

Evaluation of restriction fragment length polymorphism in *Cucumis melo*

S.L. Neuhausen *

NPI, 417 Wakara Way, Salt Lake City, UT 84108, USA

Received August 5, 1990; Accepted June 11, 1991 Communicated by J. Beckmann

Summary. The objectives of this study were to assess the degree of restriction fragment length polymorphism (RFLP) in *Cucumis melo* and to determine interrelationships among cultivated varieties. Initial screening of a genomic *PstI* library revealed that approximately 40% of the clones were repetitive. A total of 162 unique and low-copy sequence clones were hybridized to seven diverse accesions of *C. melo* and *a C. sativus* cultivar 'Pacer' to evaluate RFLP variation. Of these, 130 probes (80%) detected a polymorphism between *C. melo* accessions and *C. sativus,* and the majority were polymorphic with more than one enzyme digest. In contrast, only 53 probes (33%) were useful in differentiating at least one of the seven accessions. Of those, only 9% were informative with more than one enzyme digest. This indicates that within *C. melo,* the differences among accessions are due to infrequent base substitutions, whereas between the two species, differences are mainly due to genome rearrangements such as insertions and deletions or numerous base substitutions. Of the informative probes, 34 were used in analyzing 44 *C. melo* lines to establish a data base of RFLP hybridization patterns. Percent similarity based on RFLP profiles was computed among lines and analyzed by principal component analysis, to visualize relationships among lines. There were clear demarcations among, but not within, muskmelon and honeydew groups.

Key words: Restriction fragment length polymorphism (RFLP) - *Cucumis melo -* Muskmelon - Honeydew melon

Introduction

Restriction fragment length polymorphisms (RFLPs) are differences in hybridization patterns among DNAs revealed after digestion with a restriction endonuclease and subsequent probing with a labelled DNA clone. Tanksley (1983) and Soller and Beckmann (1983) have reviewed the potential usefulness of RFLPs in plant improvement studies, as well as in basic plant genetic research. Many uses of RFLPs are the same as for isozymes. The primary advantage of RFLPs over isozymes is their potentially unlimited number.

The usefulness of RFLP markers is largely dependent on the degree of polymorphism. RFLP variation has been detected in a wide range of crops (Helentjaris and Burr 1989). Levels of variability vary greatly among species, with allogamous species generally exhibiting more polymorphism than autogamous species.

Isozyme variability has been examined in *Cucumis melo* L. populations. Esquinas-Alcazar (as reported in Dane 1983) examined 125 *C. melo* populations and six enzyme systems encompassing 11 isozyme loci. For five of the isozymes, there was no variation and the other six were dimorphic. They both concluded that little or no allelic variation was found at most of the loci in all of the populations studied. Staub et al. (1987) examined 30 accesions of *C. melo* from Iran with nine isozymes. There was no variation for five of the isozymes, two were trimorphic, and the two *Pgm* loci were dimorphic. Perl-Treves et al. (1985) examined five *C. melo* varieties and one wild variety with 29 nuclear-coded enzymes. They found 22 of the isozymes to be monomorphic. Of the other 7 isozymes, 6 were dimorphic and FDP was trimorphic. These studies indicate that there is little isozymic variation in *C. melo* germ plasm. Since there are only a finite number of isozymes that can be examined in order

^{*} Current address: Biology Department, Slippery Rock University, Slippery Rock, PA 16057, USA

to find sufficient variation in molecular markers for establishing linkage maps and for breeding, more markers need to be identified. RFLP markers are potentially unlimited and the number evaluated is only limited by the time and resources of the researcher. Evaluation of RFLP variation may allow us to identify sufficient variation in *C. melo* for further phylogenetic research and for breeding studies. The objectives of this study were to assess the variation in RFLPs within *C. melo* and to determine interrelationships among cultivated varieties.

