-733R), while the other five revealed two bands (1F-733R, 108F-958R, 421F-366R, 421F-958R and 6F-971R). The first of two products fell within the expected size range but the second was much shorter and faint. Surprisingly, primers 421F and 366R resulted in two, clearly visible but relatively short amplicons (350 bp, 490 bp) although the target sites' orientation rather excluded amplification (Table 1). *In silico* PCR used for quick primer analyzing on the *rpgip1* template demonstrated that some of additional, shorter products could result from unspecific amplification of the transgene. Likely, this explanation is valid for the 420 bp product revealed by 108F-958R but not for the other pairs of primers, for which even very weak searching criteria (7 mismatches, initial word size of 2) did not result in virtual, unspecific amplification. Therefore, it was assumed, that additional bands were probably derived from unintended amplification of other plasmid sequences. Pairs of primers amplifying additional plasmid sequences included 1F-733R (480 bp), 421F-958R (300 bp), 421F-366R (350 bp and 490 bp), and 6F-971R (740 bp).

3.2 Efficiency of primer combinations in identification of the *rpgip1* transgene in *P. sativum*

Amplification patterns of the pSCP1 vector demonstrated that all primer pairs could theoretically detect the transgene and eventually other sequences within an insert. Surprisingly, among ten pairs of primers checked, only 50% of combinations gave the plasmid banding pattern in transgenic PGIP and PGIP+VST plants and no amplification in the parent cultivar, Baroness (Fig. 1). This first group included 1F-366R, 1F-733R, 1F-958R, 421F-733R, 421F-958R pairs. They can further be used in identifying the transgene in various backgrounds as confirmed by fingerprints of *rpgip1* hemizygotes similar to these of transgenic lines and no amplification in negative controls (Table 2).

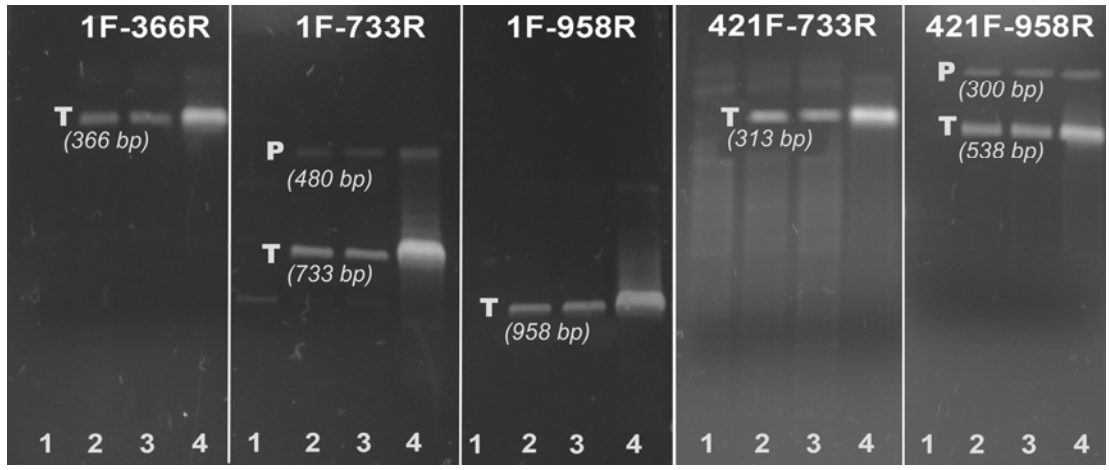

Fig 1. Identification of the *rpgip1* transgene and plasmid sequences in *P. sativum*. 1 - Baroness – negative control, 2 - PGIP, 3 - PGIP+VST, 4 - pSCP1 plasmid. T - products corresponding to the transgene, P – products corresponding to other insert sequences.

