Theor Appl Genet (1995) 90: 1101–1105

B. Grebenstein · O. Grebenstein · W. Sauer · V. Hemleben

Characterization of a highly repeated DNA component of perennial oats (*Helictotrichon*, Poaceae) with sequence similarity to a A-genome-specific satellite DNA of rice (*Oryza*)

Received: 1 May 1994 / Accepted: 8 December 1994

Abstract The taxonomic relationships among perennial oats (*Helictotrichon* Besser ex Schultes & Schultes, Avenae, Aveneae, Poaceae) have been studied using highly repeated satellite DNA as a molecular marker. Highly repetitive sequences were isolated from restriction endonuclease digests of nuclear DNA of *Helictotrichon convolutum*, and satellite repeats (approximately 365 bp in length) were cloned, sequenced and compared among each other. They exhibited an intraspecific sequence variability of 6–9%. This satellite DNA, CON1, is differentially distributed within the genus *Helictotrichon*. In species of the subgenus *Helictotrichon* a high copy number is detectable, whereas in representatives of the subgenera *Pratavenastrum* and *Pubavenastrum* the number of copies per genome is rather low. Surprisingly, the satellite DNA repeat CON1 shows 74% sequence similarity to an A-genome specific repetitive DNA of *Oryza* (rice).

Key words *Helictotrichon* · *Oryza sativa* · Satellite DNA · Genome-specificity · Phylogenetic relationship

Introduction

Highly repetitive DNA represents a major component of the genome of higher plants. Detailed information on its structure, genomic organization and evolution has already been provided. Highly repetitive DNA is either organized in long tandem arrays of relatively short repeated elements (Flavell 1986) or is interspersed between unique or other repetitive sequences (Bedbrook et al. 1980 a, b). Genus- and/or species-specific DNA repeated elements arranged in tandem arrays were found; therefore, detection and characterization of these genome components may help to elucidate the phylogenetic relationships of species within a genus or at the family level (Hemleben et al. 1992; Hemleben 1993). The existence of repetitive DNA sequences in the grass family (Poaceae) which appear to be either species-specific or else characteristic for certain species has been well described, especially within the Triticeae (Appels et al. 1978; Rimpau et al. 1978; Bedbrook et al. 1980 a, b; Appels and McIntyre 1985; McIntyre et al. 1988). On the other hand, Appels et al. (1987) and Xin and Appels (1988) demonstrated that a 350-bp rye-specific repeated sequence is prominently represented in *Agropyron* species. Another highly repeated 120-bp sequence of rye (*Secale cereale*) is also present in *Hordeum* species (Gupta et al. 1989). The genera mentioned, *Secale*, *Agropyron* and *Hordeum*, all belong to the Triticeae.

For molecular studies, our interest is focussed on the taxonomic relationships of species within the genus *Helictotrichon*. This genus has been well studied by classical morphological and karyological methods (Sauer and Heubl 1984; Röser 1989). Moreover, the species which are adapted to different habitats all over the world may represent an important genetic pool for the agriculturally important *Avena sativa* (common oat, Avenineae, Aveneae). In terms of morphological characters, the genus *Helictotrichon* is subdivided into four subgenera (*Helictotrichon*, *Pratavenastrum*, *Pubavenastrum*, and *Tricholemma*; Röser 1989, 1992). However, species with a mosaic of morphological characteristics are difficult to interpret and to place into defined subgenera. Therefore, evolutionary relationships among these species were investigated by molecular methods to obtain more information on their taxonomical position.

Interestingly, the satellite repeat type shown here to be characteristic for the subgenus *Helictotrichon* exhibits high sequence similarity to the repeated element which has been shown to be specific for the A-genome of rice (Zhao et al. 1989; Wu et al. 1991).

