Theor Appl Genet (1994) 88: 845–851

W. J. Phillips · C. G. D. Chapman · P. L. Jack

Molecular cloning and analysis of one member of a polymorphic family of GACA-hybridising DNA repeats in tomato

Received: 1 September 1993 / Accepted: 25 November 1993

Abstract Simple sequence repeat oligonucleotides were used to probe the tomato genome for elements displaying variability amongst commercial cultivars. The oligonucleotide $(GACA)_4$ was found to be particularly informative on genotype screening blots, hybridising to a highly polymorphic family of elements, and was used to clone one such member from a lambda library. The GACA-hybridisation was localised to a 1.3-kb *Hin*fI fragment within the original 15-kb lambda insert. This 1,349-bp subclone (pTGACA-2:1.3) was used to probe 27 Californian processing varieties and found to be capable of distinguishing all from each other, thus demonstrating its utility as a genetic fingerprinting probe for cultivar identification. Hybridisation occurred to approximately 10 major high molecular weight (> 4-kb) bands, most of which segregated independently in F_2 populations, as well as a large number of less clearly resolvable smaller fragments. Sequence analysis of the cloned element reveals that it is almost entirely composed of GACA or GATA repeats. These tetranucleotides are organised into distinct repetitive domains, consisting either of tandem arrays of each tetranucleotide or interspersions of GACA and GATA to form dodecanucleotides that are then further repeated. The boundaries between domains contain sufficient departures from the concensus repeat to allow construction of unique polymerase chain reaction (PCR) primers. Amplification from two such contiguous regions identifies length variation in both, thus yielding a genotype screen appropriate for high-throughput applications, such as assessment of purity in F_1 hybrid seed lots.

Key words Tomato · *Lycopersicon esculentum* · GACA Genetic fingerprinting · RFLP · PCR

Communicated by J. W. Snape

W. J. Phillips · C. G. D. Chapman · P. L. Jack (✉)
Plant Breeding International Cambridge Limited,
Maris Lane, Cambridge CB2 2LQ,
UK

Introduction

Low copy DNA elements displaying allelic length variation following restriction enzyme digestion (RFLPs) have proven to be powerful tools for mapping plant genomes and have considerable potential as markers for specific traits in breeding programmes. Such DNA markers are highly effective in some crops e.g. maize, brassicas and potatoes, though their utility is severely compromised by the low levels of RFLP encountered in crops of narrow genetic base, notably tomato (Miller and Tanksley 1990), wheat (Chao et al. 1989) and soybean (Keim et al. 1989). Nevertheless, these crops do contain substantial variation for yield, disease resistance, quality characters and a host of other traits which the plant breeder wishes to manipulate but for which informative RFLP markers cannot easily be identified. One consequence of this low level of variation is that linkage maps have not yet been constructed within the cultivated tomato (*Lycopersicon esculentum*) but only indirectly through inter-specific crosses such as those between *L. esculentum* and *L. pennellii*, where over 90% of markers are informative (Helentjaris et al. 1985, Tanksley et al. 1992). Unfortunately, such maps have limited utility in crosses between cultivars in breeding programmes. Given the difficulty in identifying informative (low copy) RFLP markers within such species, it is reasonable to question whether they always offer the best means of following genetic traits. For example, repeated DNA sequences, especially in combination with the polymerase chain reaction (PCR), have proven to be powerful markers in exploring genetic diversity within human populations (Jeffreys et al. 1991; Weissenbach et al. 1992).

A number of workers have demonstrated the presence of length-variable repeat arrays in plants (Dallas 1988; Weising et al. 1989; Nybom and Schaal 1990; Beyermann et al. 1992; Broun et al. 1992; Vosman et al. 1992), though in most cases these studies have been restricted to whole genome Southern analysis. Exceptions include Daly et al. (1991), who have cloned variable tandem repeat elements

