Isolation, characterization and RFLP linkage mapping of a DNA repeat family of *Solanum spegazzinii* by which chromosome ends can be localized on the genetic map of potato

C. GEBHARDT*, B. EBERLE, C. LEONARDS-SCHIPPERS, B. WALKEMEIER
AND F. SALAMINI
Max-Planck-Institut für Züchtungsforschung, Carl-von-Linné-Weg 10, D-50829 Köln, Germany

Summary

In a random sample of 2263 cloned genomic DNA fragments of the wild potato species *Solanum spegazzinii* six related, highly repetitive fragments (SPG repeat family) were identified that were present in much higher copy numbers in *S. spegazzinii* when compared with the closely related cultivated potato *S. tuberosum*. The SPG repeat family was organized in long arrays of multiple copies. Cross hybridization experiments with 29 wild and cultivated *Solanum* species and with the related tomato showed specificity of the SPG repeat family for tuber-bearing *Solanum* species. Among tuber bearing *Solanum* species a high degree of variation was observed for restriction fragment length and copy number. The variation in copy number was not correlated with established taxonomic relationships between tuber-bearing *Solanum* species. DNA sequence analysis revealed a subrepeat structure of 120–140 base pairs embedded in longer repeat units of variable length. Length polymorphisms between highly repeated restriction fragments detected by the SPG probes were used for segregation- and linkage analysis in four mapping populations of potato, for which RFLP maps had been constructed. Twelve loci were identified, eleven of which mapped to the distal ends of nine linkage groups. All the evidence suggested that the SPG repeat family represents a satellite repeat members of which are localized in the subtelomeric region of potato chromosomes. The SPG repeat family could be used, therefore, for completing the genetic map of potato.

1. Introduction

Among the highly repetitive DNA fraction of plant genomes sequences have been identified and molecularly cloned that are 'species specific' (Bedbrook *et al.* 1980; Metzlaff *et al.* 1986; Ganal *et al.* 1988; Junghans & Metzlaff, 1988; Schweizer *et al.* 1988; Pehu *et al.* 1990; Schmidt *et al.* 1990; Schweizer *et al.* 1993). Based on DNA homology these sequences are detected in one species or in a group of closely related species but not in other species of the same taxonomic family. They can be used for studying the introgression of foreign germplasm (Bedbrook *et al.* 1980; Schmidt *et al.* 1990), for the identification of somatic fusion hybrids (Schweizer *et al.* 1988; Pehu *et al.* 1990) and for the analysis of genome structure by *in situ* hybridization (Bedbrook *et al.* 1980; Ganal *et al.* 1988).

The wild tuber-bearing *Solanum* species *S. spegazzinii* ($2n = 2x = 24$) can readily be crossed with diploid ($2n = 2x = 24$) breeding lines of *S. tuberosum* and has been used to introduce into the cultivated potato resistance alleles to the root cyst nematode *Globodera rostochiensis* (Ross, 1986; Barone *et al.* 1990). According to classical taxonomy, *S. spegazzinii* and *S. tuberosum* are closely related, both belonging to the series Tuberosa (Hawkes, 1990). Molecular taxonomy based on RFLPs separated the species of the series Tuberosa into two clusters, one including wild species such as *S. spegazzinii* and the other one including mainly the cultivated species such as *S. tuberosum* (Debener *et al.* 1990). We questioned whether species-specific repetitive DNA sequences could be isolated that are detectable in *S. spegazzinii* but not in *S. tuberosum* as was possible for the more distantly related species *S. acaule* versus *Lycopersicon esculentum* (Schweizer *et al.* 1988), *S. tuberosum* versus *S. brevidens* (Pehu *et al.* 1990) or *S. demissum* versus *S. tuberosum* (Schweizer *et al.* 1993). Such repetitive sequences would be useful markers in following the genome of *S. spegazzinii* after the introgression of nematode resistance alleles into the cultivated potato.

* Corresponding author.

We did not find a repeat family that was specific for *S. spegazzinii*. We did find, however, a repeat family that was present in higher copy number in *S. spegazzinii* when compared with *S. tuberosum* and that had the structural properties of satellite repeats found in plants as well as in animals (Miklos, 1986; Lapitan, 1992). Using the repeat sequence as a RFLP marker probe, it was possible to determine the position of telomeric regions on the potato RFLP map.

2. Materials and methods

(i) *Plant material*

Seeds of 28 wild and cultivated *Solanum* species were obtained from the germplasm collection of the Bundesforschungsanstalt für Landwirtschaft (FAL) of Braunschweig, Germany, by courtesy of J. R. Hoekstra. Between one and three accessions per species were acquired and a single plant of each accession number was analysed. The species were *S. acaule* (*acl*), *S. alandiae* (*aln*), *S. berthaultii* (*ber*), *S. brevicaule* (*bre*), *S. bukasovii* (*buk*), *S. canasense* (*can*), *S. chacoense* (*chc*), *S. commersonii* (*cmm*), *S. demissum* (*dms*), *S. etuberosum* (*etb*), *S. gourlayi* (*grl*), *S. hondelmannii* (*hdm*), *S. kurtzianum* (*ktz*), *S. leptophyes* (*lph*), *S. lignicaule* (*lgl*), *S. maglia* (*mag*), *S. microdontum* (*mcd*), *S. megistacrolobum* (*meg*) *S. neorossii* (*nrs*), *S. oplocense* (*opl*), *S. phureja* (*phu*), *S. tuberosum* ssp. *andigena* (*adg*) *S. spegazzinii* (*spg*), *S. sparsipilum* (*spl*), *S. stenotomum* (*stn*), *S. stoloniferum* (*sto*), *S. verrucosum* (*ver*) and *S. vernei* (*vrn*). The cultivated potato, *S. tuberosum* ssp. *tuberosum* (*tbr*), was represented by vegetatively propagated breeding lines of the Max-Planck-Institut für Züchtungsforschung and cv. Bintje. The tomato, *Lycopersicon esculentum*, was represented by cv. Estrella. Abbreviations of species names are according to Huaman & Ross (1985).

Segregation of RFLPs detected by the cloned DNA fragments SPG24, SPG108, SPG125 and SPG130, respectively, was followed in four RFLP mapping populations derived from crosses between diploid potato lines. Three of these crosses have been described in Gebhardt *et al.* (1991) and in Leonards-Schippers *et al.* (1994). The fourth population in which an RFLP map has been constructed (unpublished results from this laboratory) consisted of 157 F1 hybrids derived from crossing line H80.577/1 as female parent with line H80.576/16 as male parent (Gebhardt *et al.* 1989).