Communicated by J. W. Snape

B. Grebenstein · O. Grebenstein · W. Sauer
Lehrstuhl für Spezielle Botanik, Botanisches Institut der Universität Tübingen, Auf der Morgenstelle 1, 72076 Tübingen, Germany

V. Hemleben (✉)
Lehrstuhl für Allgemeine Genetik, Biologisches Institut, Universität Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

Table 1 List of investigated plants, their accessions and collectors

Species	Accession	Collector
Oryza sativa	Botanical Garden Tübingen	
Arrhenatherum elatius	North Caucasus, Khevsureti	Leg. Sauer and Grebenstein
Avena sativa	var. "Alfred", NR. D/S 312514	
Genus *Helictotrichon*		
Subgenus *Helictotrichon*		
Helictotrichon sempervirens	France, Dauphine Alps	Leg. Melzer
Helictotrichon convolutum	former Jugoslavia, Dalmatia	Leg. Rheder
Helictotrichon filifolium	Spain, Granada	Leg. Röser
Helictotrichon cantabricum	Spain, Huesca	Leg. Röser
Subgenus *Pubavenastrum*		
Helictotrichon pubescens	Transcaucasus, Ossetia	Leg. Sauer and Grebenstein
Subgenus *Pratavenastrum*		
Helictotrichon gervaisii	Spain, Malaga	Leg. Röser
Helictotrichon cinncinatum	Italy, Sicilia	Leg. Röser
Helictotrichon occidentale	Portugal, Baixo Alentejo	Leg. Röser
Helictotrichon hackelii	Portugal, Baixo Alentejo	Leg. Röser
Helictotrichon compressum	Turkey, Vilajet Bolu	Leg. Sauer and Grebenstein

Materials and methods

Living plant material of wild perennial oats (*Helictotrichon sempervirens*, *H. convolutum*, *H. filifolium*, *H. cantabricum*, *H. pubescens*, *H. gervaisii*, *H. cinncinatum*, *H. occidentale*, *H. hackelii*, *H. compressum*) and *Arrhenatherum elatius* (Table 1) were collected in their natural habitats and were grown in pots in the greenhouse (23°C in summer, 17°C in winter). Seeds of *Oryza sativa* and *Avena sativa* were germinated in Petri dishes, small seedlings were transferred to pots and cultivated in the greenhouse under the same conditions.

Finely chopped leaf tissue from plants was frozen in liquid nitrogen. Nuclei were isolated, and DNA was purified by CsCl buoyant density gradients as described by Hemleben et al. (1982). Total nuclear DNA of *Helictotrichon convolutum* was completely digested with *Dra*I, fractionated electrophoretically on a 1.5% agarose gel, and stained with ethidium bromide. A prominent band of DNA fragments (see Fig. 1a), approximately 365 bp in length, was eluted from the gel, ligated into the *Sma*I cloning site of pUC19 (Maniatis et al. 1982), and transformed into *E. coli* PLK-F′ (Stratagene). Recombinant clones were screened by colony filter hybridization with digoxigenin-labelled genomic DNA of *H. convolutum*. Four recombinant clones (pCON1_1-pCON1_4) were sequenced with the Sequenase Version 2.0 Sequencing-Kit from USB. Data were analyzed using the GCG-programm (Devereux et al. 1984) and "Align/sequence alignment program" (Myers and Miller 1988).

Digested DNA samples were fractionated by electrophoresis on a 1.5% agarose gel. DNA fragments were transferred to a Hybond N membrane (Amersham). The insert of clone pCON1_1, used as hybridization probe, was random-primed labelled with digoxigenin-dUTP. Hybridization and detection with AMPPD (Tropix) was carried out according to the manufacturer's manual (Boehringer, Mannheim).

Results and discussion

Southern hybridization of *Dra*I-digested nuclear DNA of *H. convolutum* with the insert of clone pCON1_1 (cloned from *H. convolutum*) revealed that hybridization occurred in a series of bands corresponding to multimeric lengths of a 365-bp repeat (Fig. 1). Such a pattern is typical for a tandemly organized repeated element; it arises by mutational loss of the restriction sites defining the repeat unit.