Materials and methods

Preparation of clones

Genomic library. A partial genomic DNA library was prepared from 10-day-old leaf tissue DNA of the muskmelon 'PMR45.' Total DNA was digested with the endonuclease restriction enzyme *PstI* (BRL) and electrophoresed in a 0.7% agarose gel. DNA fragments 500-2,000 bp in length were eluted from the gel and subcloned into pUC19 plasmid vector, which had been previously digested with *PstI* and treated with calf-intestinal alkaline phosphatase. The recombinants were used to transform JM83, and the resulting colonies were stored in Luria broth with glycerol at -70 °C. Highly repetitive sequences were identified by colony lifts probed with $32P$ -radiolabelled total genomic DNA (Maniatis et al. 1982) and were not examined further.

cDNA library. Total RNA was prepared by homogenization of fresh-frozen leaf tissue in 50 mM TRIS-HCl (pH 8.8), 150 mM NaCl, 5 mM EDTA, and 5% SDS, followed by several phenol and phenol-chloroform-isoamyl extractions and ethanol precipitation (King et al. 1988). The poly $A+$ fraction was isolated by chromatography over oligo-dT cellulose (Aviv and Leder 1972). The cDNA was prepared using the BRL cDNA synthesis system and then blunt-end ligated into pUC19 plasmid vector, which had been previously digested with *SmaI* (BRL) and treated with calf-intestinal alkaline phosphatase. As with the genomic library, recombinants were used to transform JM83 and the resulting colonies were stored in freezing media at -70 °C.

Plant DNA preparation, restriction enzyme digestion, and blotting

Leaves from 24 plants per accession were harvested 10 days after germination and lyophilized. Preparation of total genomic DNA was as described by Lichtenstein and Draper (1985), with modifications to increase the starting mass of plant material and the volume of reagents. Procedures for digestion by restriction enzymes, agarose gel electrophoresis, and Southern blotting have been described previously (Helentjaris et al. 1985). Plant DNAs extracted from *C. melo* germ plasm accessions were digested with *EcoRI, HindIII,* and *EeoRV.* Nylon membrane used as the blotting matrix was obtained from MSI.

Preparation of radioactive probes and hybridization

Procedures for DNA plasmid minipreparation and nick-translation of genomic probes, as well as for the prehybridization, hybridization, washing, and autoradiography of blots were as previously described (Helentjaris et al. 1985). DNA plasmid minipreparations were also made by the alkaline lysis method (Maniatis et al. 1982). Inserts were excised from low-meltingpoint agarose gels. For the cDNA plasmid minipreparations, inserts were labelled using the random primer labelling method of Feinberg and Vogelstein (1983, 1984).

Clones that detected polymorphisms on screening blots were prepared using a large-scale plasmid procedure as previously described (Helentjaris et al. 1985).

Cucumis melo germ plasm used in assessing polymorphism

For the initial screening of clones to determine which were effective in detecting polymorphisms, eight accessions were examined, including *a C. sativus* cultivar 'Pacer.' The seven C. *melo* lines were 'PMR45,' 'Topmark,' Pl190184 (Mexico), PI124113 (India), PI357775 (Yugoslavia), PI267085 (U.S.S.R), and PI323427 (Austria). The pool of U.S. muskmelon cultivars is narrow (J. McCreight, personal communication) and isozyme studies revealed little polymorphism (Esquinas-Alcazar as reported in Dane 1983; Perl-Treves et al. 1985; Staub et al. 1987). Therefore, in order to increase the opportunity of identifying clones that detected polymorphisms, accessions from outside the United States were used. A clone-enzyme combination was considered informative in detecting polymorphism if different hybridization patterns were observed among the accessions being compared. Once informative clone-enzyme combinations were identified, 44 *C. melo* lines were evaluated (Table 1) to form a data base (an additional four lines were examined but deleted due to missing data).

Analysis of data base

A total of 34 probes were used to map the 44 *C. melo* lines. For each probe, absence of a fragment was assigned a 0, presence of only that fragment was assigned a 1, and presence of that fragment and a second fragment was assigned a 0.5 for each fragment. Accordingly, each fragment was a variable for a total of 87 variables for the 34 probes with the 44 *C. melo* lines.