(ii) *Cloning of random genomic DNA fragments of S. spegazzinii*

Total genomic DNA (20 μg) of *S. spegazzinii* (BGRC 016929) at a concentration of 0·5 μg/μl was digested with 5×10^{-4} units/μl of DNAse I for 10 min at 15 °C (Anderson, 1981). The reaction was stopped by adding EDTA to a final concentration of 20 mM and by heat denaturation at 65 °C for 15 min. The DNA fragments were size fractionated on 2·5% agarose. Fragments between 200 and 2000 bp in length were isolated by electroelution and concentrated by precipitation with ethanol. The DNA fragments were then treated with 0·05 units/μl of T4 DNA polymerase for 1 h at 11 °C according to Anderson (1981), purified by phenol/chloroform extraction and precipitated with ethanol. *S. spegazzinii* DNA fragments (2 μg in 50 μl) were C-tailed according to Heidecker & Messing (1983) with 0·5 units/μl of terminal transferase at 37 °C for 15 min. 8 μg of the vector pUC9 (Messing & Vieira, 1982) were digested with *Pst* I and G-tailed in 20 μl with 0·05 units/μl of terminal transferase at 37 °C for 7 min. The tailing reactions were phenol/chloroform extracted and ethanol precipitated. Annealing between 1 ng G-tailed pUC9 and 10 ng C-tailed *S. spegazzinii* DNA was performed in 10 μl 10 mM Tris HCl, pH 7·8, 1 mM EDTA, 0·1 mM NaCl (Sambrook *et al.* 1989) for 5 min at 65 °C and 2 h at 57 °C. *E. coli* cells (strain TG-2) were transformed with the product of the annealing reaction following standard procedures (Sambrook *et al.* 1989).

(iii) *Identification of highly repetitive clones*

Colony filter hybridization of *E. coli* transformants was performed according to Grundstein & Hogness (1975). Colonies were screened in ordered 96-well microtitre plate arrays. Colony filters (Amersham, Hybond N) were hybridized against *c.* 100 ng of total DNA of *S. spegazzinii*, labelled with 30 μCi [α-^{32}P]dCTP (Amersham, Braunschweig, Germany) in 20 μl reaction volume using the random priming method of Fainberg & Vogelstein (1984). Hybridization and washing conditions were as described in Gebhardt *et al.* (1989).

Differential hybridization with total DNA of S. tuberosum ssp. tuberosum and S. spegazzinii

Plasmid DNA of clones with inserts of repetitive *S. spegazzinii* DNA was prepared according to Birnboim & Doly (1979). Equal amounts of plasmid DNA (1 μg) were restricted with *Eco*R I plus *Hin*d III, separated by agarose gel electrophoresis, blotted onto Hybond N membrane (Amersham) according to Southern (1975) and hybridized in turn with total genomic DNA of *S. spegazzinii* (*c.* 5×10^6 Bq/μg) and *S. tuberosum* ssp. *tuberosum* (*c.* 10^7 Bq/μg) labelled as above. Membranes were washed three times for 15 min in $1 \times$ SSC, 0·1% SDS at 65 °C, once for 30 min in $0.5 \times$ SSC, 0·1% SDS at 65 °C and exposed to Kodak X-OMAT AR X-ray film for 3 d. In between hybridizations, probes were removed by boiling the membranes for 30 min in 0·1% SDS.
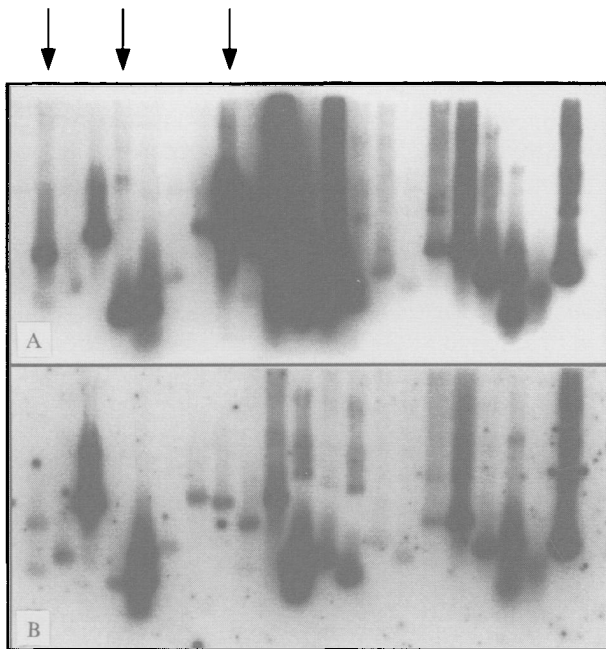
Fig. 1. Differential hybridization of repeated sequences containing clones to total genomic DNA of *S. spegazzinii* (A) and *S. tuberosum* ssp. *tuberosum* (B) as probe. 1 μg plasmid DNA of each clone was restricted with *Eco*R I plus *Hin*d III, separated on agarose gel, blotted onto Nylon membrane and hybridized first to *S. spegazzinii* and subsequently to *S. tuberosum* ssp. *tuberosum* DNA. Differentially hybridizing clones are indicated by arrows.

Southern blot analysis

Total DNA of single potato and tomato plants was isolated as described previously (Gebhardt *et al.* 1989). DNA (1 μg) of each genotype was digested with 5 units of *Eco*R I, *Hin*d III and *Taq* I, respectively, according to the supplier's instructions (Boehringer, Mannheim, Gibco BRL, Gaithersburg). *Eco*R I and *Hin*d III restricted DNA fragments were separated in 0·7% agarose and blotted onto Nylon membrane (Hybond N, Amersham) using standard procedures (Sambrook *et al.* 1989).

Taq I restricted DNA fragments were separated in a size range between 50 and 300 nucleotides on 6% denaturing polyacrylamide gels and electroblotted as described in Gebhardt *et al.* (1989). Filters were hybridized against *c.* 100 ng highly repetitive SPG probe labelled with 10 μCi [α-^{32}P]dCTP (Feinberg & Vogelstein, 1984), washed first at a moderate stringency (5 mM sodium phosphate, 1 mM EDTA, 0·2% SDS, pH 7·0, room temperature, Gebhardt *et al.* 1989), and then, after exposure, at a high stringency (0·1 × SSC, 0·1% SDS, 65 °C, 30 min) and reexposed.