Southern hybridization with pCON1_1 on DNA of various other grass species (Table 1) demonstrated that this satellite DNA family is distributed throughout the genus *Helictotrichon* (Fig. 1b). However, the intensity of hybridization signals indicate different copy numbers (or a modification of the repeats) in species of the subgenera investigated. Strong hybridization signals were obtained with DNA of species of the subgenus *Helictotrichon* (*H. sempervirens*, *H. convolutum*, *H. filifolium*, and *H. cantabricum*). A rather weak reaction with a slightly visible ladder pattern occurred with DNA of species of two other subgenera, *Pubavenastrum* (*H. pubescens*) and *Pratavenastrum* (*H. gervaisii*, *H. cinncinatum*, *H. occidentale*, *H. hackelii*, *H. compressum*). Satellite probe pCON1 is not detectable in species of two other genera assumed to be closely related to the genus *Helictotrichon*: *Arrhenatherum* and *Avena* (Fig. 1b). Remarkably, pCON1_1 also hybridized to repeats present in *O. sativa* with a characteristic ladder pattern (Fig. 1b).

The basic repeat length of CON1 (pCON1_1-pCON1_4) is 360–369 bp, with a G+C content of 43% and an intraspecific variability of 6–9% (Fig. 2). No internal duplications were detectable. On comparing these sequences with entries in the EMBL data bank a strong sequence similarity of CON1 was found to a repetitive DNA component of rice, first described as RC48 (isolated from *O. sativa*, cultivar "Labelle"; Wu and Wu 1987) and which almost corresponds to the sequence of OS7 (*O. sativa*, cultivar "Cigalon"; DeKocho et al. 1991; Ohtsubo et al. 1991). Satellite repeats of CON1 exhibit 74% sequence similarity to these A-genome specific 354-bp repeats of *Oryza*. Sequence alignment detected a 15-bp deletion in the rice repeat at position 172–187 compared to the sequences of CON1 (Fig. 2).

