Acta Bot. Neerl. 41(4), December 1992, p. 397–406

Repetitive sequences are valuable as molecular markers in studies of phylogenetic relationships within the genus *Cucumis*

U. ZENTGRAF, K. KING and V. HEMLEBEN*

Department of Genetics, Biological Institute, University of Tübingen, Auf der Morgenstelle 28, W-7400 Tübingen, Germany

SUMMARY

Different species of the genus *Cucumis* were characterized with molecular markers. Highly repetitive satellite DNA and different elements of the middle repetitive ribosomal RNA genes were used as hybridization probes. The RFLP analyses and hybridizations with different satellite DNAs and rDNA probes confirmed the division of the genus *Cucumis* into an African and an Asian subgenus and revealed that *C. melo* of the African subgenus is more closely related to *C. sativus* and *C. hardwickii* of the Asian subgenus than to *C. anguria* which also belongs to the African subgenus.

Key-words: Cucumis, Cucurbitaceae, highly and middle repetitive sequences, RFLP, ribosomal DNA, satellite DNA.

INTRODUCTION

Morphological traits, sexual crossability and cytogenetic studies are traditionally used in taxonomic and evolutionary studies of plants (Dane *et al.* 1980; Singh & Yavada 1984; Ramachandran & Narayan 1985; Den Nijs & Visser 1985; Van Raamsdonk *et al.* 1989). The new techniques of RFLP analysis and the comparison of protein and DNA sequences can complement our view of genetic variability, the basis of every comparative analysis. Besides the investigation of isozymes by gel electrophoresis (Perl-Treves *et al.* 1985; Staub *et al.* 1987) and the analysis of the plastome (Perl-Treves & Galun 1985), repetitive sequences could be valuable in studies of phylogenetic relationships between different species and cultivars within a plant genus or between genera of a family (Beckmann & Soller 1986; Crawford 1990; Hemleben *et al.* 1992).

The genus *Cucumis* includes about 30 species and was recently subdivided into two subgenera (Jeffrey 1980). The Asian group of the genus, the subgenus *Cucumis*, including as most important species *C. sativus* L., has a basic chromosome number $n=7$. The second subgenus *Melo* with $n=12$ as basic chromosome number comprises four groups (*metuliferus, anguria, melo, hirsutus*); it is mainly distributed in Africa. The species of the two subgenera are not cross-compatible (Deakin *et al.* 1971; Den Nijs & Visser 1985); although this would be helpful in some cases because several interesting resistances are found in African wild species (Robinson & Munger 1976). New techniques like protoplast fusion or asymmetric fusion might help to overcome this barrier and the identification of

*To whom correspondence should be addressed.

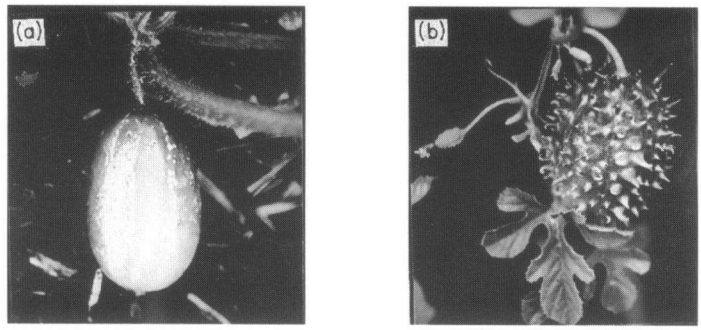

Fig. 1. Fruits of the two wild species. (a) *Cucumis hardwickii* Royle (b) *Cucumis anguria* ssp. *anguria*.

the fusion partners by isozymes, RFLPs or species-specific markers in a very early stage of regeneration of the hybrids might be valuable (Malepszy 1988; Negrutiu *et al.* 1989).

Highly and middle repetitive genome components have already been analysed for different species in the genus *Cucumis* (Ganal & Hemleben 1986, 1988; Ganal *et al.* 1986). Further characterization by RFLPs and species-specific satellite DNA which is investigated in this paper might also be helpful as an additional tool for analysing phylogenetic relationships within the genus *Cucumis*.