RFLP mapping

Southern blot filters of the mapping populations were prepared as described in Gebhardt *et al.* (1989). Filters were hybridized against SPG24, SPG108, SPG125 and SPG130 probes, respectively, radio-
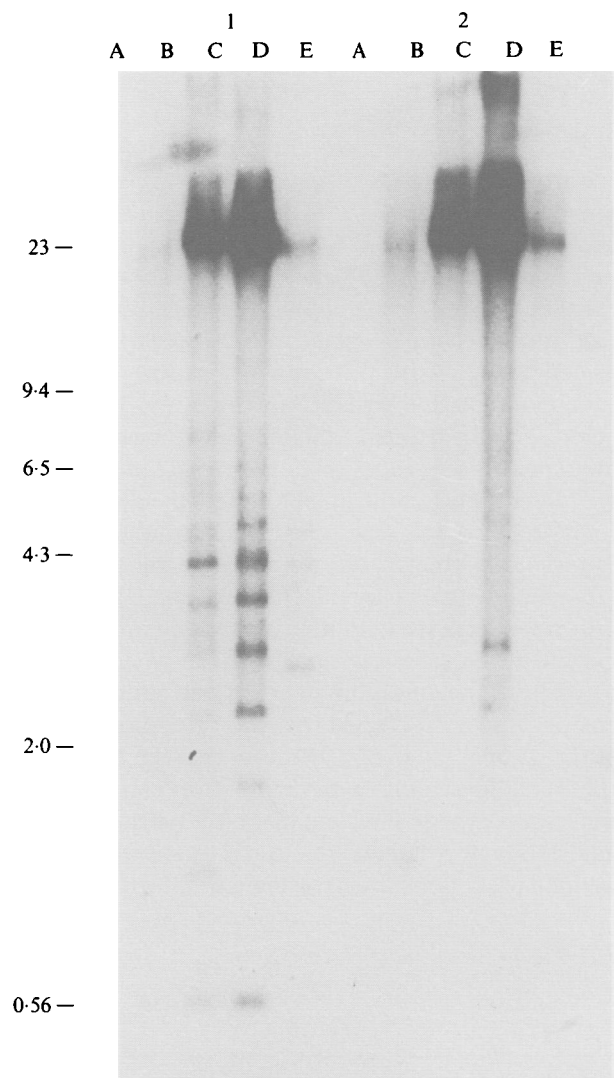

Fig. 2. Southern gel blot analysis of total DNA of *S. tuberosum* ssp. *andigena* (A), *S. demissum* (B), *S. spegazzinii* (C), *S. tuberosum* ssp. *tuberosum* (D) and *Lycopersicon esculentum* (E). Total DNA (1 μg) per genotype was restricted with *Eco*R I (1) or *Hin*d III (2), separated on 0·7% agarose, blotted, and hybridized to the SPG125 probe. The membrane was washed under high stringency condition (0·1 × SSC, 0·1% SDS, 1 h at 65 °C). Size markers are indicated on the left.

labelled as above, washed at moderate stringency (see above) and exposed to Kodak X-OMAT AR X-ray film for various times. Segregating fragments were mapped as described previously (Gebhardt *et al.* 1989, Ritter *et al.* 1990) using computer programs developed by E. Ritter.

DNA sequence analysis

*Eco*R V fragments of clones SPG24, SPG108 and SPG130 were subcloned using standard procedures (Sambrook *et al.* 1989). DNA sequences were determined according to Sanger *et al.* (1977) using the T7 polymerase kit (Pharmacia, Freiburg, Germany). Sequence data were analysed with the GCG programs

3. Results

Random genomic DNA fragments of *S. spegazzinii* ranging in size from 300 to 2000 bp were obtained by limited digestion with DNAse I and cloned in a plasmid vector. From 2263 clones screened for highly repeated sequences by colony hybridization with labelled total DNA of *S. spegazzinii*, between 168 and 291 – depending on the selection stringency – gave a positive hybridization signal. Between 7 and 13% of the clones were selected, therefore, as harbouring highly repeated sequences of *S. spegazzinii*.

By differentially hybridizing blots with the restricted plasmid DNA of 168 repetitive clones to total DNA of *S. spegazzinii* and *S. tuberosum*, respectively (Fig. 1), 16 clones were selected that showed stronger hybridization signals with the *S. spegazzinii* probe than with the *S. tuberosum* probe. Even when using high stringency conditions for hybridization, no repetitive clone was detected that was recognized only by the DNA of *S. spegazzinii*. Further characterization of the 16 differentially hybridizing clones by cross hybridization experiments and slot blot hybridization to total DNA of *S. spegazzinii*, *S. etuberosum* and *L. esculentum* (data not shown) resulted in the identifica-

tion of a group of six related clones that hybridized consistently more strongly to DNA of *S. spegazzinii* than to *S. tuberosum* and not at all to *S. etuberosum* or *L. esculentum* DNA. The insert size of this type of repetitive clone ('SPG clones') ranged from 300 to 1200 bp. The frequency was six of 2263 or 0·26%.

Southern gel blot analysis was performed with DNA of *S. tuberosum* ssp. *tuberosum*, *S. spegazzinii*, *L. esculentum* and a series of other wild and cultivated *Solanum* species (Fig. 2–4). When filters were hybridized to different SPG clones, very similar hybridization patterns were obtained.