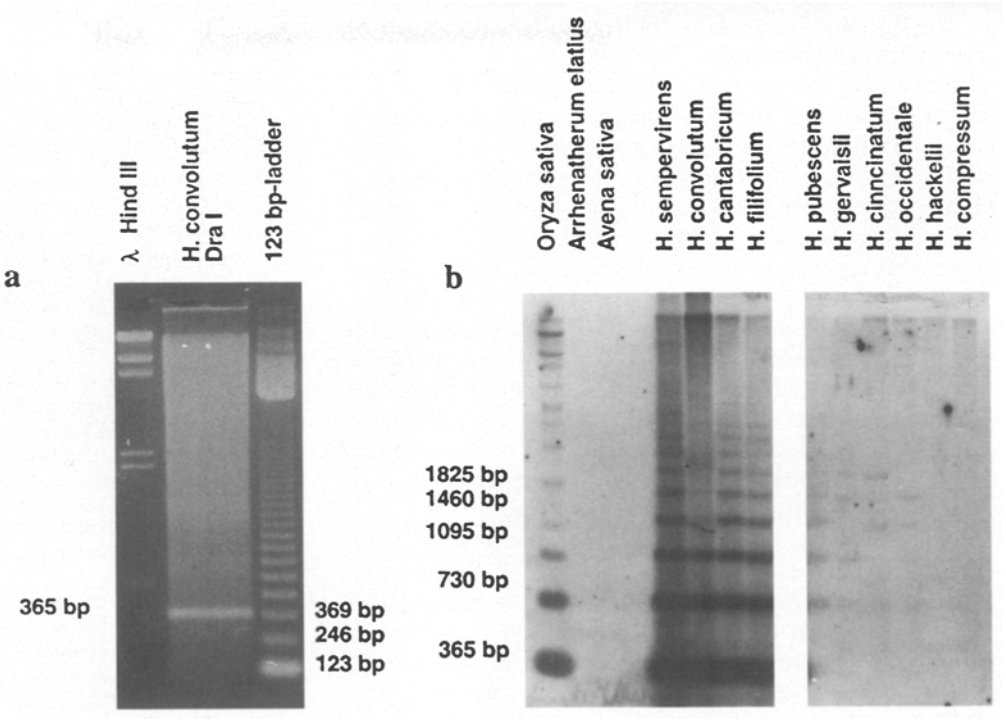

Fig. 1a, b Characterization of satellite DNA CON1 cloned from *Helictotrichon convolutum* and its distribution in several grass species. **a** Nuclear DNA of *H. convolutum* was digested with *Dra*I and electrophoretically separated in a 1.5% agarose gel. The gel was stained with ethidium bromide and photographed. The DNA fragments of a prominent band, approximately 365 bp in length, were eluted from the gel, cloned and sequenced (pCON1). **b** DNA was digested with *Eco*RI (*Oryza sativa*) or *Dra*I (*Avena sativa*, *Arrhenatherum elatius,* and *Helictotrichon* species), electrophoretically separated in a 1.5% agarose gel, and hybridized with digoxigenin-labelled pCON1_1. Exposure time was 4 h

```
                    10        20        30        40        50        60        70        80        90
                     *         *         *         *         *         *         *         *         *
pCON1_1     AAGTTTAAAACCTTGAGAACTTTGCCTTCTGGACATGGAATCGAGCTAGGTTTTTTTTCATAGCTGCAGAATCACATGCCTCACCCTA-GAGAACCTGAA
pCON1_2     ................C.....................................A.....A.....................-..A........
pCON1_3     .T...........A...........A...C.A.........A.....-...........G.....................---..........
pCON1_4     ..................................................----------.....................T...-........A..
RC48        ..-A............T...........C..T.........T..............A.CG.............G........GA.-....GA...G.
RG1         ..-A............T...........C..T.........T..............A.CG.............G........GA.-....GA...G.
RG2         ..-A............T...........CA.T.........T..............A.CG.............A....G...GA.-....GA...G.
RG3         ..--........-...T...........C..T.........................A.CG.............G........GG.A....GA...G.

                  100       110               120       130       140       150       160       170       180       190
                   *         *                 *         *         *         *         *         *         *         *
pCON1_1     TTGGCAACATTGACCC------GATATTTCTAAAAATCC-CTTGTCATAGGCCACAAAGCAGTGTTTGTTTCAGTGAAATCATGCAATGCAATTCCGACG
pCON1_2     .......A.G......TCTAGA.........A.........-...A.........G.........CA............A.....................
pCON1_3     .......A.G...T..TCTAGA.........A.........-...A..............T.....................A...........T..T.
pCON1_4     .......A.G......TCTGGA......A..........-.....................CA.....................................
RC48        ..T.G.GAG...GA..CCCAGA...T.......G....TA-...AC..........G......C.G.......G.C....----------------..T..
RG1         ..T.G.GAG...GA..CCCAGA...T.......G....TA-...AC..........G......C.G.......G.C....----------------..T..
RG2         ..T.G.GAG...GA..CCCAGA...T.......G....TTA...AC..........G......C.G.......GAC....----------------CG.
RG3         ..T.G.GAG...GA..CCCAGA...T.......G....TA-...ACT.........G......C.G.......G.C....----------------..T..

                       200       210       220       230       240       250       260       270       280       290
                        *         *         *         *         *         *         *         *         *         *
pCON1_1     AAGATGTCTG--GCTCCAGCCTTGATCGAAAATGAGTTGGTTAGGATAGGAATGGCACATAGAAAGTTAGGGTGGATAGTTTAGATGATGCTAGTATGAA
pCON1_2     .........A--.........-.....C.....................................-...........T.....................C..
pCON1_3     ..........--.............C.....................................................................C..
pCON1_4     ..........--.............C.....................................................................C..
RC48        ..T..C.GA.--.G....CAAC..G.TTTC..C...G.......T.....G..........G.C.............C.....C....TCG..G.C..
RG1         ..T..C.GA.--.G....CAAC..G.TTTC..C...G.......T.....G..........G.C.............C.....C....TCG..G.C..
RG2         ..TC.A....AG.G....CAAC..G.TTTC..C...G.......T.....G..........G.C.............C.....C....-.C..G.C..
RG3         ..T..C.AA.--.G....CAAC..G.-TTC..C...G.......T.....G..........G.CA............C.....C....-.C..G.C..

                             300       310       320       330       340       350       360
                              *         *         *         *         *         *         *
pCON1_1     TTACACATT-GCCTAACCACATTGTTATGGTATGTTGGTGGCCGGGACACCAGGGACTGCCGATGCAGATTTCAGGGC
pCON1_2     ..--.....-.................A....C.....C.....................................
pCON1_3     ..--.....-...C.............A....C.....C..-.....................................
pCON1_4     ..--..G..-T................A....C......T.T.A...T..........A........T....T.....
RC48        --G.G.G.C-.A......G.GCC....A....G.C...C..G..C.....TGT...A...A....G...A...T.T..
RG1         --G.G.G.C-.A......G.GCC....A....G.C...C..G..C.....TGT...A...A....G...A...T.T..
RG2         --G.G.G.C-.A......G.G.C....A....G.C...C..G..C.....TGT.A.A...A....G...A...T.T..
RG3         --G.G.G..C.A......G.GCC....A...TG.C...C..G..C.....TGT...A...A....G...A...T.T..
```