MATERIALS AND METHODS

Plant material

Seeds of *Cucumis sativus* cv. Sensation Typ Neckarruhm and *Cucumis melo* cv. Benarys Zuckerkugel were purchased commercially (Endriβ, Tübingen, Germany) whereas seeds of *Cucumis anguria* ssp. *anguria* and *Cucumis hardwickii* Royle were obtained from the Zentralinstitut für Genetik und Kulturpflanzenforschung, Gatersleben, Germany (Fig. 1). *Cucumis hardwickii* Royle grows wild in the foothills of the Himalayas and is suggested to be a progenitor of the cultivated cucumber (Whitaker & Bemis 1976).

DNA isolation

Seeds were cultivated under sterile conditions at 25°C. Seedlings were harvested after 7–10 days or leaves were taken from 4–6 weeks old plants. Nuclei were isolated and DNA was purified as described by Hemleben *et al.* (1982).

Analysis of DNA

The DNA samples were digested overnight with restriction enzymes according to manufacturers' specifications. The DNA restriction fragments were separated on 0·7 or 1·2% agarose gels and transferred to nylon filter (Hybond N^+, Amersham) for 4 h using 0·4 M NaOH as transfer buffer. The filters were pre-hybridized for 3 h at 65°C in pre-hybridization buffer containing 5 × SSPE (0·5 M NaCl, 5 mM EDTA, 0·5 M NaH_2PO_4, pH 7·0), 5 × Denhardts (0·1% Ficoll, 0·1% bovine albumin, 0·1% Polyvinylpyrollidon), 0·1% SDS, and 50 μg ml^{-1} denatured salmon sperm DNA. Hybridizations were carried out for 16 h at 65°C in pre-hybridization buffer in the presence of oligolabelled radioactive probes (Feinberg & Vogelstein 1983). After hybridization, the filters were washed for 10 min at room-temperature in 2 × SSPE, 0·1% SDS; then for 20 min at 65°C in 1 × SSPE, 0·1% SDS, and finally for 20 min at 65°C in 0·1 × SSPE, 0·1% SDS.

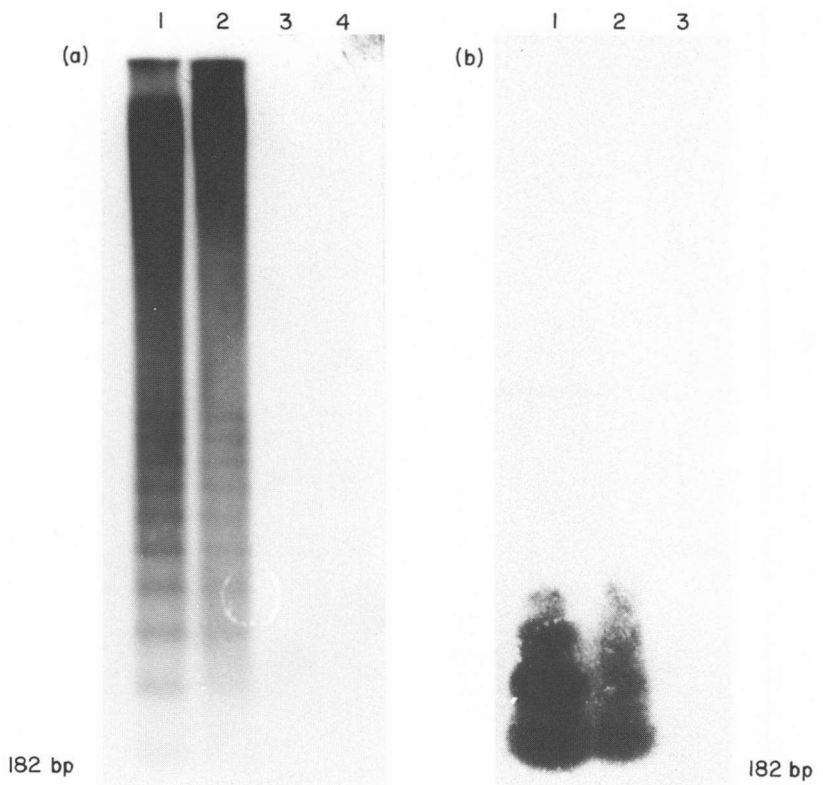

Fig. 2. Distribution of the type I satellite DNA of *C. sativus*. DNA (3 μg) of *C. sativus* (lane 1), *C. hardwickii* (lane 2), *C. anguria* (lane 3), *C. melo* (lane 4) were digested with SmaI (A) and TaqI (B) and hybridized with the 182-bp type I satellite DNA of *C. sativus*.