Percent similarity for every pair of entries was computed as follows. For each probe, a pair of entries was assigned 0 if it had no fragment in common, 0.5 if the two entries shared a fragment, and 1.0 if the fragment pattern was identical. For each pair of entries, these values were then summed across all the probes, divided by the number of probes, and multiplied by 100 to obtain percent similarity. An example is shown below.

Percent similarity for:

Lines 1 and 2: $(1 + 0.5 + 0)/3 \times 100 = 50\%$

Lines 1 and 3: $(0+0+1)/3 \times 100 = 33\%$

Lines 2 and 3: $(0+0.5+0)/3 \times 100 = 17\%$.

The computed percent similarity for each pair was placed in a matrix of each line to every other line. Principal component analysis was then performed on this matrix (SAS 1985). The first several principal components were plotted to visualize the associations among accessions.

Entry name	Source	Entry name	Source
Muskmelons		Honeydew melons	
1 AR5 4 Aurora 6 Burpee Hybrid	J. McCreight Asgrow Burpee	3 Asgrow Honeydew 19 Honeydew Green 25 Limelight	Asgrow Harris Moran Burpee
7 Bushstar Cinco 11 Early Dawn	Burpee J. McCreight Harris Moran	38 Tam Dew Improved 43 Venus	Harris Moran Burpee
12 Edisto 47 13 Explorer	J. McCreight Northrup King	Casaba melons	
16 Hales Best Jumbo 17 Hilene 18 Honeybush	J. McCreight Asgrow Burpee	15 Golden Beauty 22 Juanes de Canary	PetoSeed Harris Moran
20 Intrepid 21 Itsy Bitsy Sweetheart Bush $23 L-20$	Northrup King SunSeeds G. Elmstrom	Middle East melons	
24 Laguna 26 Magnum .45 27 Mission 30 Perlita	Asgrow PetoSeed Asgrow J. McCreight	5 Birdnest 14 Gavlia 28 Ogen	J. McCreight Thompson & Morgan Thompson & Morgan
31 PMR45 32 PMR5 34 Samson 35 Saticoy 36 Shipmaster 37 Superstar	J. McCreight J. McCreight SunSeeds PetoSeed Northrup King Harris Moran	Ananas melons 2 Ananas 8 Cameo	SunSeeds Thompson & Morgan
39 Tania 40 Topmark 41 Topnet S.R. 42 Topscore 44 W-6	Asgrow J. McCreight Harris Moran PetoSeed J. McCreight	Other types 10 Crenshaw 29 PI414723 33 Resistant Joy	J. McCreight, USDS G. Elmstrom, Univ. Fla. Thompson & Morgan

Table 1. *Cucumis melo* accessions used in assessing the level of restriction fragment length polymorphisms

Results

Degree of polymorphism among seven Cucumis melo accessions and 'Pacer'

A total of 500 genomic clones were screened by colony hybridization, and 335 of these clones appeared to contain low-copy-number DNA sequences. Of these putative low-copy-number clones, 244 were probed against Southern blots of genomic DNAs of the eight screening lines, and 29 had hybridization patterns characteristic of repeated sequences. Thus, approximately 40% of the genomic melon library probes were high-copy-number sequences.

A total of 138 genomic and 24 cDNA clones were subsequently used to determine the degree of polymorphism among the seven *C. melo* accessions and the C. *sativus* accession. These 162 clones showed strong hybridization to a small number of fragments, fulfilling the criteria of a useful clone, as outlined by Helentjaris et al. (1986). Twenty-seven probes (17%) detected no polymorphisms among the eight accessions. One hundred thirty probes (80%) detected a fragment in the *C. sativus* variety 'Pacer' not present in any of the seven *C. melo* accessions. 'Pacer' had the same hybridization pattern as one of the'C, *melo* accessions for only five probes.