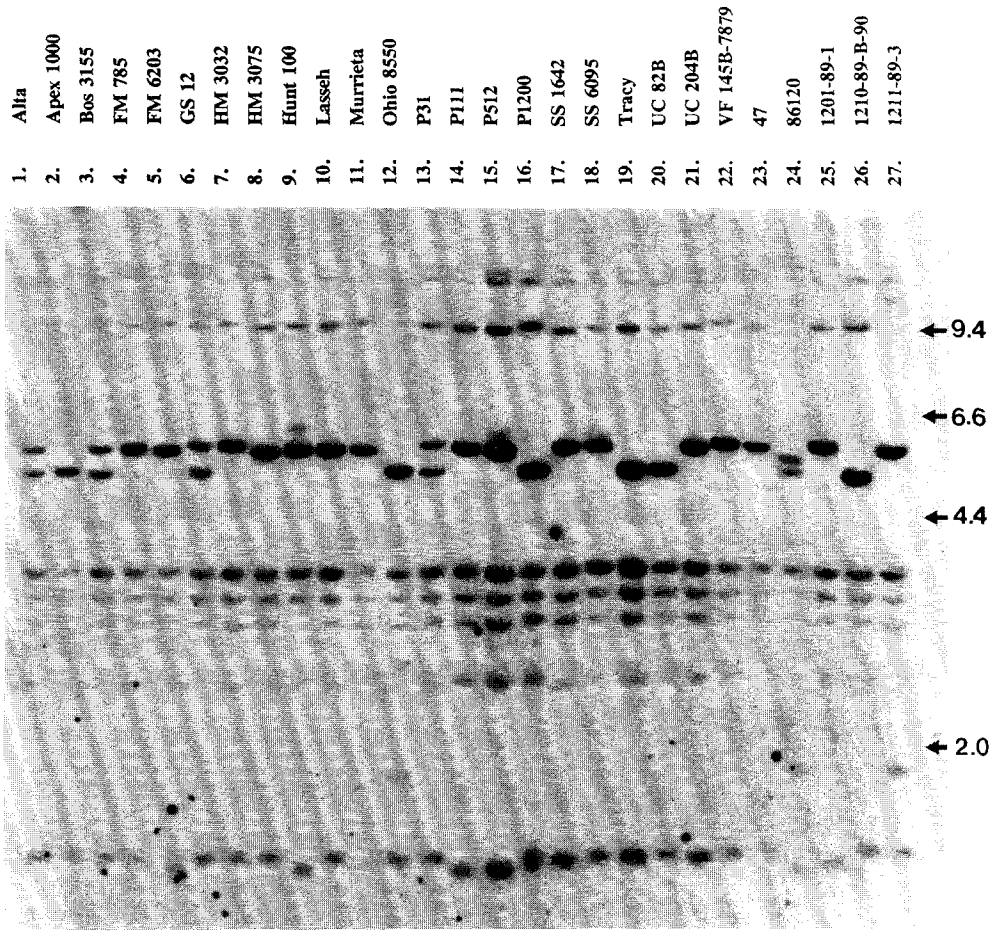

Fig. 1 Southern blot showing variation revealed by probing with oligonucleotide $(GACA)_4$ within processing tomato cultivars

from rice through homology with human minisatellite core sequences, and Akkaya et al. (1992), who have demonstrated the utility of microsatellite sequences in soybean. We wish to develop similar marker systems for tomato and report here the identification and characterisation of a 1.3-kb element, composed entirely of tetranucleotide repeats, which hybridises to a family of loci displaying considerable length variation and demonstrate that such variation can be accessed directly through PCR amplification.

Materials and methods

Plant material and southern analysis

The following accessions were obtained from C.M. Rick (University of California, Davis): *L. chmielewskii* (acc. no. LA1028), *L. hirsutum* (LA173), *L. pennellii* (LA1277), *L. esculentum var. cerasiforme* (LA1323) and *L. esculentum* land-race LA172. All other materials were supplied by A. El-Sheikh, RAGU Foods, Stockton, California. DNA was isolated from leaf material using the method of Dellaporta et al. (1983). Restricted DNA (2.5 µg) was electrophoresed through 0.8% agarose gels and transferred to Hybond N nylon membranes (Amersham) by capillary blotting.

DNA labelling and hybridisation

Purified plasmid insert DNA (30 ng) was radiolabelled with [α-^{32}P]-dCTP (110 TBq/mmole, Amersham International) using the random primed method of Feinberg and Vogelstein (1983), and oligonucleotides were end-labelled with [γ-^{32}P]ATP (Sambrook et al. 1989). Filters were pre-hybridised at 65 °C for 3 h in 6×SSPE, 0.5% dried milk and either 100 µg/ml pre-boiled salmon sperm DNA or Denhardt's solution (Sambrook et al. 1989) for plasmid insert or oligonucleotide hybridisation, respectively. Labelled probe DNA was then added and incubations continued at 65 °C overnight. Filters were washed twice in 2×SSC, 0.1% SDS at 20 °C and autoradiographed at –70 °C for an appropriate period.

Genomic cloning and sequencing

Size-fractionated (10–20 kb) partial Sau3A1 DNA fragments were ligated to partially filled XhoI arms of Lambda Gem 11 vector (Promega), packaged *in vitro* and grown on *E. coli* ER1647. Plaques were transferred onto Hybond N membranes, hybridised with $(GACA)_4$ oligonucleotide probe and phage DNA isolated from plate lysates using the LambdaSorb protocol (Promega).