Hybridization probes

Characterized, cloned sequences were isolated by restriction digestion from recombinant plasmids:

(a) 352-bp satellite repeat of *Cucumis melo* (pSat107/Hemleben *et al.* 1982; Brennicke & Hemleben 1983),
(b) 182-bp satellite repeat of *Cucumis sativus* (pSG12/Ganal *et al.* 1986),
(c) 360-bp satellite repeat (type IV) of *Cucumis sativus* (Ganal & Hemleben 1988),
(d) 5S rDNA repeat of *Vigna radiata* (pVR51/Hemleben & Werts 1988),
(e) GC-rich cluster of the IGS of *Cucumis sativus* rDNA (Ganal *et al.* 1988),
(f) 3′ end of the 25S rDNA coding region of *Cucumis sativus* (Ganal *et al.* 1988),
(g) 330-bp fragment of the 5′ ETS repeats of *Cucumis sativus* rDNA (Zentgraf *et al.* 1990).

RESULTS

Hybridization with a type I satellite-DNA repeat of *C. sativus* with a repeat length of 182 bp resulted in a typical restriction pattern for satellite DNA for two different (a

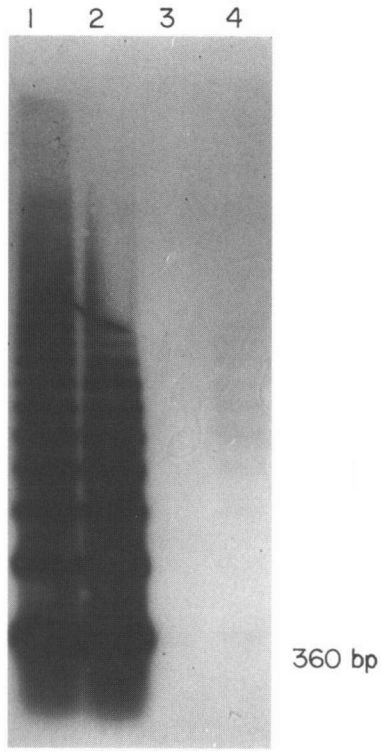

Fig. 3. Distribution of type IV satellite DNA of *C. sativus*. DNA (3 µg) of *C. sativus* (lane 1), *C. hardwickii* (lane 2), *C. anguria* (lane 3), *C. melo* (lane 4) were digested with Sau3A and hybridized with the 360-bp type IV satellite DNA of *C. sativus*.

methylation sensitive and an insensitive) enzymes in *C. sativus* as well as in *C. hardwickii*. A hybridization signal was not detectable for either of the African subgenus members (Fig. 2). Using a type IV satellite-DNA repeat of *C. sativus* with a repeat length of 360 bp, a strong signal was obtained for the DNA of *C. sativus* and *C. hardwickii* but, in contrast to type I, an additional weak signal could be recognized for *C. melo* (Fig. 3). The 352-bp satellite repeat isolated from *C. melo* hybridized exclusively with DNA from *C. melo* (Fig. 4).

Further information was gathered by hybridization with a 5S rDNA probe of *Vigna radiata* (Fabaceae). In contrast to the 5S rRNA coding regions, the spacers separating the genes are variable in length and sequence, so that hybridization with the 5S rDNA probe can often result in a RFLP. The members of the Asian subgenus have a repeat length of *c.* 300 bp, whereas *C. anguria* and *C. melo* of the African subgenus can be distinguished from the Asian subgenus having both a repeat length of *c.* 330 bp (Fig. 5).