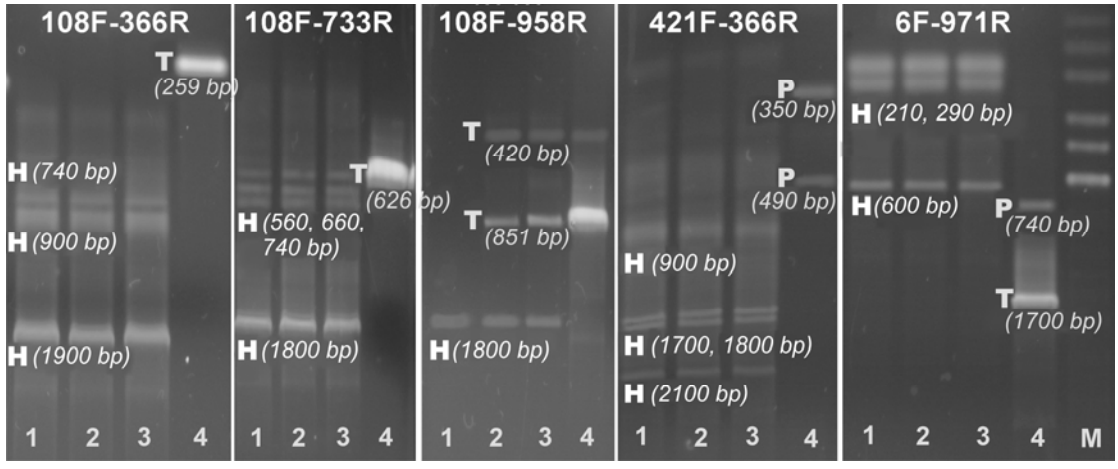

Fig 2. Unspecific amplification of *P. sativum* genomic DNA revealed by primers complementary to the *rpgip1* transgene. 1 - Baroness – negative control, 2 - PGIP, 3 - PGIP+VST, 4 - pSCP1 plasmid. T - products corresponding to the transgene, P – products corresponding to other insert sequences, H – products corresponding to endogenous pea sequences homologous to the transgene (transposons or *rpgip1* homologues), M – DNA marker.

Table 2. Validation of primer combinations in identifying the *rpgip1* transgene in *P. sativum*.

Primer combination	Estimated product length [bp]	Negative controls			Transgenic lines		*rpgip1* hemizygotes						Donor species
		Baroness	Bar x VST	ram x Bar	PGIP	PGIP+VST	Bar x PGIP	arth x PGIP	PGIP x ram	Bar x PGIP+VST	PGIP+VST x ram	PGIP+VST x Sok	*Rubus idaeus*
1 Primer pairs identifying the *rpgip1* transgene or/and plasmid sequences													
1F-366R	366	0	0	0	T	T	T	T	T	T	T	T	366
1F-733R	480	0	0	0	P	P	0	0	P	P	P	P	1000
	733	0	0	0	T	T	T	T	T	T	T	T	
1F-958R	958	0	0	0	T	T	T	T	T	T	T	T	1100
421F-733R	313	0	0	0	T	T	T	T	T	0	T	T	450
421F-958R	300	0	0	0	P	P	P	P	P	0	P	P	538, 600
	538	0	0	0	T	T	T	T	T	0	T	T	
2 Primer pairs revealed unintended amplification on pea homologous sequences													
108F-366R	740, 900, 1900	H	H	H	H	H	H	H	H	H	H	H	0
108F-733R	560, 660, 740, 1800	H	H	H	H	H	H	H	H	H	H	H	500
108F-958R	420	0	0	0	T	T	T	T	T	T	T	T	1000, 1500
	851	0	0	0	T	T	T	T	T	T	T	T	
	1800	H	H	H	H	H	H	H	H	H	H	H	
421F-366R	900, 1700, 1800, 2100	H	H	H	H	H	H	H	H	H	H	H	1000, 1800, 2100
6F-971R	210, 290, 600	H	H	H	H	H	H	H	H	H	H	H	580
3 Multiplex PCR													
Set1: 1F, 366R, 733R 958F	366	0	0	0	T	T	T	T	T	T	T	T	366, 1000, 1100
	480	0	0	0	P	P	P	P	P	P	P	P	
	733	0	0	0	T	T	T	T	T	T	T	T	
	958	0	0	0	T	T	T	T	T	T	T	T	
Set2: 108F, 421F, 958R	300	0	0	0	P	P	P	P	P	0	0	0	538, 1000, 1500
	420	0	0	0	P	P	P	P	P	0	0	0	
	538	0	0	0	T	T	T	T	T	0	T	T	
	851	0	0	0	T	T	T	T	T	T	T	T	
	1800	H	H	H	H	H	H	H	H	H	H	H	