In genomic restriction digests with *Eco*R I and *Hin*d III, respectively, the SPG probe hybridized under high stringency conditions, mostly to fragments larger than 20 kb (Fig. 2). After digestion with *Hae* III no hybridizing fragment larger than *c*. 560 bp was detected (data not shown). When *S. tuberosum* DNA was restricted with *Eco*R V, hybridization to SPG probes revealed seven major repeated fragments of 185, 320, 340, 505, 520, 650 and 670 basepairs (data not shown). It was concluded from this result, that the SPG repeat units were not dispersed in the genome but organized in large tandem arrays of multiple copies. Fig. 1 also demonstrates the large difference in copy number between *S. tuberosum* (ssp. *andigena* and *tuberosum*) and *S. spegazzinii* and *S. demissum*, respectively, and the absence of the repeated sequence
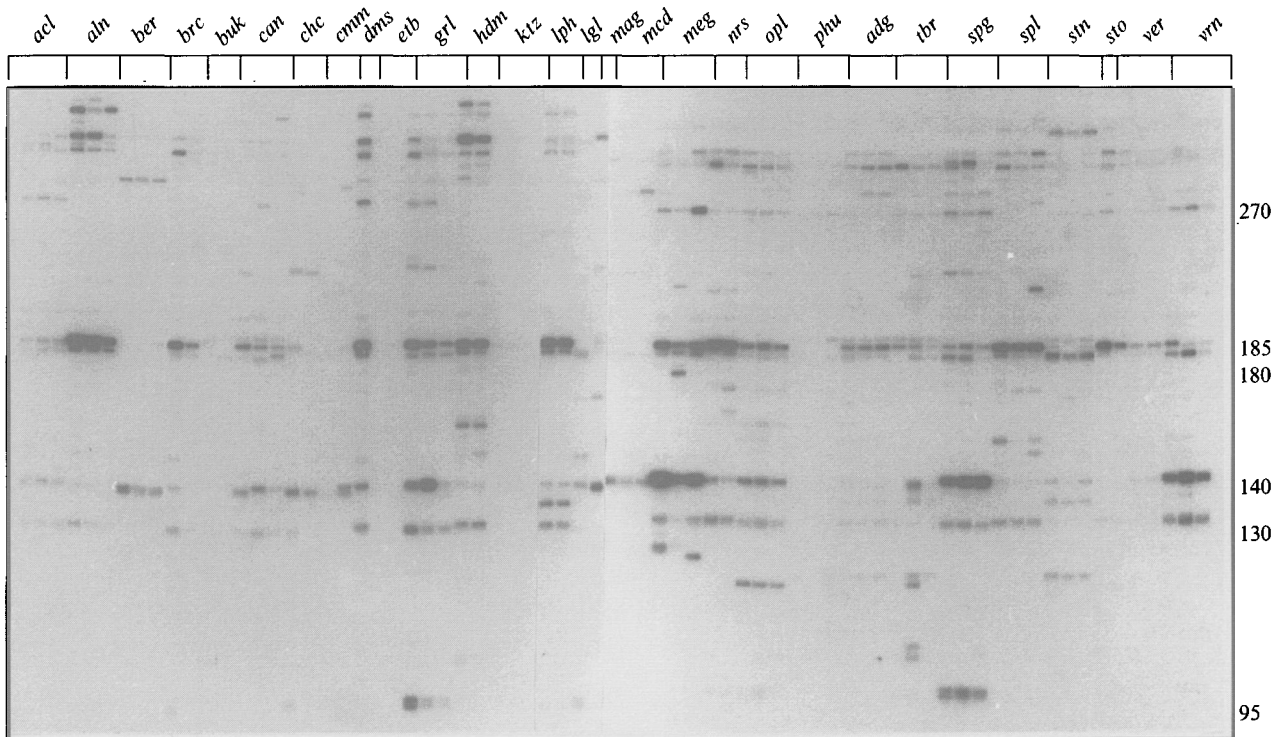

Fig. 3. Restriction fragment patterns of wold and cultivated *Solanum* species after hybridization with the SPG125 probe. Species names are abbreviated by a three letter code (Materials and Methods). Each species is represented by DNA of a single plant of one to three different accession numbers. Total DNA (1 μg) per genotype was restricted with *Taq* I, separated on a 6% denaturing polyacrylamide gel, electroblotted and hybridized to the SPG probe. The membranes were washed under high stringency condition (0·1 × SSC, 0·1% SDS, 30 min at 65 °C). The sizes in basepairs of major hybridizing restriction fragments are indicated on the right.

in *Lycopersicon esculentum.* In a control experiment using the same Southern filters and cloned ribosomal DNA as probe (Landsmann & Uhrig 1985) strong signals of similar size and intensity were detected in all five species (data not shown).

In a high resolution Southern gel blot analysis *Taq* I restricted DNA of 29 *Solanum* species, each represented by one individual plant of one to three accessions, was hybridized to the SPG probes SPG24, SPG108, SPG125 and SPG130, respectively, at moderate and high stringency.

The hybridization patterns obtained at high stringency with probe SPG125 are shown in Fig. 3. A high degree of variation was observed among the species in terms of fragment length polymorphism and copy number. Within the species, restriction fragment length variation was also detected, although the major hybridizing bands were preserved (see also Fig. 4). The non-tuber-bearing *S. etuberosum* (*etb*) was the only *Solanum* species which did not hybridize at all to the probe. *S. acaule* (*acl*), *S. bukasovii* (*buk*), *S. kurtzianum* (*ktz*), *S. microdontum* (*mcd*) and *S. phureja* (*phu*) showed low signal intensities when compared, for example, with *S. alandiae* (*aln*), *S. demissum* (*dms*), *S. gourlayi* (*grl*), *S. hondelmannii* (*hdm*), *S. leptophyes* (*lph*), *S. megistacrolobum* (*meg*), *S. neorossii* (*nrs*), *S. spegazzinii* (*spg*), *S. sparsipilum* (*spl*) and *S. vernei* (*vrn*), that hybridized most strongly to the probe. The remaining species showed intermediate hybridization signals. Among the species of the genus *Solanum* tested, the repeated sequence was restricted, therefore, to tuber-bearing species.

When the same filters as shown in Fig. 3 were hybridized at moderate stringency to SPG24, SPG108, SPG125 and SPG130, very similar patterns were observed with the exception of two fragments, *c.* 130 and 150 bp long. This is shown in Fig. 4 for the three plants of *S. spegazzinii.* The 150 bp fragment was not detected by probes SPG125 and SPG130 and the 130 bp fragment was hardly detected by SPG24, whereas both fragments were detected with SPG108.