Different spacer elements of the 18S, 5·8S, and 25S rDNA cluster were used as hybridization probes analysing sequence homology and structural organization. An EcoRV digestion exhibiting a conserved restriction site in the 5·8S rRNA gene often results in restriction fragments representing the repeating unit of the rRNA genes (Fig. 6a). The length heterogeneity of the rDNA repeat within and between species is often caused by a different number of repeated elements in the spacer. *C. sativus* shows two hybridization signals with the 5′ ETS repeats in the EcoRV digestion corresponding to

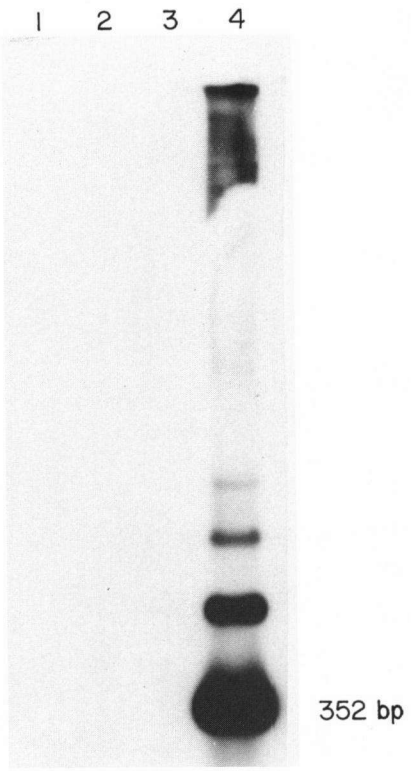

Fig. 4. Distribution of the 352-bp satellite DNA of *C. melo*. DNA (3 μg) of *C. sativus* (lane 1), *C. hardwickii* (lane 2), *C. anguria* (lane 3), *C. melo* (lane 4) were digested with HindIII and hybridized with the 352-bp satellite DNA of *C. melo*.

two prominent repeat classes, 11·5 kbp and 10·5 kbp in length. The length of the rDNA repeats of *C. melo* is described as 10·7 kbp and 10·55 kbp, respectively (Ganal & Hemleben 1986). In *C. anguria* one prominent rDNA repeat was detected with a length of *c.* 9·3 kbp.

In *C. sativus* an extraordinary feature has been observed in the IGS of the rRNA genes: the 3′ end of the 25S rRNA coding region and the subsequent sequences are duplicated in the spacer (see Fig. 6e; Ganal *et al.* 1988). Using the 3′ end of the 25S rRNA gene as a hybridization probe, these duplications are visible in a NsiI digestion in *C. sativus*, but not in *C. anguria* and *C. melo* (Fig. 6b). The digestion with RsaI indicates that *C. melo* also shows duplications of this region (Fig. 6c), but this has to be verified by further sequence analysis. Interestingly, the rDNA repeat of *C. anguria*, probably not exhibiting this phenomenon, is much shorter in comparison to *C. sativus* and *C. melo*. In the digestion with CfoI and the hybridization with a cluster of 30 bp GC-rich elements of the IGS of *C. sativus* (Ganal *et al.* 1988) *C. sativus*, *C. anguria* and *C. melo* are clearly distinguished from each other by RFLP (Fig. 6d). The results of these investigations are summarized in Tables 1 and 2.

DISCUSSION

Middle and highly repetitive sequences can be valuable as molecular markers in studies of phylogenetic relationships (Torres *et al.* 1989; Torres & Hemleben 1991). The combination of rDNA coding and spacer regions and highly repeated satellite-DNA markers

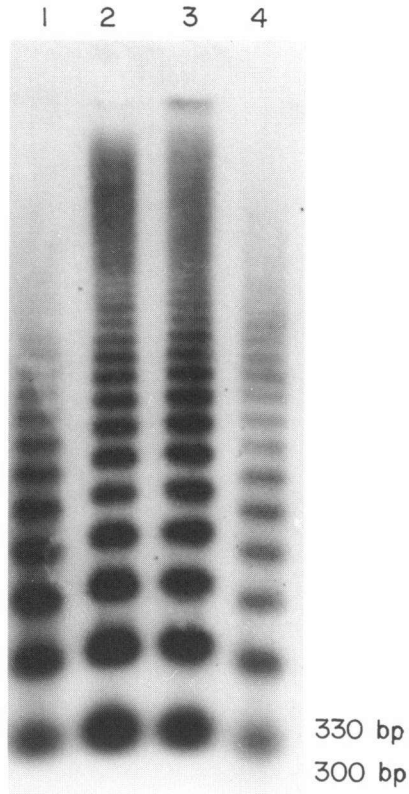

Fig. 5. Organization of the 5S rRNA genes. DNA (3 μg) of *C. sativus* (lane 1), *C. melo* (lane 2), *C. anguria* (lane 3), *C. hardwickii* (lane 4) were digested with BamHI and hybridized with the 5S rDNA repeat of *Vigna radiata*.

especially, can detect relationships between genera and species; additionally, synthetic oligonucleotides (Weising *et al.* 1989) and mini-satellites (Jeffreys *et al.* 1985) can be used to distinguish between cultivars, populations, or even individuals.