Fig. 2 Nucleotide sequence comparison of satellite repeats CON1 (cloned from *H. convolutum*) with A-genome-specific satellite repeats of *O. sativa* (RC48 and RG1-3; Wu and Wu 1987). (.) same nucleotide as in pCON1_1; (–) nucleotide is not present at this site

The distribution of the satellite family represented by the CON1 clones appears to coincide with the recent systematics of the genus *Helictotrichon* (Röser 1989). The amount of this satellite DNA component detectable in the species of *Helictotrichon* investigated so far varies considerably. Species of the subgenus *Helictotrichon* contain a high percentage of this satellite DNA whereas in species of the other subgenera (*Pubavenastrum* and *Pratavenastrum*) this repeat type is under-represented. Species of the subgenus *Tricholemma* have not yet been investigated; however, some evidence was recently obtained that this small group of species do not belong to the genus *Helictotrichon* (M. Röser, personal communication). Since this satellite DNA was not found in species of closely related genera, like *Avena* and *Arrhenatherum*, it could be assumed that this DNA component either has not been amplified or else was lost during species differentiation.

Studies on highly repeated genome fractions in *Solanum* (Schweizer et al. 1993), *Lycopersicon* (Ganal et al. 1988), *Cucumis* (Zentgraf et al. 1992), or the family Brassicaceae (Hallden et al. 1987), confirmed that satellite DNAs are valuable molecular markers for phylogenetical studies, since the distribution of different satellite DNAs can be species-, genera-, or even family-specific. These studies showed that the number of copies per genome of a certain satellite DNA can vary over a wide range. Either the satellite DNA repeats were amplified differently or else were lost during species differentiation as discussed by Dover (1982): molecular drive, through unequal crossover, gene conversion and slippage replication, often results in related satellite repeats or in one repeat population being replaced by another.

In rice, 354-bp satellite repeats were characterized as highly specific for the A-genome. Wu et al. (1991) suggested that this specific sequence was introduced into the A-genome of *Oryza* after its divergence from the other genome types of rice. However, the distribution of CON1 sequences in species of the genus *Helictotrichon*, and concomitantly in the distantly related genus *Oryza*, indicates a different origin for this satellite type. Probably, this DNA component was present in an ancestoral species or in a group of species of the Poaceae which then evolved in a distinct way during genus and species differentiation. Thus, satellite DNAs could be valuable as molecular markers in studies of phylogenetic relationships in large taxonomic groups, such as tribes and families.