Lambda DNA was digested with *Hin*fI or *Dra*I and the $(GACA)_4$-hybridising fragment (1.3 kb) blunt-ligated into a *Sma*I-cut Bluescript vector. Deletions were produced using exonuclease III (He-

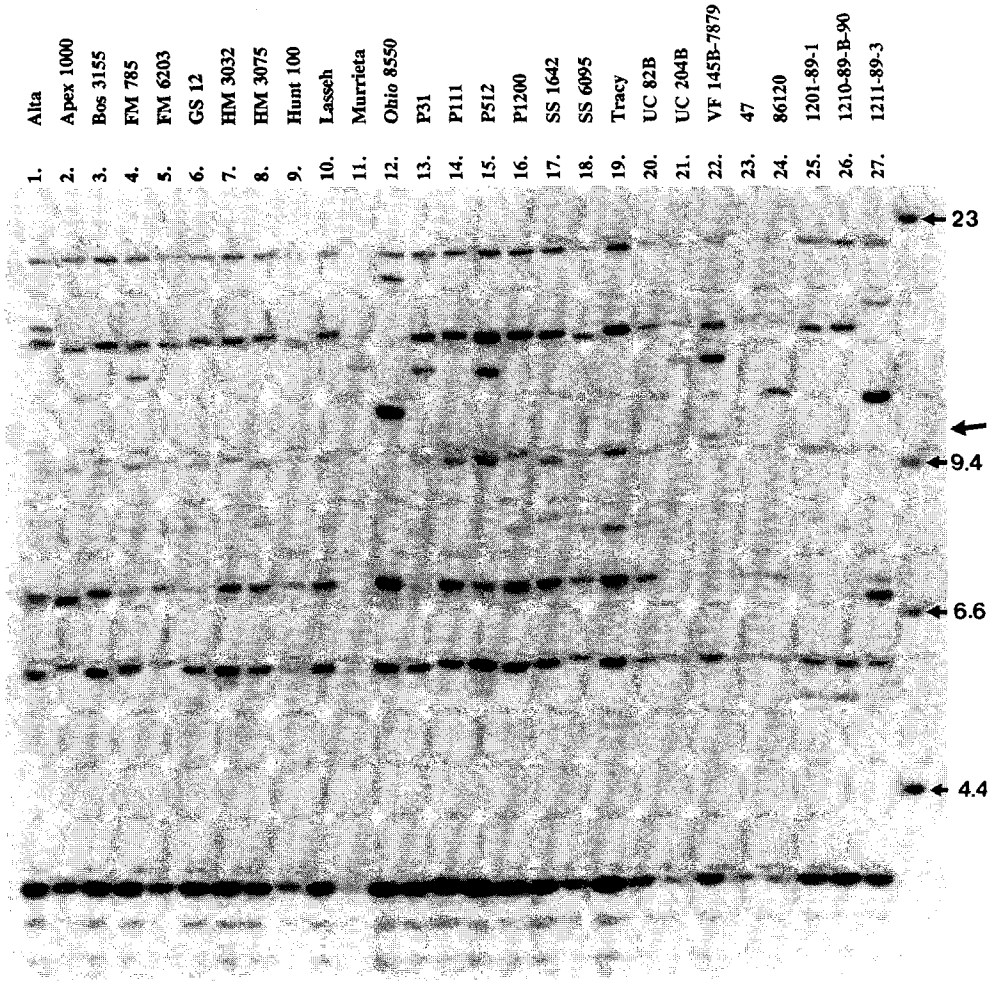

Fig. 2 Southern blot showing variation revealed by a cloned 1.3-kb tomato element containing GACA repeats (pT-GACA-2:1.3), amongst the same set of processing cultivars as shown in Fig. 1. Note that in this case the gel has run twice the distance as in Fig. 1

nikoff 1984), in some cases modified by inclusion of α-phosphorothioate dNTPs to fill 3′-recessed ends.

PCR amplification

DNA (50–200 ng) was amplified in a standard 25 µl reaction mixture containing: 0.2 mM dNTP, 0.3 µM each oligonucleotide primer, 10 mM TRIS pH 8.3, 50 m*M* KCl, 1.5 m*M* $MgCl_2$ and 0.5 units Taq DNA Polymerase (Advanced Biotech or New England Biolabs – Vent^{TM} R^{exo}). An extended 5 min 95 °C denaturation step was followed by 30 cycles of 95 °C for 45 sec, 58 °C for 60 sec, 72 °C for 60 s and a final extension at 72 °C for 10 min. Primer sequences (5′ to 3′) are as follows, P1=GGAGGTCATGGTTTTAAGGTG; P2=CTGTCTGTCAAACTGTTTGC; P3=ACAGACAGGCAAACAGATTG; P4=CTATCTGTCTGTCTATCTGG.