0 – lack of amplification. T – products corresponding to the *rpgip1* transgene, P – products corresponding to plasmid/insert sequences, H – products resulted from unintended amplification of homologous pea sequences, for *R. idaeus*, an approximate size of observed products is given.

PCR amplification from pea genomic DNA using other five pairs of primers revealed completely different fingerprints than observed in the plasmid (Fig. 2). This second group comprised pairs: 108F-366R, 108F-733R, 108F-958R, 421F-366R and 6F-971R. Unexpected products were found both in transgenic and non-transgenic plants, hence providing an evidence for unspecific amplification from homologous sequences in pea genome (Table 2). All pairs but 108F-958R failed to detect either the transgene or other insert sequences. When PCR was performed with 108F-958R primers, the combination of a plasmid pattern and additional bands was noticed (Fig. 2). Hence, this pair may be in favour when a goal is to distinguish between the transgene and its endogenous homologues.

Based on the results for each pair of primers, two sets consisting of four (set 1: 1F, 366R, 733R, 958R) and three primers (set 2: 108F, 421F, 958R) were prepared with the aim of simultaneous amplification of different *rpgip1* fragments. Fingerprints obtained from these two multiplex assays were sums of bands observed from individual pairs (Fig. 3) so the utility of using several primers at once was demonstrated. Alike individual pairs, the first set amplified only sequences of a transgenic origin (Table 2). The second one gave products of both transgene/insert and endogenous origins. Likely, this pattern was caused by the 108F primer, combinations of which amplified mainly homologues. An additional advantage of multiplex PCR over singleplex is clear differentiation between the transgene and *pgip* genes present in the donor species, *R. idaeus* (Fig. 3). For example, bands revealed by 1F-366R and 421F-958R pairs in red raspberry were indistinguishable from the *rpgip1* amplicons in transgenic pea. Simultaneous amplification by 108F, 421F and 958R (set 2) gave a three-band pattern in *R. idaeus* that was clearly distinct from the transgene fingerprint.

3.3 Structure of the *rpgip1* locus in transgenic plants

The *rpgip1* transgene authenticity was confirmed by sequencing six PCR products, each two from the PGIP, PGIP+VST lines and the pSCP1 plasmid. The sequences derived from 1F-733R primers had 678 bp in transgenic plants and 691 bp in the pSCP1. Expectedly, sequences obtained from 1F-958R primers were longer and had 903 bp and 889 bp, respectively. Irrespective of the template (plant or plasmid DNA), shorter and longer nucleotide sequences were identical on the length of ca. 690 bp, what confirmed that both primer pairs identified the *rpgip1* transgene. This fact also enabled to use only longer sequences in further comparisons. Generally, significant rearrangements were not observed in the transgene (Fig. 4). However, all sequences were shorter than the original *rpgip1* construct. They started from the 47 bp position instead of 1 bp despite the 1F primer was used in sequencing reactions. The deletion of cytosine at 59 in relation to the *rpgip1* template occurred in the PGIP+VST line and thymine was substituted by adenine at 392 in the PGIP line. In total 1806 unique base pairs were sequenced in two transgenic lines, so that the frequency of point mutations was 1.1 x 10^{-3} or one mutation per 909 nucleotides.