Of the 162 probes, 53 (33%) distinguished at least one of the seven *C. melo* accessions. Among the *C. melo* accessions, 45 of the 53 probes (85%) identified two hybridization patterns, and the remaining 8 (15%) exhibited three patterns. For 27 of the 53 informative probes (51%), only one of the seven *C. melo* lines had the second hybridization pattern. For an additional 8 clones (15%), only two of the lines showed the second hybridization pattern. Among the polymorphic clones, the majority of the polymorphisms were two banding patterns and one pattern was predominant among the lines. Approximately half of the polymorphisms were from two plant introductions. PI124113 had a unique restriction pattern for 15 of the probes, PI357775 was unique for 10 of the probes, PI267085 and PI323427 were unique for 5 probes each, and PI190184 was unique for 3 probes. Only 3 probes (2%) distinguished the two U.S. cultivars 'Topmark' and 'PMR45.'

Although three enzymes were used, restriction fragment length differences were observed with only a single enzyme for 48 of the 53 probes distinguishing *C. melo* accessions. Twenty-six of the informative probes (49%) exhibited a polymorphism with the enzyme *EcoRI.* In contrast, the majority of the probes was useful in distinguishing 'Pacer' from the *C. melo* accessions with all three enzyme digests.

Fig. 1. Plot of principal components 1 and 2 for 44 *Cucumis melo* accessions (names associated with entries are in Table 1). Clusters were drawn to highlight groupings of honeydews and muskmelons

Table 2. Number of combination of hybridization patterns among 44 *C. melo* accessions for each probe

No. of patterns	No. of probes	% Total probes
2	14	41.2
3	8	23.5
4		20.6
		8.8
6		2.9
		2.9

Degree of polymorphism among cultivated accessions of C. melo

A total of 34 of the informative probes were individually hybridized to the 44 *C. melo* lines to build a data base of RFLP hybridization patterns. The accessions were chosen to include a wide range of phenotypes, including open-pollinated and hybrid muskmelons, an unadapated plant introduction, five honeydew melons, three Middle Eastern lines, and European varieties of Casaba and Ananas (Table 1).

Hybridization with 20 of the 34 probes (59%) resulted in only two hybridization patterns. For the remaining hybridizations, 9 of the probes (26%) showed three hybridization patterns and 5 (14%) showed four hybridization patterns. With two, three, and four fragments, there are three, six, and ten possible combinations of those fragments, respectively. Only a small number of these were represented (Table 2), with 86% of the clones identifying two to four classes. Examination of the degree of polymorphism, as measured by the frequency of fragments across accessions for each probe, showed that the variation again appeared to be limited. One fragment was dominant, with a frequency greater than 0.80 for 52% of the probes.

Principal component analysis was performed to visualize relationships among 44 of the melons. Seventy-five percent of the total variance was explained by the first two principal components (Fig. 1). PI414723 (29) was clearly an outlier from the other *C. melo* accessions. It was an elongate melon with yellow skin with no true rind and white, mealy flesh. All the other melons had true rinds. The five honeydew melons clustered at the top of the figure. The other non-muskmelon types (mostly with light rind and flesh colors) clustered loosely away from the muskmelons and honeydews. Within this group, there was no clustering based on size, type of melon, or geographical area. For example, Juanes de Canary (22) and Golden Beauty (15) are both casaba melons, yet they did not appear to be closely related based on this analysis. There was no clear demarcation among muskmelon types in the cluster, i.e., based on geographical markets. When the muskmelons were analyzed alone, there was still little separation of the lines (data not presented).

With ten principal components, 97% of the variance was explained. The third principal component explained 6%, the fourth explained 5%, and the fifth explained 4% of the variation. Thus, there was little information gained by including more of the components. A three-dimensional figure of principal components 1, 2, and 3 was constructed, but the demarcation of lines was no different than that with only the first two components included (figure not presented).