Results

Cloning of $(GACA)_n$ loci

A standard *Lycopersicon* genotype blot was screened with a series of simple sequence oligonucleotides as probes. One particular 16 base oligonucleotide, namely $(GACA)_4$, was found to give both a good hybridisation signal and evidence of polymorphism (Fig. 1). The $(GACA)_4$ oligonucleotide was used to screen a representative lambda library (10^5 plaques, insert size 10–20 kb), resulting in approximately 60 positively-hybridising clones. Of these, 18 were subject to plaque purification, though only 7 could be successfully recovered, possibly because of the instability of eukaryotic repeat sequences in prokaryotic hosts (Kelly et al. 1989). Lambda recombinants were digested with *Hin*fI or *Dra*I, fragments separated by electrophoresis and probed with $(GACA)_4$ and positively-hybridising fragments were then used to probe Southern genomic blots. Five gave smears or non-polymorphic patterns, whilst 2 revealed similar patterns of substantial polymorphism between cultivars. One of these fragments (1.3 kb) was subcloned into a plasmid vector, yielding the subclone pT-GACA-2:1.3. Hybridisation obtained with this probe is shown in Fig. 2. The cultivars used in this blot represent 27 of the most widely grown Californian processing types, many of which are related, and thus represent a severe test of probe discriminatory power. Nevertheless, from careful examination of bands (including the weaker but still infor-

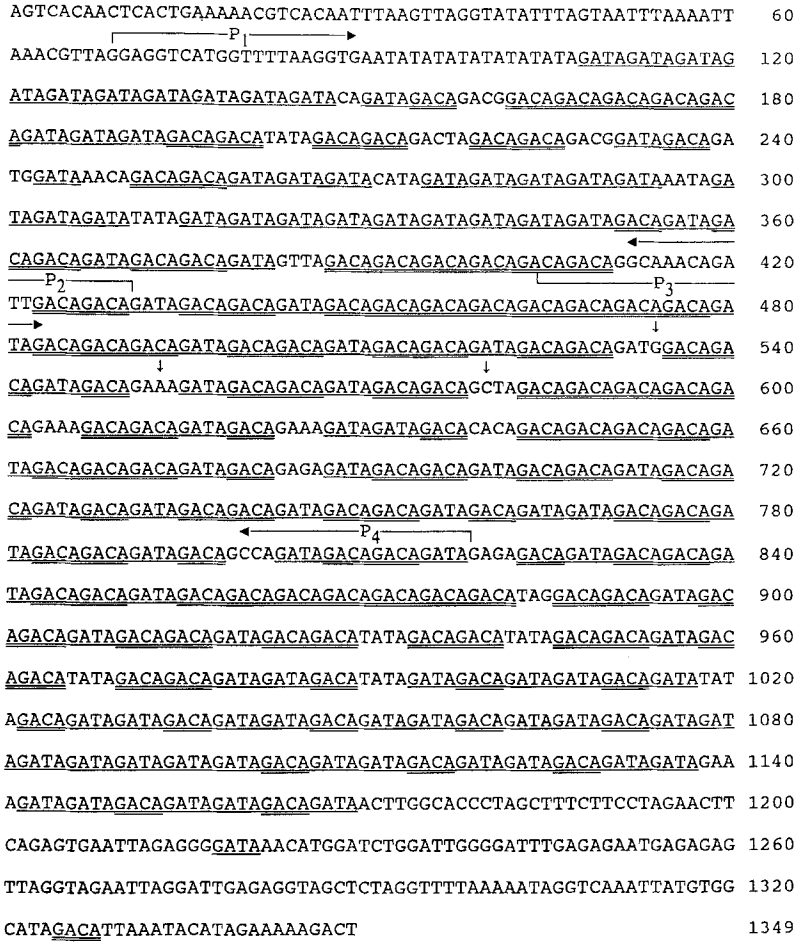

Fig. 3 DNA sequence of pT-GACA-2:1.3 element. GACA tetranucleotides are *double-underlined*, whilst GATA tetranucleotides are *single-underlined*. Examples of deviations from concensus repeats are marked by *vertical arrows*, as discussed in the text. Regions used for PCR priming are denoted by *horizontal arrows* and marked according to Fig. 4

mative ones in the 4 to 5-kb range) it is possible to unambiguously distinguish all genotypes from each other. One other notable feature of the hybridisation pattern is that bands are not randomly size-distributed through genotypes but instead appear to vary around basic lengths. For example, the region of the blot identified by the large arrow contains bands of about 10 kb in most genotypes, but lanes 20, 21 and 22 reveal variants at 9.9, 10.2 and 10.5 kb, respectively. These size differences are reproducible and most evident in cultivars where doublets are present (lanes 3, 9 and 14). This is suggestive of a tandem repeat array, with subtle length variation arising from minor changes in repeat copy number, as occurs for ribosomal DNA spacer length polymorphism in barley (Saghai-Maroof 1984).