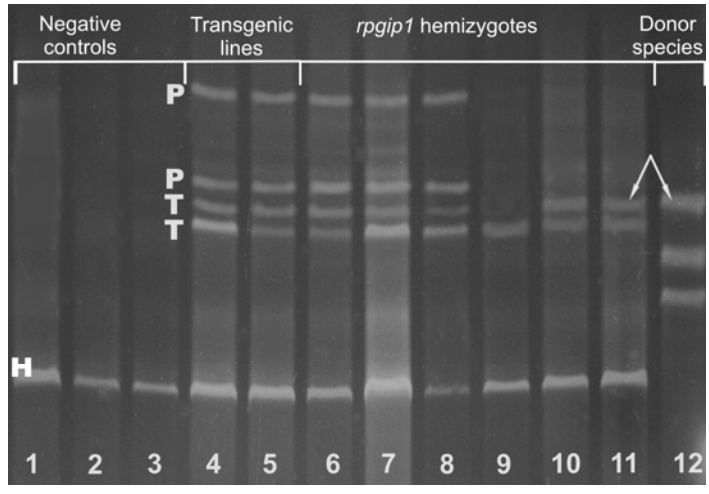

Fig. 3. Multiplex PCR profiles revealed by primers 108F, 421F and 958R (set 2). 1 - Baroness, 2 – Bar x Vst, 3 – *ram* x Bar, 4 – PGIP, 5 – PGIP+VST, 6 – Bar x PGIP, 7 – *arth* x PGIP, 8 – PGIP x *ram*, 9 – Bar x PGIP+VST, 10 – PGIP+VST x *ram*, 11 – PGIP+VST x Sok, 12 – *Rubus idaeus*. T - products corresponding to the transgene, P – products corresponding to other insert sequences, H – products corresponding to endogenous pea sequences homologous to the transgene (transposons or *rpgip1* homologues). Arrows indicate amplicons of transgenic and *R. idaeus* origin that have similar mobility.