Discussion

Approximately 40% of the genomic DNA probes screened in the *PstI* library were repetitive under our hybridization conditions. In other plant species, such as maize and *Brassica,* digestion with a methylation-sensitive enzyme in the cloning procedure has resulted in a significant enrichment of single-copy sequences (Helentjaris et al. 1985; Figdore et al. 1988). The most likely explanation for the high proportion of repetitive clones is the large amount of satellite DNA present in melons. At 10 days after germination, when we harvested leaves for DNA extraction, leaf tissue contained approximately 25% satellite DNA (Pearson et al. 1974). The remainder of the repetitive DNA is likely ribosomal and organellar DNA.

Of the 162 low-copy-number *C. melo* probes examined, 153 (94%) hybridized to the *C. sativus* cultivar 'Pacer,' indicating general homology among the two species. For 13 of these probes (8%), hybridization was detected, but the signal strength was reduced relative to that of the *C. melo* accessions, suggesting differences in sequence-copy number or reduced homology of those regions between *C. sativus* and *C. melo.* For 9 of these probes (6%), hybridization to Pacer fragments was not detected, indicating little homology of the sequence probes to 'Pacer' (see example in Fig. 2). For the majority of the polymorphic probes, the polymorphism occurred with all three enzymes, suggesting rearrangements rather than base pair substitutions. Differences in chromosome number *(C. sativus* is $x = 7$ and *C. melo* is $x = 12$) and genome size [C. *melo* has 0.95 pg per haploid nucleus and *C. sativus* has 1.35 pg (Ingle et al. 1975)] also suggest that the majority of polymorphisms between the two species are due to insertions, deletions, or rearrangements, rather than base pair substitutions. Probes that were single copy in *C. melo* were also single copy in *C. sativus.*

RFLPs are useful for studying phylogenetic relationships among related species. For example, relationships among morphotypes in *Brassica rapa* and *B. oleracea* were more closely defined using RFLPs than those based on comparative morphology, isozymes, and seed proteins (Song et al. 1988). Based on the ease of distinguishing differences between *C. melo* and *C. sativus,* RELPs could be used to more closely examine phylogenetic relationships among *Cucumis* species at a nuclear level based on fragment length differences and homology differences of probes, as well as to confirm previous studies. Using 18 enzymes, relationships among cross-compatible wild diploid species of *Cucumis* were classified biochemically (Staub et al. 1987); however, relationships within species were not well defined. Analysis of RFLPs may allow increased resolution of genotypes within each species.

Within *Cucumis melo,* polymorphism was generally detected with only one enzyme (91%), indicating that differences among accessions were associated with infrequent base pair substitutions rather than insertions, deletions, or abundant base pair substitutions. As further evidence, in an examination of sequence variation among eight melon accessions, there were only two base pair changes out of 1,572 base pairs and no rearrangements (Shattuck-Eidens et al. 1990). In contrast, in maize, where there is a high level of RFLP variation, irrespective of number of enzymes used (Helentjaris et al. 1985), there were 92 base changes and 44 insertion/deletions in the 1,809 bases sequenced (Shattuck-Eidens et al. 1990).

The level of polymorphism detected in this study was sufficient for data base analysis. Data base information allows us to classify germ plasm into related groups with-

Fig. 2. Autoradiograph of a Southern blot of EcoRI-digested total genomic DNA from 23 *Cucumis melo* accessions and 'Pacer.' The probe did not hybridize to the DNA from *C. sativus* 'Pacer'

384

out benefit of pedigree information. For muskmelons, the majority of the varieties are privately developed so that pedigree information is not available. Even when pedigree information is available, we can assay the effects of selection and pick out lines related or unrelated at specific sites. We were interested in categorizing the muskmelons to see whether or not we could distinguish Western shippers from Eastern types, and hybrids from inbreds. Within muskmelons, there were no widely spaced subgroupings. For these lines, phenotypic discrimination was as effective as using RFLPs. Hale's Best Jumbo is thought to be the progenitor of many current U.S. melon cultivars (J. McCreight, personal communication). It is therefore likely that there was a significant reduction in population size leading to a reduction in genetic variation. The honeydew, non-muskmelon, and muskmelon groups were readily distinguished from one another and PI414723 was clearly different, which indicates that there is adequate genetic variation within *C. melo,* but possibly not in narrow germplasm pools within U.S. muskmelon and honeydew. The majority of the probe-enzyme combinations exhibited two or three hybridization patterns. With this limited range of polymorphism, there were sufficient differences among hybridization patterns when summed across probes that lines of different melon types could be readily identified.