A preliminary analysis of GACA-positive band segregation using an inter-specific F_2 population demonstrated that the majority of bands segregated independently of one another.

DNA sequence analysis of pT-GACA-2:1.3

One strand of the 1.3-kb element was sequenced in its entirety, though considerable difficulties were encountered in generating and propagating some members of the deletion set from the complementary strand. This may be a reflection of plasmid instability arising from the presence of certain regions within the array. Similar observations have been reported for rice (Daly et al. 1991). In this regard it is noteworthy that original manipulations of the 1.3-kb element were frustrated by poor growth and recovery of plasmid DNA, regardless of bacterial host (e.g. DH5 α, XL1 Blue, dam$^-$ or mcr$^-$ strains). On further culture, a spontaneous event occurred, which resulted in marked improvements to growth and yield suggesting a modification to either element, vector or host. No major change could be detected in the overall length of the insert, as occurred with mouse minisatellites (Kelly et al. 1989), though this would not exclude minor deletion events as discussed below. Transfer of the insert from the high-yielding clone to a closely-related vector resulted in re-acquisition of the unstable phenotype, suggesting that the spontaneous event may have involved the vector rather than the element.

The DNA sequence of the 1,349-bp *Hin*fI fragment is presented in Fig. 3. Apart from flanking regions of 90-bp 5′ and 180-bp 3′, virtually the entire insert is composed of tetranucleotide repeats with GACA and GATA predominating (132 GACA, 105 GATA and 10 TATA repeats). To assist alignment and reveal internal structure, GACA runs are marked in Fig. 3 by double underlining and GATA by single underlining. It can be seen that the tetranucleotide repeats are further organised into higher order structures

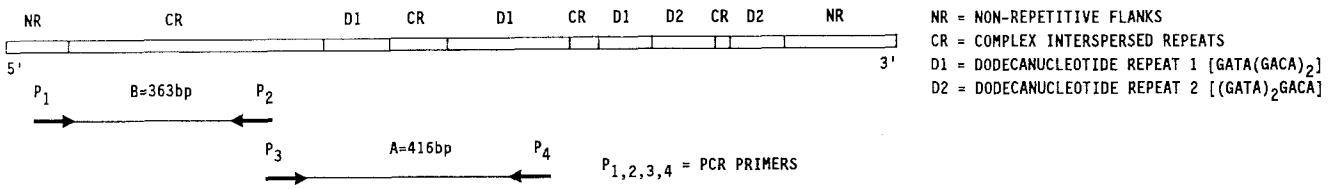

Fig. 4 Schematic representation of repetitive domains in pT-GACA-2:1.3 and the position of the primers used for PCR amplification

which themselves vary through the array; for example the region between nucleotides 495 and 578 is composed of one GATA followed by two GACA repeats, yielding a GATA $(GACA)_2$ dodecanucleotide, itself repeated seven times. This array is followed by 95 bases of less regular GACA/GATA interspersion and then the dodecanucleotide motif is further repeated (4 dodecanucleotide repeats at 675–722 bp, 11 dodecanucleotide repeats at 731–862 bp and so on). Towards the end of the element a variant dodecanucleotide of $(GATA)_2$ GACA appears (8 repeats at 978–1073 bp and 6 repeats at 1094–1165 bp).

Such simple basic repeats allow departures from the concensus to be readily identified e.g. positions 534, 553 and 579 (vertical arrows in Fig. 3), consist of A to G, C to A and A to C departures respectively from the canonical repeat and may correspond to point mutations introduced after amplification of the basic GATA $(GACA)_2$ repeat since these putative changes are unique within this dodecanucleotide array. Sequence analysis of other *L. esculentum* alleles may help confirm evolution of this array.

Curiously, this arrangement of GACA/GATA tetranucleotide repeats is capable of encoding long open reading frames (ORFs) e.g. the complementary strand to that shown in Fig. 3 contains ORFs of 371, 294 and 367 amino acids in the three reading frames commencing at positions 151, 80 and 171 bp respectively from the 5′ end. Of the 371 amino acids in the longest reading frame, over 80% are present in the canonical tetrapeptide repeat: serine, isoleucine or valine, tyrosine or cysteine, leucine. A similar coding potential has been observed by others in complements to GACA/GATA repeats (Epplen et al. 1983). In contrast, the presence of stop codons in consecutive GATA repeats results in a maximum ORF of only 72 amino acids in the top strand (i.e. that shown in Fig. 3).