Restriction analysis had shown that the insert of clone SPG24 (960 bp), SPG108 (840 bp) and SPG130 (1200 bp) contained an *Eco*R V fragment of 340 bp, 520 bp and three fragments of 300 bp, respectively. These fragment sizes corresponded to the sizes of some of the major *Eco*R V fragments detected in Southern gel blots of genomic DNA (see above). The *Eco*R V fragments present in SPG24, SPG108 and SPG130 were subcloned and sequenced. Clone SPG125 (290 bp) was directly sequenced from both ends. The aligned sequences are shown in Fig. 5. The homologous sequences were between 87 and 91% identical. Common to all four sequences was a subrepeat, *c.* 120–140 bp long, that extended upstream and downstream of the position of the *Eco*R V restriction site. The subrepeat was embedded in longer repeat units of different size as identified by the position of the *Eco*R V restriction sites in the aligned
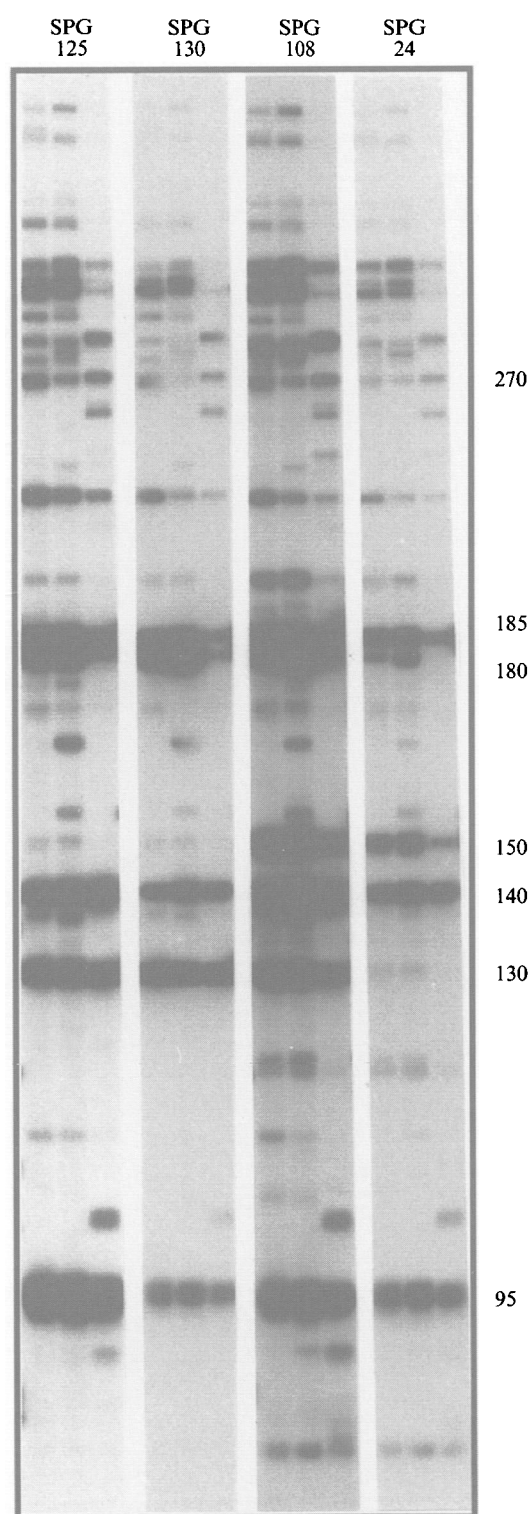

Fig. 4. Restriction fragment pattern of *Solanum spegazzinii* after hybridization with SPG125, SPG130, SPG108 and SPG24 probes. The same membranes as shown in Fig. 3 were used but washed at moderate stringency (Materials and Methods).

sequences. Clones SPG24 and SPG108 contained a repeat unit of 341 and 337 bp, respectively (position 3 to 346 in Fig. 5). The *Taq* I fragment of 150 bp that was specifically detected by SPG24 and SPG108 on genomic Southern blots (Fig. 4) corresponded to the *Taq* I restriction sites at positions 196 and 348 in Fig.

```
              EcoRV
SPG108    GATATCGAAA AACATACGCG GGTTCAAAAA AAAAATCGCC TGAAACGGAC        50
SPG 24    GATATCGAAA AA.ATACGCG GGTTCAACAA AAAAATCGCC TG.AACGGAC


SPG108    GTCCGAGGCC CCAGTTATGG CCGTTTgaag tttcacaaat tgacag.cta       100
SPG 24    GTCCGAAGCC CGGGTTATGG CCGTTTgaag tttaccaaat tcacagccta


SPG108    ccttcaatct ttaagatgct gaattttcag acggtcgtaa cttttgactc       150
SPG 24    ccttcaatct ttaagatgct gaattttcag atgttcgtag cttttgactc

                                                            Taq I
SPG108    ggacgtccgt ttgacgcaat ctttttttg aatcctcgta ttttttcgag        200
SPG 24    ggacgtccgt ttgacgcaat tttttttttg aatcctcgta ttttttcgag

                                        Taq I
SPG108    acctacttga aaaactcatc cgcttcgact gaatttcgtt tttatttttt       250
SPG 24    atatatttgg aaaactcatc cgcgtcaacc gaatttccat tttatttttt


SPG108    ca.cacattt .ccact.gta tgag.aaaac agtAAAAGGC ATTTTTGGGA       300
SPG 24    catcatattt tacgctcgta taagaaaaac agtAAAAGGC ATTTTCGGGA

                                                            Taq I
                                                            EcoRV
SPG130    .......... .......... .......... .......... ...GATATCG      350
SPG125    .......... .......... .......... ......TTGC GTAGATATCA
SPG108    AAATCG.... GCGCAAACCG CGTCACGGTG GTTGGCTGC GTAGCTATCG
SPG 24    AAATCGGCAC GCGCAAACCG CGTCACGGTG ATTTGGTTGC GTAGATATC.

                                        Taq I
SPG130    AAAAAATACG CGGGTTCAAA AAAAAAATCG CGTCAAACGG A.CGTACGAA       400
SPG125    AAAAAATACG CGGTTTCAAA AATAGAATCG CGTCGAACGG A.CACCCGAA
SPG108    AAAAAATACG CAGGTTCAAA CAAAAAATTG CATCAAACGG AGCGTCCGAG

          Taq I
SPG130    GCTCGAGTTA TGGCCGTTTg aatttgcttt tttcctttct catcttaatt       450
SPG125    GCTCGAGTTA TGGCCGTTTg aattt.cctt tttcctttct catcttaatt
SPG108    GCTC.AGATA TGGCCGTTTg aattt.cctt tttcctttct catcttattt


SPG130    ctacaattca aagaatgagc AAAATGCGTT TTTTGAAAGA TCGGCTCGCT       500
SPG125    ctactattca aggaatgggt AAAATGCGTT TTTTGCAAAA TCGGCTCGCT
SPG108    ctaccattca aagaatgagt GAAATGCGTT TTTTTGGAAA TCGGCTCGCT

                                        Taq I
                                        EcoRV
SPG130    CAAACCACCT TACGGTAGTT TGGTTGCGTA CGTATCGTAA AAATACGCGG       550
SPG125    CAAACCACCT AACGGTGGTT TGGTTGCGTA GATATCGAAA AAATACGCGG
SPG108    CAAACCACCT CACGGCAGTT TGGTTGCGTA GATATC.... ..........


SPG130    ATTCAAAAAA TAAATCGCCT GAAACGGACG TCCGAAGCCC gagtttggga       600
SPG125    GTTCAAAAAA AAAATCGCCT GAAACGGACG TCCGAAGCCC AAGTTATGAC

                                                  EcoRV
SPG130    aattcggcat GCGCAAATCG GGAAACGGTA GTTTGGTTGT GTAGATATC.       650
SPG125    CGTTTGAAGT TTcccaaatc cacagcta.. .......... ..........
```