The RFLP analysis performed with the middle repetitive 5S rDNA probe renders it possible to distinguish between the two subgenera *Cucumis* and *Melo* (Jeffrey 1980), formerly defined and confirmed by morphological traits, sexual crossability and cytogenetic studies (Dane *et al.* 1980; Singh & Yavada 1984; Ramachandran & Narayan 1985; Den Nijs & Visser 1985; Van Raamsdonk *et al.* 1989). Using different probes of the 18S, 5·8S, and 25S rDNA spacer, *C. anguria* and *C. melo* of the subgenus *Melo* can clearly be distinguished from each other by their spacer organization resulting in RFLPs. This is in agreement with the division of the subgenus *Melo* into the 'anguria group' and the 'melo group' by the analysis of iso-enzymes and plastomes of 21 *Cucumis* species (Perl-Treves *et al.* 1985).

Highly repetitive DNA often occurs as tandem arrays of a short repeating unit which are strongly conserved among each other. Several mechanisms, for example unequal crossover, gene conversion, 'slippage replication' are held responsible for the specific manner of evolution of these sequences, leading to 'concerted evolution' (Smith 1976; Dover 1982). This molecular drive often results either in related satellite repeats or in one

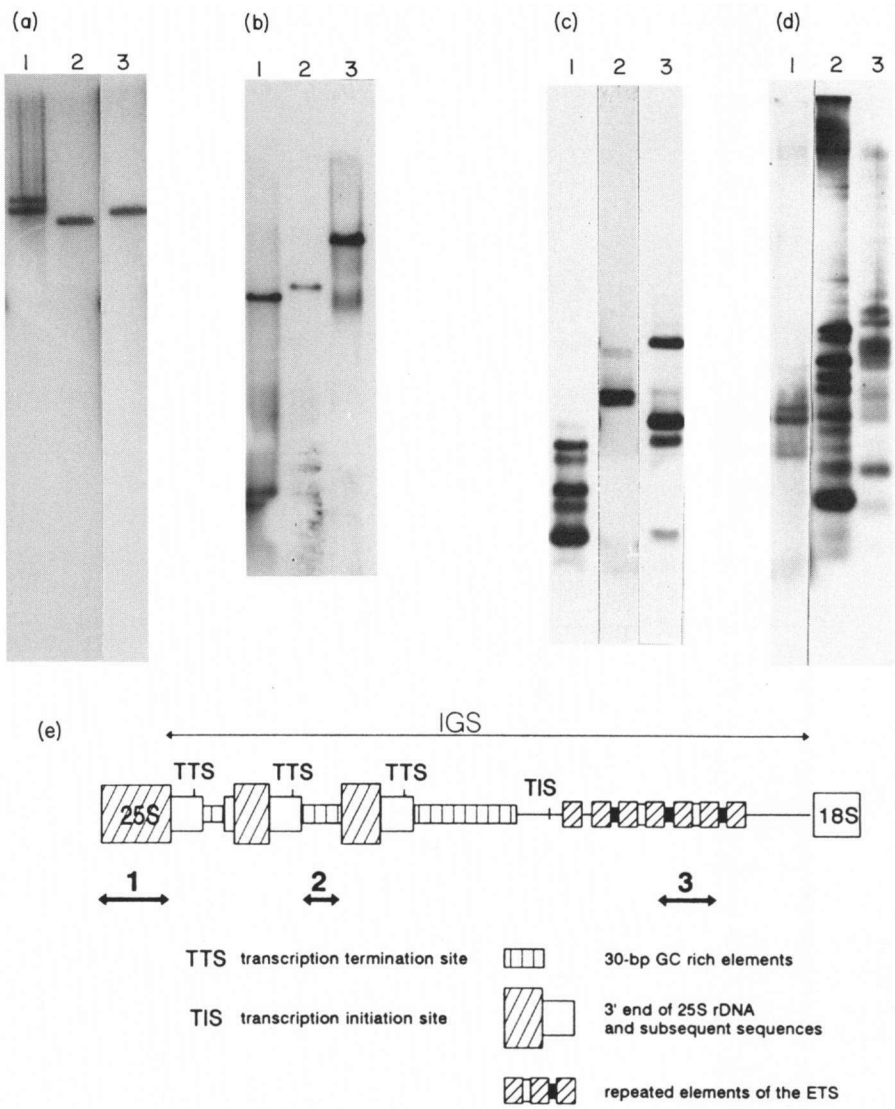

Fig. 6. RFLP analysis with various rDNA probes. DNA (3 µg) of *C. sativus* (lane 1), *C. anguria* (lane 2), *C. melo* (lane 3) were digested with EcoRV (a), NsiI (b), RsaI (c) and CfoI (d) and hybridized with probe 3 (a), probe 1 (b and c) and probe 2 (d) indicated in a schematic drawing of the intergenic spacer of the rDNA repeat of *C. sativus* by arrows (e).

repeat population being replaced by another (Dover 1986). New variants of satellite repeats can also originate from transposable elements indicated by terminal inverted repeats (Flavell 1986). Thereby, species-, genus- or family-specific satellite DNA can appear as described for various plants (Hallden *et al.* 1987; Schweizer *et al.* 1988; Zhao *et al.* 1989; Crowhurst & Gardner 1991; Schmidt *et al.* 1991). These satellite sequences can be a useful tool for the identification of species, as demonstrated here in the hybridization with the 352-bp satellite DNA of *C. melo* and 182-bp type I satellite DNA of *C. sativus*. The 182-bp type I satellite is not present in *C. melo* and, the 352-bp satellite of *C. melo*

Table 1. Comparison of middle and highly repetitive sequences in the genus *Cucumis*