PCR amplification of domains within pT-GACA-2:1.3

The presence of variant repeat arrays allows the division of the element into distinct repetitive domains (Fig. 4) and, where the boundary between domains departs sufficiently from the canonical repeat, unambiguous primers can be constructed, thus allowing PCR amplification of regions from within the element. Figure 4 shows two such regions, (marked "A" and "B"), which have been amplified from different genotypes. Figure 5 shows amplification products from region "A" analysed either by agarose (Fig. 5A) or acrylamide (Fig. 5B) gel electrophoresis. Two major allelic variants can be identified at approximately 426 and 466-bp and a third rare allele at approximately 436-bp is also present (sample 5). Curiously, the PCR amplification product from region A of the genotype sequenced is 10-bp longer than the predicted size from sequencing (426 and 416 bp respectively), which may reflect a deletion event during cloning. Lane 15 contains a F_1 hybrid, where both major alleles can be seen as a dimer. Figure 5B shows amplification of region B on an acrylamide gel, and again two major variants are present (363 and 379 bp) as is a third rare variant (367 bp). The size differences in this case suggest amplification/deletion through intact tetranucleotide or dodecanucleotide units. Length polymorphism in regions "A" and "B" divides the cultivars in an identical fashion and may have implications concerning allelic selection at this locus during domestication of the species. It is noteworthy that primer 4 departs from the canonical GATA$(GACA)_2$ sequence by a single A to C transition at the 3′ end, yet this is sufficient to specify amplification of this precise region of this particular GACA-repeat locus.

Discussion

Structure of a GACA-hybridising locus and practical utility as a fingerprinting probe

We have reported here the identification of a family of elements within the tomato genome all of which share hybridisation with the simple sequence repeat $(GACA)_4$, and display substantial length polymorphism. Cloning of one family member (pT-GACA-2:1.3) yields a stable DNA probe for estimating the degree and nature of this length variation. Typically, only 7% of low copy *Pst*I clones display RFLPs when assessed against a battery of tomato genotypes (Miller and Tanksley 1990), however the majority of the readily resolvable, high molecular weight bands hybridising to pT-GACA-2:1.3 show polymorphism (Fig. 2). The combination of multiple bands, most of which are independently-segregating, together with high levels of variability, yields a probe with considerable discriminatory power. Even a highly restricted set of genotypes, such as the Californian processing cultivars shown in Fig. 2, display substantial banding differences and, with careful analysis, all can be distinguished from each other, thus demonstrating practical utility as a genetic fingerprinting probe.

PCR amplification of contiguous internal domains in a series of genotypes reveals length polymorphism (Fig. 5), indicating a dynamic structure capable of variation at multiple points. Such PCR amplification exploits sequence