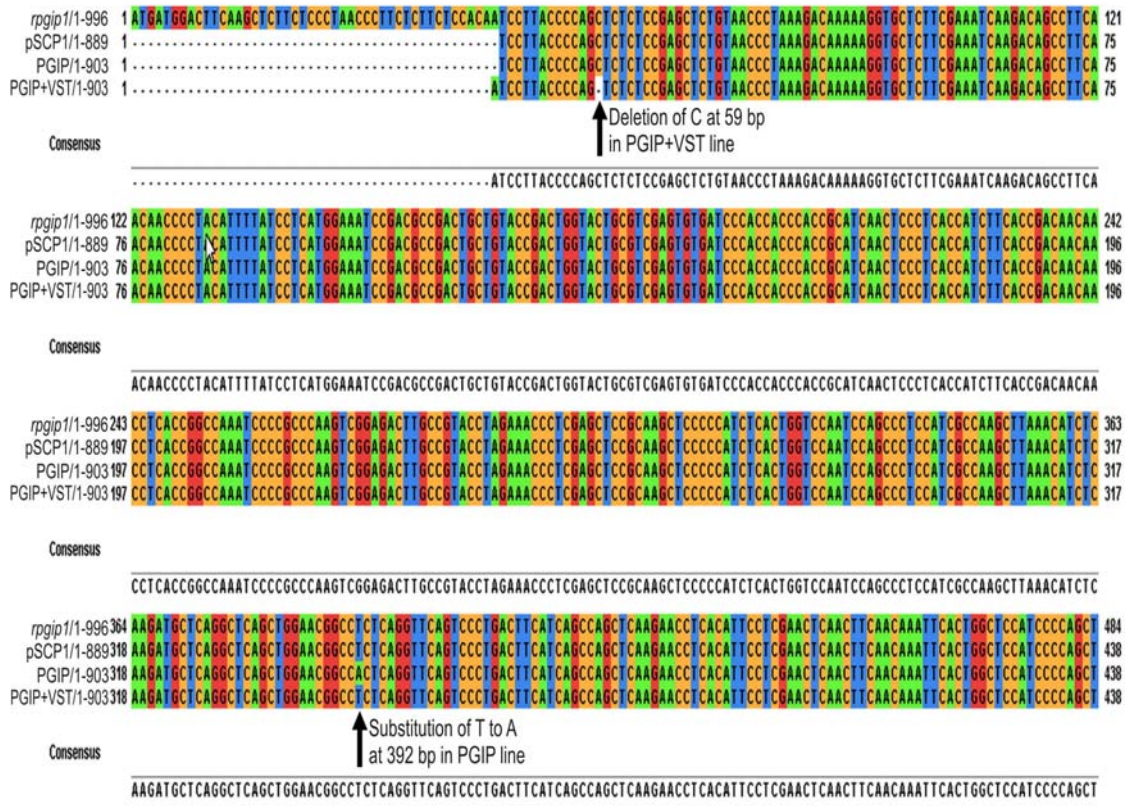

Fig. 4. Alignment of the *rpgip1* sequences obtained from the PCR product revealed by primers 1F-958R. *rpgip1*/1-996 – a sequence used for transformation, pSCP1/1-889 – a sequence from the pSCP1 plasmid, PGIP/1-903 – a sequence from the PGIP line, PGIP+VST/1-903 – a sequence from the PGIP+VST line. Arrows indicate point mutations.

3.4 Endogenous homologous sequences responsible for unintended amplification on pea genomic DNA

To find out the nature of unintended amplification by several primer pairs, the site specificity of primers was checked by performing a sequence homology search through all known template sequences for *Pisum sativum* at NCBI. Four pairs that otherwise were effective in *rpgip1* identification could also amplify other sequences from pea genome. Pairs 1F-733R, 1F-958 could identify the *Ogre* retrotransposon (AY299397.1) while 421F-733R and 421F-958R demonstrated affinity to an unknown sequence (CU655881.1) and the *Tfl* gene responsible for late flowering (AY430579.1). Among primers unsuccessful in transgene identification, two pairs, 108F-733R and 108F-958R could identify the *Psmar-2 – Mariner* retrotransposon (AY833551.1).

Because the *pgip* genes are ubiquitous in plants, unintended amplification of endogenous homologues can not be excluded. Two putative pea homologues of *rpgip1* were identified by NCBI searching, i.e., AB0877839 and AJ749705. Both sequences were derived from mRNA and they share of 65% and 60% of similarity to the *rpgip1* transgene. Alignment of both sequences and the *rpgip1* transgene at a protein level revealed several conserved regions between 164 and 331 amino acids corresponded to conserved residues at the nucleotide level. On the other hand, *in silico* PCR conducted on both pea homologues using all primer pairs complementary to *rpgip1* from *R. idaeus* produced none amplicons even though weak criteria were applied (4 - 7 mismatches allowed at the 3' end of primers).

4 DISCUSSION

The classical approach to GMO screening involves amplification of sequences common to many transgenic plants, such as the cauliflower mosaic virus 35S promoter or the *Agrobacterium tumefaciens nos* terminator (Bonfini *et al*., 2001). Transgenic peas expressing *rpgip1*, like most GM plants, contain these

sequences what enables to use numerous commercial kits for their routine screening. It should be stressed however, that the detection of common GMO markers only indicates that a sample contains DNA from any GM plant and provides neither trait nor transgene confirmation. An alternative is to focus on target sequences but then any PCR-based strategy depends on a transgene and organism. Besides the majority of studied primer pairs identified the *rpgip1* transgene, some of them revealed additional bands corresponded to pea homologous sequences. Thus, the choice will depend very much on the objective of the PCR analysis. For routine screening purposes five pairs of primers (1F-366R, 1F-733R, 1F-958R, 421F-733R, 421F-958R) identifying unequivocally the *rpgip1* transgene can be recommended to detect as many *rpgip1* fragments as possible. If only a single pair is used, a small rearrangement within the transgene may prevent amplification and a GMO material may pass unnoticed. Notwithstanding several primer pairs can overcome this problem, the method is troublesome if many singleplex PCRs have to be used to test hundreds of samples as it is typical of food testing or GMO spread in the environment. Each pair needs a separate PCR reaction and sizing by agarose gel electrophoresis. Two proposed multiplex PCR assays allowing one-step identification of up to five GM-derived products provide promising simplification in detecting transgenic pea expressing antifungal genes. This procedure is in agreement with the current trends as emphasized by multiplex assays developed for the simultaneous detection of GM maize and soybean lines (Forte *et al*., 2005; Yoke-Kqueen *et al*., 2011). Rapidity and cost-efficiency are in favour of multiplex assays.