	C. sativus	*C. hardwickii*	*C. melo*	*C. anguria*
Type I satellite	182 bp	182 bp	—	—
Type IV satellite	360 bp	360 bp	(360 bp)	—
Satellite of *C. melo*	—	—	352 bp	—
5S rDNA	300 bp	300 bp	330 bp	330 bp
rDNA repeat	11·5 kbp 10·5 kbp	ND	10·7 kbp 10·5 kbp	9·3 kbp
ETS repeats	+	ND	+	+
GC-rich cluster	+	ND	+	+
Duplication 3′ end 25S rDNA	+	ND	(+)	(−)

ND: not determined.

Table 2. Combination of hybridization probes and restriction enzymes producing RFLPs to distinguish between the *Cucumis* species

Hybridization probe	EcoRV	NsiI	RsaI	CfoI	BamHI
5S rDNA					+
ETS repeats	+				
GC-rich cluster				+	
3′ end of 25S rDNA		+	+		

cannot be found in *C. sativus* and *C. hardwickii*. *C. anguria* has neither of these repetitive elements. However, the satellite hybridization with the 360-bp type IV satellite DNA of *C. sativus* shows a weak signal with DNA of *C. melo* but not with DNA of *C. anguria*. This means that either a small number of type IV satellite is still present in the *C. melo* genome or a satellite with only weak sequence similarity has been amplified. *C. anguria* seems to have totally different repetitive elements, which are not yet identified. Additionally, the hybridization with the 3′ end of the 25S rDNA implicates that a duplication of this region is probably present in *C. melo* but not in *C. anguria*. Various bands appear in the RsaI digestion after hybridization with a very small probe of the 3′ end of the 25S rDNA in *C. melo*, whereas for *C. anguria* only one restriction fragment is visible. This is also in agreement with the shorter rDNA repeat of *C. anguria* (9·3 kbp) in comparison to *C. melo* and *C. sativus* (10·5–11·5 kbp). Therefore, the organization of the rDNA spacer seems also to be more similar in *C. melo* and *C. sativus*. From these data, the closer relationship *C. melo* to *C. sativus* and *C. hardwickii* of the Asian subgenus than to *C. anguria* of the African subgenus can be predicted. Remarkably, the crosses *C. anguria* × *C. sativus*, *C. anguria* × *C. melo* and *C. sativus* × *C. melo* always resulted in fruits with inviable seeds; in the case of *C. sativus* × *C. melo* pollen germination, pollen tube growth and penetration are often normal and even fertilization and some embryo development up to the globular stage may take place (Kroon *et al.* 1979; Kho *et al.* 1980; Den Nijs & Visser 1985).