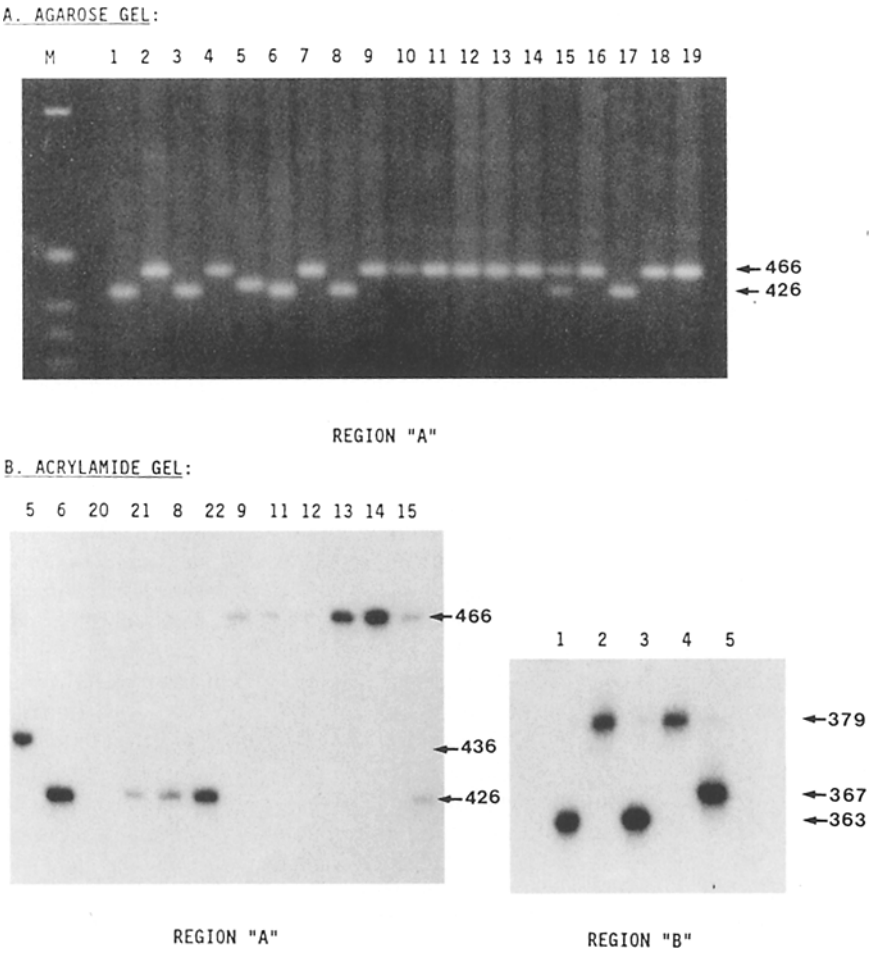

Fig. 5 A, B PCR amplification products. **A** Agarose gel separation of products from *region A* of Fig. 4. *Lanes 1–19* contain cvs 'FM6203' (*1*), 'SS6095' (*2, 10*), 'P111' (*3, 8*); 'H7998' (*4*); 'Red Cherry Small' (*5*), 'Castlong UCG' (*6*), '47' (*7*), 'Ohio 8550' (*9*),'1211-89-3' (*11*), '1210-89-B-90' (*12*), 'G512' (*13*), 'Alta' (*14*), 'P1200' (*15*), 'HM3032' (*16*), 'P512' (*17*), 'P31' (*18*), 'HM3075' (*19*). **B** Acrylamide gel separation of *region A* and *region B* products from some of the cultivars shown in Fig. 5 A. *Lanes 20–22* contain cvs 47, 1201-89-1 and 86120, respectively

variation at the boundaries of these domains and allows conversion from a multi-locus RFLP probe to single-locus PCR analysis, thus yielding a simply-interpretable, high-throughput screening system. This has immediate practical applications, for example the purity assessment of F_1 hybrid seed lots.

Relationship of pT-GACA-2:1.3 to other repeated sequences

Variable tandem repeat arrays used in mammalian genotyping fall into two major classes: either long minisatellite arrays, cloned members of which typically contain 15- to 60-bp units repeated 10–50 times (Jeffreys 1985), or much shorter microsatellite arrays of 2- to 4-bp units, repeated 15–40 times (Love et al. 1990). The tomato element reported here shares features of both, in that it can either be thought of as a series of contiguous microsatellites, each based upon a short (4-bp) repeat, or minisatellite-like, both in overall length (1.3 kb) and the presence of higher order (12-bp) repeats. Such descriptions may be quite arbitrary however, in that it is possible that those microsatellites and minisatellites used in genotyping merely represent extremes of a continuum of structures and that the element reported here represents an intermediate form. Mapping and sequence analysis of other members of the GACA-hybridising family (particularly the longer elements) may provide insight into such relationships.

There have been a number of studies where hypervariable genetic elements have been identified in plant genomes using probes previously utilised in animal work, e.g. the M13 repeat (Nybom 1990), a human 33.6-bp polycore (Rogstad et al. 1988) as well as simple sequence repeats (Weising et al. 1989, Vosman et al. 1992), though there have been few reports where such loci have been characterised in plants at the sequence level. One exception is the study of Daly et al. (1991) who, following the observations of Dallas (1988) that human minisatellite core repeats hybridise to variable domains in rice, cloned a number of such loci and found that the variable domains were always associated with an open reading frame of unknown function. They speculate that linkage of repeats with coding domains may allow interplay between members of multigene families at non-homologous loci. They also observe that, although the rice minisatellite domain shares homology with a 16-bp human core repeat, it possesses an even simpler cryptic repeat of $(GGN)_n$, possibly reflecting a simple sequence origin.

A curious feature of the element described here is the intimate association of GACA and GATA repeats. This arrangement is likely to be reiterated at other loci in tomato,

since we and others (Vosman et al. 1992) have found that the $(GACA)_4$ hybridisation pattern is very similar to that obtained with $(GATA)_4$. This association of GACA and GATA repeats has been observed in a variety of different eukaryotes. For example Epplen et al. (1982) reported that a 2.5-kb fragment of sex-specific satellite DNA from a Colubrid snake contained 26 and 12 copies respectively of GATA and GACA repeats. This element also hybridises to mammalian genomes where it was again found capable of generating sex-specific hybridisation patterns; furthermore it identified transcribed regions of mouse DNA, one of which contained 130 interspersed GACA and GATA repeats in a long open reading frame (Epplen et al. 1983). Similar findings have been reported for *Drosophila* (Singh et al. 1984). The widespread occurrence of GACA/GATA repeats may reflect either selection for some hitherto unidentified function, or simply that amplification of such sequences occurs at high frequency in most or all eukaryotic, though not prokaryotic, genomes. Evidence for the proclivity of such sequences to evolve rapidly comes from neoplastic transformation studies where it has been shown that human gliomas are frequently associated with altered simple sequence repeat hybridisation patterns (Nürnberg et al. 1991). Such a predisposition of GACA/GATA repeats to amplification may explain the high level of polymorphism described here for tomato. Further studies are now in progress to establish the stability of such loci within defined tomato cultivar lineages.