As more and more traits are introduced into plants, a challenge is how genetic variation of both donors and recipients influences the identification of GMOs. Risks rely on the false-positive results arising from wild ecotypes of a donor species. Primers 1F-366R gave a product of the same size in transgenic peas and wild raspberry. Similar fingerprints were observed for 421F-958R primers. This means that samples containing raspberry products may be misidentified as GM pea derivatives. It is therefore necessary to employ multi-primers assays that often entail clearer differentiation as the present studies demonstrated. Surprisingly, such considerations have been scarce, presumably because the majority of so far marketed GM plants carry bacterial or viral genes that are avoided in detection procedures (Bonfini *et al*., 2001).

A major aspect of the *rpgip1* detection in transgenic pea is the distinguishing between transgenes and endogenous homologues as shown for transgenic peas. This problem is rarely raised because most transgenic plants harbour insect or herbicide resistance genes that have not counterparts in plant genomes. Recently, genes from more or less related plant taxa are employed in plant transformation. A favourite example involves the wheat HMG gene, *Dy10* encoding glutenin subunit and introduced into several wheat cultivars (Abdalla, 2007). But even more distant relationships do not prevent a transgene from interfering with endogenous homologues. The *rpgip1* donor species, *R. idaeus* belongs to the Rosaceae family whereas the recipient, *P. sativum* is a member of Fabaceae. Nevertheless, five primer pairs amplified pea homologous sequences instead of the transgene. False-positive results originated from other genome homologues are critical in breeding transgenic crops. A common practice involves transformation of well responding genotypes and then, the transgene is introduced to other breeding lines by ordinary genetic crosses. Different allelic variants of homologues can have different affinity to primers designed on transgenes as well as a transgene may be rearranged during breeding. At this point, one can imagine employing one multiplex for preliminary analyses of the *rpgip1* transgene structure (e.g., set 1) and another to reveal the *rpgip1* transgene, insert sequences and homologues at once (e.g., set 2).

Important outcomes from the present studies is that some primers can identify homologues instead of the *rpgip1* transgene. Expected patterns of the pSCP1 template in addition to primer specificity check and lack of unspecific products by *in silico* PCR enable to exclude the well-known points of consideration for unspecific amplification. What remains contentious is single-primer binding to the DNA template (Ma *et al*., 2011) However, each primer was used in different combinations and only some of them resulted in unspecific amplification. On the other hand, only sequences with few mismatches (1 or 2) at the 3'end of primers can be used for effective PCR. A single mismatch at the last 10 bases can reduce the primer binding and can cause unintended amplification. Although positions of point mutations recognized at the *rpgip1* transgene are apart from primer binding sites, they demonstrate such possibilities. Of course, the prerequisite for unspecific products is the presence of sequences complementary to *rpgip1* primers in the pea genome. Otherwise, the amplification fails giving false-negative results. For plant genomes, highly copied short or long direct repeats can become the target for non-specific amplification. Indeed, four primers showed partial homology to pea retrotransposons. Just as transposons, endogenous *pgip*s and especially their conserved regions are good templates for primers directed towards the transgene. Unexpectedly, *in silico* PCR on two known pea homologues did not confirm this thesis. This simulation inevitably entails the presence of so far unknown *pgip* genes in pea. Plants have evolved many PGIPs differing in inhibition

profiles and recognition specificity to counterpart many PGs secreted by pathogens. Obviously, whether observed amplicons represent so far unknown pea *pgip*s or other genomic sequences has to be clarified by sequencing and phylogenetic analyses.