ACKNOWLEDGEMENT

The support of this work by a grant of the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

REFERENCES

Beckmann, J.S. & Soller, M. (1986): Restriction fragment length polymorphisms and genetic improvement of agriculture species. *Euphytica* **35:** 11–124.

Brennicke, A. & Hemleben, V. (1983): Sequence analysis of the cloned *Cucumis melo* (highly repetitive satellite DNA. *Z. Naturforschg.* **38c:** 1062–1065.

Crawford, D.J. (1990): *Plant Molecular Systematics. Macromolecular Approaches.* J. Wiley & Sons, New York.

Crowhurst, R.N. & Gardner (1991): A genome-specific repeat sequence from kiwifruit (*Actinida deliciosa* var. *deliciosa*). *Theor. Appl. Genet.* **81:** 71–78.

Dane, F., Denna, D.W. & Tschuchiya, T. (1980): Evolutionary studies of the wild species in the genus *Cucumis*. *Z. Pflanzenzüchtung* **85:** 89–109.

Deakin, J.R., Bohn, G.W. & Whitaker, T.W. (1971): Interspecific hybrids in *Cucumis*. *Econ. Bot.* **25:** 195–211.

Den Nijs, A.P.M. & Visser, D.L. (1985): Relationships between African species of the genus *Cucumis* L. estimated by the production, vigour and fertility of F_1 hybrids. *Euphytica* **34:** 279–290.

Dover, G. (1982): Molecular drive: A cohesive mode of species evolution. *Nature* **299:** 11–117.

Dover, G. (1986): Molecular drive in multigene families: How biological novelties arise, spread and are assimilated. *Trends Genet.* **2:** 159–165.

Feinberg, A. & Vogelstein, B. (1983): A technique for radio-labelling DNA restriction fragments to high specific activity. *Anal. Biochem.* **132:** 6–13.

Flavell, R.B. (1986): Repetitive DNA and chromosome evolution in plants. *Phil. Trans. R. Soc. Lond. B.* **312:** 227-242.

Ganal, M. & Hemleben, V. (1986): Comparison of the ribosomal RNA genes in four closely related *Cucurbitaceae*. *Pl. Syst. Evol.* **154:** 63–77.

Ganal, M. & Hemleben, V. (1988): Integration and amplification of a DNA sequence in satellite DNA of *Cucumis sativus* L. (cucumber). *Theor. Appl. Genet.* **75:** 357–361.

Ganal, M., Riede, I. & Hemleben, V. (1986): Organization and sequence analysis of two related satellite DNAs in cucumber (*Cucumis sativus* L.). *J. Mol. Evol.* **23:** 23–30.

Ganal, M., Torres, R. & Hemleben, V. (1988): Complex structure of the ribosomal spacer of *Cucumis sativus* (cucumber). *Mol. Gen. Genet.* **212:** 548–554.

Hallden, C., Bryngelsson, T., Säll, T. & Gustafson, M. (1987): Distribution and evolution of a tandem repeated DNA sequence in the family Brassicaceae. *J. Mol. Evol.* **25:** 318–323.

Hemleben, V. & Werts, D. (1988): Sequence organization and putative regulatory elements in the 5S rRNA genes of two higher plants (*Vigna radiata* and *Matthiola incana*). *Gene* **62:** 65–69.

Hemleben, V., Leweke, B., Roth, A. & Stadler, I. (1982): Organization of highly repetitive satellite DNA of two cucurbitaceae species (*Cucumis melo* and *Cucumis sativus*). *Nucleic Acids Res.* **10:** 631–644.

Hemleben, V., Zentgraf, U., King, K., Borisjuk, N., Schweizer, G. (1992): Middle repetitive and highly repetitive sequences detect polymorphisms in plants. In: *DNA-Polymorphisms in Eucaryotic Genomes.* Hüthig, Heidelberg, pp. 157–170.

Jeffrey, C. (1980): A review of the Cucurbitaceae. *Bot. J. Linn. Soc.* **81:** 233–247.