Acknowledgements We wish to thank Adel El Sheikh, Van den Bergh Foods, Stockton, California, for support and advice during the course of this work and Rebecca Stratford, Plant Breeding International Cambridge, for valuable advice on DNA sequencing.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132: 1131–1139

Beyermann B, Nürnberg P, Weihe A, Meixner M, Epplen JT, Börner T (1992) Fingerprinting plant genomes with oligonucleotide probes specific for simple repetitive DNA sequences. Theor Appl Genet 83:691–694

Broun P, Ganal MW, Tanksley SD (1992) Telomeric arrays display high levels of heritable polymorphism among closely related plant varieties. Proc Natl Acad Sci USA 89:1354–1357

Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) Genetic maps of wheat homoeologous group 7 chromosomes. Theor Appl Genet 78:495–504

Dallas JF (1988) Detection of DNA "fingerprints" of cultivated rice by hybridization with a human minisatellite DNA probe. Proc Natl Acad Sci USA 85:6831–6835

Daly A, Kellam P, Berry ST, Chojecki AJS, Barnes SR (1991) The isolation and characterisation of plant sequences homologous to human hypervariable minisatellites. In: Burke T, Dolf G, Jeffreys AJ, Wolff R (eds) DNA fingerprinting: approaches and applications. Birkhäuser Verlag, Basel, Switzerland, pp 330–340

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep (vol 1) 4:19–21

Epplen JT, McCarrey JR, Sutou S, Ohno S (1982) Base sequence of a cloned snake W-chromosome DNA fragment and identification of a male-specific putative mRNA in the mouse. Proc Natl Acad Sci USA 79:3798–3802

Epplen JT, Cellini A, Romero S, Ohno S (1983) An attempt to approach the molecular mechanisms of primary sex determination: W- and Y-chromosomal conserved simple repetitive DNA sequences and their differential expression in mRNA. J Exp Zoo 228:305–312

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Helentjaris T, King G, Slocum M, Siedenstrang C, Wegman S (1985) Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant Mol Biol 5:109–118

Henikoff S (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351–359

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. Nature 314:67–73

Jeffreys AJ, MacLeod A, Tamaki K, Neil DL, Monckton DG (1991) Minisatellite repeat coding as a digital approach to DNA typing. Nature 354:204–209

Keim P, Shoemaker RC, Palmer RG (1989) Restriction fragment length polymorphism diversity in soybean. Theor Appl Genet 77:786–792

Kelly R, Bulfield F, Collick A, Gibbs M, Jeffreys AJ (1989) Characterization of a highly unstable mouse minisatellite locus: evidence for somatic mutation during early development. Genomics 5:844–856

Love JM, Knight AM, McAleer MA, Todd JA (1990) Towards construction of a high resolution map of the mouse genome using PCR-analysed microsatellites. Nucleic Acids Res 18:4123–4130

Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor Appl Genet 80:437–448

Nürnberg P, Barth I, Fuhrmann E, Lenzner C, Losanova T, Peters C, Pöche H, Thiel G (1991) Monitoring genomic alterations with a panel of oligonucleotide probes specific for various simple repeat motifs. Electrophoresis 12:186–192

Nybom H, Schaal BA (1990) DNA "fingerprints" applied to paternity analysis in apples (*Malus* x *domestica*). Theor Appl Genet 79:763–768

Rogstad SH, Patton JC, Schaal BA (1988) A human minisatellite probe reveals RFLPs among individuals of two angiosperms. Nucleic Acids Res 16:11378

Sahar-Maroof MA, Soliman KM, Jorgenson RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Singh L, Phillips C, Jones KW (1984) The conserved nucleotide sequences of Bkm, which define *Sxr* in the mouse, are transcribed. Cell 36:111–120

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Vosman B, Arens P, Rus-Kortekaas W, Smulders MJM (1992) Identification of highly polymorphic DNA regions in tomato. Theor Appl Genet 85:239–244

Weising K, Weigand F, Driesel AJ, Kahl G, Zischler H, Epplen JT (1989) Polymorphic simple GATA/GACA repeats in plant genomes. Nucleic Acids Res 17:10128

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. Nature 359:794–801