Fig. 5. Aligned DNA sequences of SPG clones SPG130, SPG125, SPG108 and SPG24. The subrepeat regions are shown in capital letters. The positions of *Eco*R V and *Taq* I restriction sites are indicated.
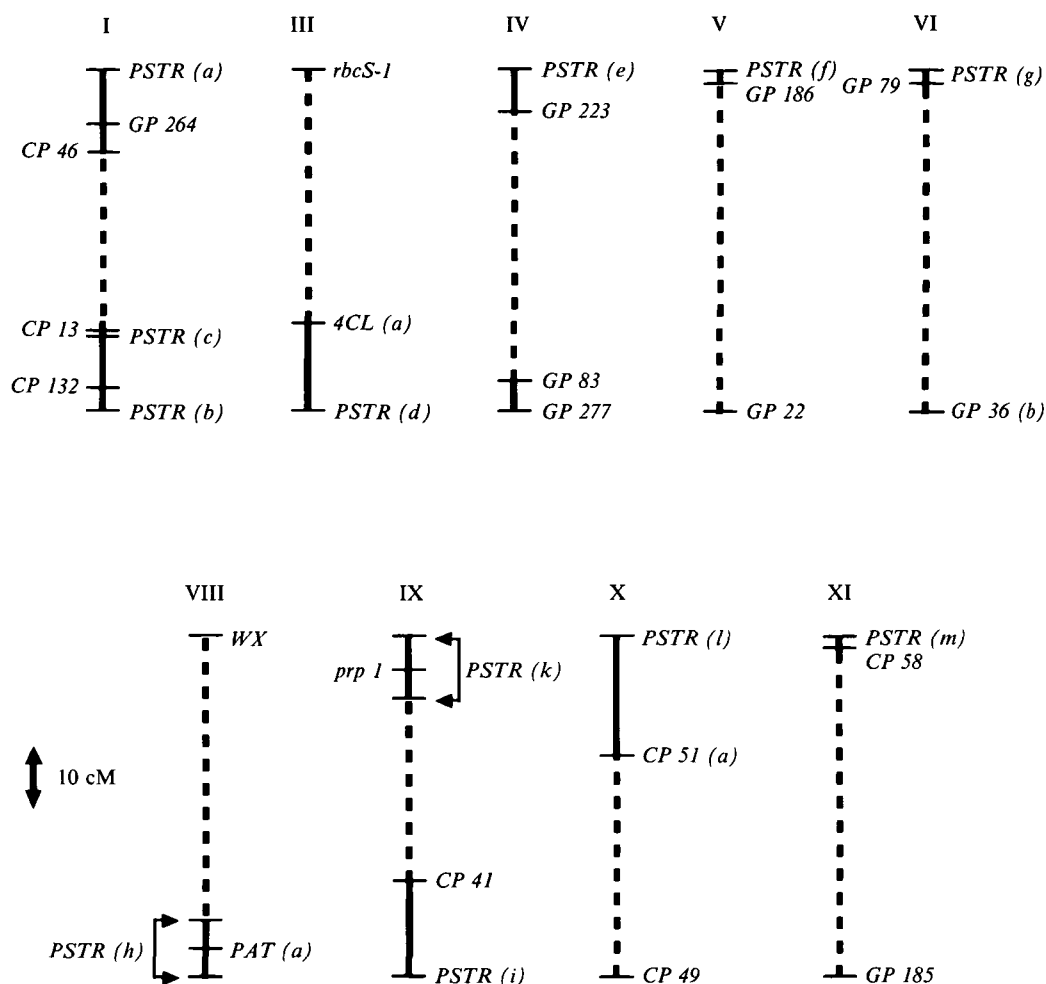
Fig. 6. Positions of twelve *PSTR* loci on the potato RFLP map. Only the ends of nine of the twelve potato linkage groups are shown as indicated by the distal RFLP markers on each linkage group according to the current status of the potato RFLP map (Gebhardt *et al.* 1994). The genetic positions of *PSTR* loci are shown relative to the most closely linked RFLP loci found in four different mapping populations. The orientation of *PSTR*(h) and *PSTR*(k) relative to *PAT*(a) (chromosome VIII) and *prp* 1 (chromosome IX) could not be determined unequivocally.

5. Clones SPG125, SPG130 and also SPG108 contained a repeat unit of 185 or 186 bp (position 347 to 533 in Fig. 5). The *Taq* I fragment of 130 bp that was detected by SPG125, SPG130 and SPG108, but not by SPG24 on genomic Southern blots (Fig. 4) is accounted for by *Taq* I restriction sites at positions 403 and 535 in Fig. 5. The sequences of the 185 bp repeat unit were between 82 and 87% homologous to the sequence of the 183 bp 'satellite' repeat (clone pSA287) isolated from *S. acaule* (Schweizer *et al.* 1988). The structure of SPG130 was, to some extent, more different from the other three SPG clones as it contained shorter or at the ends more highly diverged versions of the subrepeat (position 471 to 590 and position 611 to 649 in Fig. 5).

When Southern blots carrying *Taq* I restricted DNA of different genotypes of *S. tuberosum* were probed with the SPG clones under moderate stringency conditions, restriction fragment length polymorphism was observed. Several highly repeated restriction fragments segregated in mapping populations used for constructing RFLP linkage maps and could be

allocated, therefore, to chromosomal positions relative to RFLP loci of known position. Linkage analysis in four independent mapping populations using the SPG clones as marker probes identified 12 loci on nine potato chromosomes (*PSTR*(a)–*PSTR*(m)). Nine loci mapped clearly distal to the most distal RFLP markers of potato chromosomes I, III, IV, V, X and XI (Fig. 6, Gebhardt *et al.* 1991, 1994). Two loci (*PSTR*(h) and *PSTR*(k) could not be unequivocally orientated relative to the most distal RFLP locus of chromosomes VIII and IX, respectively, and one (*PSTR*(c)) mapped proximal to the most distal RFLP locus. Two chromosomes, I and IX, carried *PSTR* loci at both ends. The genetic distance between *PSTR* loci and the most closely linked RFLP loci ranged from 2 to 21 cM with a mean of 8·3 cM.

4. Discussion

Between 7 and 13% of the genome of the tuber bearing wild species *S. spegazzinii* consists of highly repetitive DNA. This estimate was derived from the

fraction of highly repeated sequences containing clones in a random sample of cloned genomic DNA fragments. The percentage of highly repetitive nuclear DNA might be even lower as chloroplast and mitochondrial sequences were not excluded from the library by screening with chloroplastic and mitochondrial DNA probes. These numbers are similar to the ones found for the related tomato (Zamir & Tanksley 1988).