One final point to consider is point mutations that may arise at transgenic loci through successive reproductive generations. Data from *Arabidopsis thaliana* (Papazova *et al*., 2008) and maize MON810 (La Paz *et al*., 2010) suggest the high stability of transgenes and flanking sequences from one side but from the other, an example of oat lines has demonstrated multiple transgene rearrangements, truncated sequences and complex transgene loci (Makarevitch *et al*., 2003). These contrary results can simply mean that transgene behaviour depends on a gene and organism. The *rpgip1* transgene sequenced from the GM peas did not show any large rearrangements apart from a truncated fragment of 46 bp at 5'end. However, this may result from a sequencing procedure and explaining this lack needs further identification of transgene junctions. Noteworthy, two point mutations, one deletion and one substitution were recognized at the *rpgip1* locus in transgenic peas. Likewise, 34 point mutations such as small deletions and base pair substitutions have occurred in the transgene coding region of Roundup Ready soybean transgenic plants during 10 years since their release (Ogasawara *et al*., 2005). In both cases the rate of mutations is comparable, 1.1 x 10^{-3} at the *rpgip1* transgene and 0.87 x 10^{-3} at a transgene in soybean. Remarkably, the same research on Roundup Ready soybean has demonstrated similar mutation rates (0.9 x 10^{-3}) at *Cong* gene, the host locus encoding conglycinin storage proteins. According to the neutral theory of molecular evolution, the frequency of mutations at the transgene should be comparable with that for host genes but whether or not this hypothesis is true for the *rpgip1* transgene corroborates further comparisons.

To conclude, using several pairs of primers targeted towards different fragments of the *rpgip1* transgene is an advantage over a single pair because at once they identify the transgene, distinguish between the transgene and endogenous homologues as well as enable to avoid false positive results due to contamination from donor species genes. Mutliplex assays provide further cost effective simplification of the procedure. Mutations at transgenic loci may complicate GMO identification but they do not seem to be more frequent than at host genes.

5 ACKNOWLEDGEMENTS

The research was supported by the European Union, the Marie Curie Host Fellowship for the Transfer of Knowledge programme under the project GenCrop, MTKD-CT-2004-509834. We are grateful to Sylwia Ciaglo-Androsiuk for making crosses and isolating DNA from F_1 plants.

6 REFERENCES

Abdalla, K.S. 2007. A simple and reliable PCR-based method for detection and screening of transgenic plants transformed with the same endogenous gene. Arabian Journal of Biotechnology 10: 155-160.

Bonfini, L., Heinze, P., Kay, S., Van den Eede, G. 2001. Review of GMO detection and quantification techniques. Ispra: European Commission, Joint Research Centre, pp. 67.

Cousin, R. 1997. Peas (*Pisum sativum* L.). Field Crops Research 53: 111-130.

Di Matteo, A., Bonivento, D., Tsernoglou, D., Federici, L., Cervone, F. 2006. Polygalacturonase-inhibiting protein (PGIP) in plant defence: a structural view. Phytochemistry 67: 528-533.

Eurostat, 2011. European Commission eurostat [Database on the Internet]. [cited 2011 August]. Available from http:// epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/

Forte, V.T., di Pinto, A., Martino, C., Tantilo, G.M., Grasso, G., Schena, F.P. 2005. A general multiplex-PCR assay for the general detection of genetically modified soya and maize. Food Control 16: 535-539.

Janni, M., Sella, L., Favaron, F., Blechl, A.E., de Lorenzo, G., D'Ovidio, R. 2008. The expression of a bean PGIP in transgenic wheat confers increased resistance to the fungal pathogen *Bipolaris sorokiniana*. Molecular Plant-Microbe Interactions 21: 171-177.

La Paz, J.L., Pla, M., Papazova, N., Puigdomenech, P., Vicient, C.M. 2010. Stability of the MON 810 in maize. Plant Molecular Biology 74: 563-571.

Ma, J., Wang, P.W., Yao, D., Wang, Y.P., Yan, W., Guan, S.C. 2011. Single-primer PCR correction: a strategy for false-positive exclusion. Genetics and Molecular Research 10: 50-159.

Makarevitch, I., Svitashev, S.K., Somers, D.A. 2003. Complete sequence analysis of transgene loci from plants transformed via microprojectile bombardment. Plant Molecular Biology 52: 421-432.