Among cultivated varieties of *C. melo,* there is a wide range of phenotypic variation in fruit size, shape, and color. In addition, there are sources of resistance to viruses and other diseases that can be exploited from less developed material (Deakin et al. 1971). With the development of a linkage map, RFLPs could be used to characterize and develop (based on QTL associations) speciality melons for which adequate variation has been identified, and to introgress disease resistance from the less developed melons (i.e., PI414723) into cultivated varieties. To generate a sufficient number of useful polymorphic clone-enzyme combinations for linkage map construction, it may be necessary to increase the number of restriction endonucleases used in screening experiments to compensate for the paucity of base pair substitution and/or to use RAPD markers (Williams et al. 1990).

Acknowledgements. I wish to thank J. Nienhuis for his help in developing the percent similarity equation. I also appreciate the contribution of J. Tivang in preparing the figures.

References

Aviv H, Leder P (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc Natl Acad Sci USA 69:1408-1414

- Dane F (1983) Cucurbits. In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, Part b. Elsevier, Amsterdam, pp 369-390
- Deakin JR, Bohn GW, Whitaker TW (1971) Interspecific hybridization in *Cucumis.* Econ Bot 25:195-211
- Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6
- Feinberg AP, Vogelstein B (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity, addendum.. Anal Bioehem 137:26
- Figdore SS, Kenard W, Song KM, Slocum MK, Osborn TC (1988) Assessment of the degree of restriction fragment length polymorphism in *Brassiea.* Theor Appl Genet 75: 833- 840
- Helentjaris T, Burr B (eds) (1989) Development and application of molecular markers to problems in plant genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor/NY
- 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
- Helentjaris T, Slocum M, Wright S, Schaefer A, Nienhuis J (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor Appl Genet 72:761-769
- Ingle J, Timmis JN, Sinclair J (1975) The relationship between satellite deoxyribonucleic acid, ribosomal ribonueleic acid, gene redundancy, and genome size in plants. Plant Physiol 55:496-501
- King G, Turner V, Hussey C, Wurtele E, Lee M (1988) Isolation and characterization of a tomato eDNA clone which codes for a salt-induced protein. Plant Mol Biol 10:401-412
- Lichtenstein C, Draper J (1985) Genetic engineering of plants. In: Glover DM (ed) DNA cloning, vol II. A practical approach. IRL Press, Oxford Washington DC, pp 67-119
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor/NY
- Pearson GG, Timmis JN, Ingle J (1974) The differential replication of DNA during plant development. Chromosoma 45:281-294
- Perl-Treves R, Zamir D, Navot N, Galun E (1985) Phylogeny of *Cucumis* based on isozyme variability and its comparison with plastome phylogeny. Theor Appl Genet 71:430-436
- SAS (1985) SAS user's guide: statistics, Version 5. SAS Institute Inc., Cary/NC, pp 956
- Shattuck-Eidens DS, Bell RN, Neuhausen SL, Helentjaris T (1990) DNA sequence variation within maize and melon: observations from PCR amplification and direct sequencing. Genetics 126:207-217
- Soller M, Beckmann JS (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor Appl Genet 67:25-33
- Song KM, Osborn TC, Williams PH (1988) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). Theor Appl Genet 76:593-600
- Staub JE, Frederick L, Marty TC (1987) Electrophoretic variation in cross-compatible wild diploid species of *Cucumis.* Can J Bot 65:792-798
- Tanksley SD (1983) Molecular markers in plant breeding. Plant Mol Biol Rep $1:3-8$
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531- 6535