The six SPG clones were equivalent to *c.* 0·26 % of the total genome and contained a repetitive sequence of *S. spegazzinii* that was present in lower copy numbers in *S. tuberosum*. The difference in copy number was the basis of their identification. A homologous repetitive sequence has been isolated from *S. acaule* based on a different approach: highly repetitive sequences were identified among randomly cloned genomic *Taq* I restriction fragments and analysed for species specificity. The frequencies of 0·2 % to 0·4 % estimated for the *S. acaule* repeat sequence in potato cultivars, breeding lines and tuber-bearing wild species of *Solanum* (Schweizer *et al.* 1993) were in the same order of magnitude as the frequency of the SPG repeat family. The subtelomeric TGRI repeat family of tomato had, however, with 1·75 % a considerably higher frequency (Ganal *et al.* 1988, Lapitan *et al.* 1989).

A repetitive sequence that was present in *S. spegazzinii* and completely absent in *S. tuberosum* was not found. Taking into consideration the number and size of the randomly cloned genomic DNA fragments that were surveyed (Ganal *et al.* 1988), it is unlikely, that such sequences would have escaped detection, at least at frequency levels above 0·2 % of the total genome. This suggests that this type of repetitive sequence does not exist in *S. spegazzinii*. The SPG repeat family showed, similar to the homologous repeat unit of *S. acaule* (Schweizer *et al.* 1993), specificity for tuber-bearing *Solanum* species (subsection *Potatoe*, section *Petota*, genus *Solanum* (Hawkes, 1990), as it was not detectable, even at moderate hybridization stringency, in the non-tuber-bearing species *S. etuberosum* (subsection *Estolonifera*) and in tomato belonging to the most closely related genus *Lycopersicon*.

Whereas the presence or absence of the SPG repeat family can be used, therefore, as a molecular taxonomic criterion at the subsection level, the variability in copy number and the polymorphisms between highly repetitive restriction fragments may be used for the identification of tuber-bearing *Solanum* species. Although RFLPs were also observed between individual plants of different accessions of the same species, there seemed to be characteristic patterns for the major repeated fragments in some species, for example *S. alandiae*, *S. berthaultii*, *S. hondelmannii*, *S. leptophytes*, *S. oplocense* and *S. stenotomum* (Fig. 3). Others, for example *S. spegazzinii* and *S. gourlayi*, could not be distinguished from each other but from

most other species by the SPG repeat pattern (Fig. 3). The stability and specificity of the patterns, however, has to be tested in more individual plants of each species.

The variation observed between tuber-bearing *Solanum* species with respect to the copy number of the SPG repeat family did not correspond to the molecular taxonomy of the same species based on nuclear single copy RFLPs (Debener *et al.* 1990) or to classical taxonomy based on morphological criteria (Hawkes 1990). For example, most of the wild species analysed with SPG probes in this paper are members of the same series Tuberosa and are clustered as a group when genetic distance is measured by nuclear RFLPs (Debener *et al.* 1990). *S. spegazzinii*, *S. gourlayi* and *S. vernei* containing high copy numbers of the SPG repeat family are in the same cluster with *S. kurtzianum* and *S. berthaultii* having very low copy numbers. On the other hand, *S. demissum* and *S. acaule* are members of the series Demissa and Acaulia, respectively, and they are well separated from series Tuberosa (Debener *et al.* 1990). *S. demissum* has, however, high copy numbers similar to *S. spegazzinii*, whereas *S. acaule* has a low copy number comparable with *S. berthaultii* (Fig. 2, 3). The fluctuation in copy number independent from the genetic distance between related species is characteristic for satellite DNA which evolves at different rates from the rest of the genome (John & Miklos 1979).

Southern blot analysis, DNA sequencing and segregation analysis revealed further structural properties of the SPG repeat family that are common to 'satellite' repeats found in most plant species (Lapitan, 1992). The SPG sequences are organized in large arrays of tandem repeat units as shown by the lack of restriction sites for some enzymes and presence at regular intervals of restriction sites for other enzymes. The lengths of the repeat units correspond to the ones found in other plant species (Bedbrook *et al.* 1980, Martinez-Zapater *et al.* 1986, Ganal *et al.* 1988, Pehu *et al.* 1990). Sequence analysis of different repeat units revealed the presence of a subrepeat, *c.* 120–140 bp in length, that is part of at least three different types of higher order repeats. A similar organization has been described for major repeated sequence families of *Secale cereale* telomeric heterochromatin (Bedbrook *et al.* 1980). The sequences of independently cloned SPG repeat units were approximately 10 % diverged. Besides sequence divergence, polymorphisms were detected between highly repeated genomic restriction fragments. This can be explained by point mutations followed by amplification of the mutated repeat unit. Several repeated DNA fragments segregated as single Mendelian factors indicating that, after amplification, uniform blocks of specific repeat variants have been maintained within the species as alleles.

The genetic loci that were identified by using SPG repeat units as probes in several RFLP mapping populations were – with one exception – all located at

the ends of the RFLP linkage groups. This suggests that the SPG repeat family as a whole or at least some of its variable members are localized in the telomeric region of the potato chromosomes similar to the subtelomeric repeat TGRI of tomato (Ganal *et al.* 1991, 1992). Eleven of the 24 ends of the 12 potato chromosomes could be identified by RFLP mapping of SPG repeat units. The mean genetic distance between *PSTR* loci and the most closely linked RFLP loci was 8·3 cM. The RFLP map of potato which comprises approximately 400 RFLP loci and *c.* 1050 cM (Gebhardt *et al.*, 1991, 1994) covers now, when the eleven *PSTR* loci are also considered, approximately 1140 cM (1050 cM + (11 × 8·3 cM)). Assuming that *c.* 108 cM (13 × 8·3 cM) are still missing from the remaining 13 chromosome ends, approximately 90 % of the potato genome are covered with molecular markers.

References

Anderson, S. (1981). Shotgun DNA sequencing using cloned DNaseI generated fragments. *Nucleic Acids Research* **9**, 3015–3017.

Barone, A., Ritter, E., Schachtschabel, U., Debener, T., Salamini, F. & Gebhardt, C. (1990). Localization by restriction fragment length polymorphism mapping in potato of a major dominant gene conferring resistance to the potato cyst nematode *Globodera rostochiensis*. *Molecular & General Genetics* **224**, 177–182.

Bedbrook, J. R., Jones, J., O'Dell, M., Thompson, R. D. & Flavell, R. B. (1980). A molecular description of telomeric heterochromatin in *Secale* species. *Cell* **19**, 545–560.