McDonald, G.K., Peck, D. 2009. Effects of crop rotation, residue retention and sowing time on the incidence and survival of Ascochyta blight and its effect on grain yield of field peas (*Pisum sativum* L.). Field Crops Research 111: 11-21.

Morino, K., Olsen, O.A., Shimamoto, K. 1999. Silencing of an aleurone-specific gene in transgenic rice caused by a rearranged transgene. The Plant Journal 17: 275-285.

Ogasawara, T., Chikagawa, Y., Arakawa, F., Nozaki, A., Itoh, Y., Sasaki, K., Umetsu, H., Watanabe, T., Akiyama, H., Maitani, T., Toyoda, M., Kamada, H., Goda, Y., Ozeki, Y. 2005. Frequency of mutations of the transgene, which might result in the loss of the glyphosate-tolerant phenotype, was lowered in Roundup Ready soybeans. Journal of Health Science 51: 197-201.

Papazova, N., Ghedira, R., van Glabeke, S., Bartegi, A., Windels, P., Taverniers, I., Roldan-Ruiz, I., van Bockstaele, E., Milcamps, A., van den Eede, G., Depicker, A., de Loose, M. 2008. Stability of the T-DNA flanking regions in transgenic *Arabidopsis thaliana* plants under influence of abiotic stress and cultivation practises. Plant Cell Reports 27: 749-757.

Polok, K. 2007. Molecular evolution of the genus *Lolium* L. Olsztyn: "SQL", 317 p. Available from http://www.uwm.edu.pl/katgenbiol

Powell, A.L., van Kan, J., ten Have, A., Visser, J., Greve, L.C., Bennet, A.B., Labavitch, J.M. 2000. Transgenic expression of pear PGIP in tomato limits fungal colonization. Molecular Plant-Microbe Interactions 13: 942-950.

Prioul-Gervais, S., Deniot, G., Receveur, E.M., Frankewitz, A., Fourmann, M., Rameau, C., Pilet-Nayel, M.L., Baranger, A. 2007. Candidate genes for quantitative resistance to *Mycosphaerella pinodes* in pea (*Pisum sativum* L.). Theoretical and Applied Genetics 114: 971-984.

Richter, A., de Kathen, A., de Lorenzo, G., Briviba, K., Hain, R., Ramsay, G., Jacobsen, H.-J., Kiesecker, H. 2006. Transgenic peas (*Pisum sativum*) expressing polygalacturonase inhibiting protein from raspberry (*Rubus idaeus*) and stilbene synthase from grape (*Vitis vinifera*). Plant Cell Reports 25: 1166-1173.

Shanmugam, V. 2005. Role of extracytoplasmic leucine rich repeat proteins in plant defence mechanisms. Microbiological Research 160: 83-94.

Svitashev, S.K., Pawlowski, W.P., Makarevitch, I., Plank, D.W., Somers, D.A. 2002. Complex transgene locus structure implicates multiple mechanisms for plant transgene rearrangement. The Plant Journal 32: 433-445.

Szankowski, I., Briviba, K., Fleschhut, J., Schönherr, J., Jacobsen, H.-J., Kiesecker, H. 2003. Transformation of apple (*Malus domestica* Borkh.) with the stilbene synthase gene from grapevine (*Vitis vinifera* L.) and a PGIP gene from kiwi (*Actinidia deliciosa*). Plant Cell Reports 22: 141-149.

Timmerman-Vaughan, G.M., Frew, T.J., Russel, A.C., Khan, T., Butler, R., Gilpin, M., Murray, S., Falloon, K. 2002. QTL mapping of partial resistance to field epidemics of Ascochyta blight of pea. Crop Science 42: 2100-2111.

Yoke-Kqueen, C., Yee-Tyan, C., Siew-Ping, K., Son, R. 2011. Development of multiplex-PCR for genetically modified organism (GMO) detection targeting EPSPS and Cry1Ab genes in soy and maize samples. International Food Research Journal 18: 515-522.