Jeffreys, A.J., Wilson, V. & Thein, S.L. (1985): Individual-specific 'fingerprints' of human DNA. *Nature* **316:** 76–79.

Kho, Y.O., Den Nijs, A.P.M. & Franken, J. (1980): Interspecific hybridization in *Cucumis* L. II. The crossability of species, an investigation of the in vivo pollen tuber growth and seed set. *Euphytica* **29:** 661–672.

Kroon, G.H., Custers, J.B.M., Kho, Y.O., Den Nijs, A.P.M. & Varekamp, H.Q. (1979): Interspecific hybridization in *Cucumis* L. I. Need for genetic variation, biosystematic relations and possibilities to overcome crossability barriers. *Euphytica* **28:** 723–728.

Malepszy, S. (1988): Cucumber (*Cucumis sativus* L.). In: Bajaj, Y.P.S. (ed.): *Biotechnology in Agriculture and Forestry, Vol. 6: Crops II.* Springer, Berlin, Heidelberg.

Negrutiu, I., Mouras, A., Gleba, Y.Y., Sidorov, V., Hinnisdaels, S., Famelaer, Y. & Jacobs, M. (1989): Symmetric versus asymmetric fusion combinations in higher plants. In: Bajaj, Y.P.S. (ed.): *Biotechnology in Agriculture and Forestry. Vol. 3: Potato.* Springer, Berlin, Heidelberg.

Perl-Treves, R. & Galun, E. (1985): The *Cucumis* plastome: physical map, interageneric variation and phylogenetic relationship. *Theor. Appl. Genet.* **71:** 417–429.

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.

Ramachandran, C. & Narayan, R.K.J. (1985): Chromosomal DNA variation in *Cucumis*. *Theor. Appl. Genet.* **69:** 497–502.

Robinson, R.W. & Munger, H.M. (1976): Genes of the Cucurbitaceae. *HortScience* **11:** 554–564.

Schmidt, T., Jung, C. & Metzlaff, M. (1991): Distribution and evolution of two satellite DNAs in the genus *Beta*. *Theor. Appl. Genet.* **82:** 793–799.

Schweizer, G., Ninnemann, H. & Hemleben, V. (1988): Species specific DNA sequences for identification of somatic hybrids between *Lycopersicon esculentum* and *Solanum acaule*. *Theor. Appl. Genet.* **75:** 679–684.

Singh, A.K. & Yavada, K.S. (1984): An analysis of interspecific hybrids and phylogenetic implications in *Cucumis* (Cucurbitaceae). *Pl. Syst. Evol.* **147:** 237–252.

Smith, G.P. (1976): Evolution of repeated DNA sequences by unequal crossover. *Science* **191:** 528–535.

Staub, J.E., Fredrick, L. & Marty, T.L. (1987): Electrophoretic variation in cross-compatible wild diploid species of *Cucumis*. *Canad. J. Bot.* **65:** 792–798.

Torres, R.A. & Hemleben, V. (1991): Use of ribosomal DNA spacer probes to distinguish cultivars of *Cucurbita pepo* L. and other Cucurbitaceae. *Euphytica* **53:** 11–17.

Torres, R.A., Zentgraf, U. & Hemleben, V. (1989): Species and genus specificity of the intergenic spacer (IGS) in the ribosomal RNA genes of Cucurbitaceae. *J. Biosciences (Z. Naturforschg.)* **44:** 1029–1034.

van Raamsdonk, L.W.D., Den Nijs, A.P.M. & Jongerius, M.C. (1989): Meiotic analyses of *Cucumis* hybrids and an evolutionary evaluation of the genus *Cucumis* (Cucurbitaceae). *Pl. Syst. Evol.* **163:** 133–146.

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

Whitaker, T.W. & Bemis, W.P. (1976): Cucurbits. In: Simmonds, N.W. (ed.): *Evolution of Crop Plants*. Longman, London.

Zentgraf, U., Ganal, M. & Hemleben, V. (1990): Length heterogeneity of the rRNA precursor in cucumber (*Cucumis sativus*). *Plant Mol. Biol.* **15:** 465–474.

Zhao, X., Wu, T., Xie, Y. & Wu, R. (1989): Genome-specific repetitive sequences in the genus *Oryza*. *Theor. Appl. Genet.* **78:** 201–209.