Birnboim, H. C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**, 1513.

Debener, T., Salamini, F. & Gebhardt, C. (1990). Phylogeny of wild and cultivated *Solanum* species based on nuclear restriction fragment length polymorphisms (RFLPs). *Theoretical & Applied Genetics* **79**, 360–368.

Feinberg, A. F. & Vogelstein, B. (1984). A technique for radiolabelling DNA restriction fragments to high specific activity. *Addendum Analytical Biochemistry* **137**, 266–267.

Ganal, M. W., Broun, P. & Tanksley, S. D. (1992). Genetic mapping of tandemly repeated telomeric DNA sequences in tomato (*Lycopersicon esculentum*). *Genomic* **14**, 444–448.

Ganal, M. W., Lapitan, N. L. V. & Tanksley, S. D. (1988). A molecular and cytogenetic survey of major repeated DNA sequences in tomato (*Lycopersicon esculentum*). *Molecular & General Genetics* **213**, 262–268.

Ganal, M. W., Lapitan, N. L. V. & Tanksley, S. D. (1991). Macrostructure of the tomato telomeres. *The Plant Cell* **3**, 87–94.

Gebhardt, C., Ritter, E., Barone, A., Debener, T., Walkemeier, B., Schachtschabel, U., Kaufmann, H., Thompson, R. D., Bonierbale, M. W., Ganal, M. W., Tanksley, S. D. & Salamini, F. (1991). RFLP maps of potato and their alignment with the homeologous tomato genome. *Theoretical & Applied Genetics* **83**, 49–57.

Gebhardt, C., Ritter, E., Debener, T., Schachtschabel, U.,

Walkemeier, B., Ugrig, H. & Salamini, F. (1989). RFLP analysis and linkage mapping in *Solanum tuberosum*. *Theoretical & Applied Genetics* **78**, 65–75.

Gebhardt, C., Ritter, E. & Salamini, F. (1994). RFLP map of the potato. In *DNA-based Markers in Plants* (ed. R. L. Phillips, I. K. Vasil), pp. 271–285, The Netherlands: Kluwer Academic Publishers.

Genetics Computer Group (1991). *Program Manual for the GCG Package*, Version 7, April 1991, 575 Science Drive, Madison, Wisconsin, USA 53711.

Grundstein, M. & Hogness, D. S. (1975). Colony hybridization: a method of the isolation of cloned DNAs that contain a specific gene. *Proceedings of the National Academy of Sciences USA* **72**, 3961–3965.

Hawkes, J. G. (1990). *The Potato–Evolution, Biodiversity and Genetic Resources*. London: Belhaven Press.

Heidecker, G. & Messing, J. (1983). Sequence analysis of zein cDNAs obtained by an efficient mRNA cloning method. *Nucleic Acids Research* **11**, 4891–4906.

Huaman, Z. & Ross, R. W. (1985). Updated listing of potato species, names, abbreviations and taxonomy status. *American Potato Journal* **62**, 629–641.

John, B. & Miklos, G. L. G. (1979). Functional aspects of satellite DNA and heterochromatin. *International Review of Cytology* **58**, 1–114.

Junghans, H. & Metzlaff, M. (1988). Genome specific, highly repeated sequences of *Hordeum vulgare*: cloning, sequencing and squash dot test. *Theoretical & Applied Genetics* **76**, 728–732.

Landsmann, J. & Uhrig, H. (1985). Somaclonal variation in *Solanum tuberosum* detected at the molecular level. *Theoretical & Applied Genetics* **71**, 500–505.

Lapitan, N. L., Ganal, M. W. & Tanksley, S. D. (1989). Somatic chromosome karyotype of tomato based on *in situ* hybridization of the TGRI satellite repeat. *Genome* **32**, 992–998.

Lapitan, N. L. V. (1982). Organization and evolution of higher plant nuclear genomes. *Genome* **35**, 171–181.

Martinez-Zapater, J., Estelle, M. & Somerville, C. (1986). A highly repeated DNA sequence in *Arabidopsis thaliana*. *Molecular & General Genetics* **204**, 417–423.

Messing, J. & Vieira, J. (1982). A new pair of M13 vectors for selecting either DNA strand of double-digested restriction fragments. *Gene* **19**, 269–276.

Metzlaff, M., Tröbner, W., Baldauf, F., Schlegel, R. & Cullum, J. (1986). Wheat specific repetitive DNA-sequences-construction and characterization of four different genomic clones. *Theoretical & Applied Genetics* **72**, 207–210.

Miklos, G. (1986). Localized highly repetitive DNA sequences in vertebrate and invertebrate genomes. In *Molecular Evolutionary Genetics* (ed. R. MacIntyre), pp. 241–322. New York: Plenum Press.

Pehu, E., Thomas, M., Poutala, T., Karp, A. & Jones, M. G. K. (1990). Species-specific sequences in the genus *Solanum*: identification, characterization and application to study somatic hybrids of *S. brevidens* and *S. tuberosum*. *Theoretical & Applied Genetics* **80**, 693–698.

Ritter, E., Gebhardt, C. & Salamini, F. (1990). Estimation of recombination frequencies and construction of RFLP-linkage maps in plants from crosses between heterozygous parents. *Genetics* **125**, 645–654.

Ross, H. (1986). Potato-breeding – problems and perspectives. *Journal of Plant Breeding Supplement* **13**.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning, A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* **74**, 5463–5467.

Schmidt, T., Junghans, H. & Metzlaff, M. (1990). Construction of *Beta procumbens*-specific DNA probes and their application for the screening of *B. vulgaris* × *B. procumbens* (2n = 19) addition lines. *Theoretical & Applied Genetics* **79**, 177–181.

Schweizer, G., Borisjuk, N., Borisjuk, L., Stadler, M., Stelzer, T., Schilde, L. & Hemleben, V. (1993). Molecular analysis of highly repeated genome fractions in *Solanum* and their use as markers for the characterization of species and cultivars. *Theoretical & Applied Genetics* **85**, 801–808.

Schweizer, G., Ganal, M., Ninnemann, H. & Hemleben, V. (1988). Species-specific DNA sequences for identification of somatic hybrids between *Lycopersicon esculentum* and *Solanum acaule*. *Theoretical & Applied Genetics* **75**, 679–684.

Southern, E. M. (1975). Detection of specific sequences among DNA fragment separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.

Zamir, D. & Tanksley, S. D. (1988). Tomato genome is comprised largely of fast-evolving, low copy-number sequences. *Molecular & General Genetics* **213**, 254–261.