Further investigations will show if satellite DNA CON1 is also present in genomes of other genera taxonomically grouping to the tribes *Aveneae* (e.g. *Deschampsia*, *Trisetum*, *Koeleria*, *Aira*, *Holcus*), or even in other grasses of the subfamily Pooideae. Additionally, an analysis of the distribution and a sequence comparison of one satellite DNA in several species of different genera could provide further insight into the complex genomic relationships within the agriculturally important family of Poaceae.

Acknowledgements The support of this work by a grant of the Deutsche Forschungsgemeinschaft is gratefully acknowledged. We thank Dr. Martin Röser, University of Vienna, for providing plant material.

References

Appels R, Driscoll C, Peacock WJ (1978) Heterochromatin and highly repeated DNA sequences in rye (*Secale cereale*). Chromosoma 70:67–89

Appels R, Mcintyre CL (1985) Cereal genome organisation as revealed by molecular probes. Plant Mol Cell Biol 2:235–252

Appels R, Scoles G, Chapman CGD (1987) The nature of change in nuclear DNA in the evolution of the grasses. In: Sodestrom TR (ed) Int Symp Grass Syst Evol. Smithonian Institution, Washington D.C., pp 73–97

Bedbrook JR, O'dell M, Flavell RB (1980 a) Amplification of rearranged repeated DNA sequences in cereal plants. Nature 288:133–137

Bedbrook JR, Jones J, O'Dell N, Thomson RD, Flavell RB (1980 b) A molecular description of telomeric heterochromatin in *Secale* species. Cell 19:545–560

DeKocho A, Kiefer MC, Cordesse F, Reddy AS (1991) Distribution and organisation of a tandemly repeated 352-bp sequence in the Oryzeae family. Theor Appl Genet 82:57–64

Devereux H, Haeberli J, Smithies TA (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Dover G (1982) Molecular drive: a cohesive mode of species evolution. Nature 299: 11–17

Flavell R (1986) Repetitive DNA and chromosome evolution in plants. Philos Trans R Soc Lond 312:227–242

Ganal M, Lapitian NLV, Tanksley SD (1988) A molecular and cytogenetic survey of major repeated DNA sequences in tomato (*Lycopersicon esculentum*). Mol Gen Genet 213:262–268

Gupta PK, Fedak G, Molnar SJ, Wheatcroft R (1989) Distribution of a *Secale cereale* DNA repeat sequence among 25 *Hordeum* species. Genome 32:383–388

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

Hemleben V (1993) Repetitive and highly repetitive DNA components as molecular markers for evolutionary studies and in plant breeding. Curr Topics Mol Genet (Life Sci Adv) 1:173–185

Hemleben V, Leweke B, Roth A, Stadler J (1982) Organisation of highly repetitive satellite DNA from 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: Kahl K, Appelhans H, Kömpf J, Driesel AJ (eds) DNA-polymorphisms in eukaryotic genomes. BioTechForum (BFT). Advances in Molecular Genetics, vol 5. Huethig Verlag, Heidelberg, pp 157–170

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Habour Laboratory, Cold Spring Habour, New York

McIntyre CL, Clarke BC, Appels R (1988) Amplification and dispersion of repeated DNA sequences in the Triticeae. Pl Syst Evol 160:39–59

Myers EW, Miller W (1988) Optimal alignment in linear space. CABIOS 4:11–17

Ohtsubo H, Umeda M, Ohtsubo E (1991) Organisation of DNA sequences highly repeated in tandem in rice genome. Jpn J Genet 66:241–254

Rimpau JS, Flavell RB (1978) Sequence organisation analysis of wheat and rye genomes by interspecies DNA/DNA hybridization. J Mol Biol 123:327–359

Röser M (1989) Karyologische, systematische und chorologische Untersuchungen an der Gattung *Helictotrichon* Besser ex Schultes & Schultes (Poaceae) im westlichen Mittelmeergebiet. Dissertationes Botanicae, Band 145. J. Cramer Verlag, Berlin

Röser M (1992) *Helictotrichon cintranum*, species nova, a rare southwest European oat grass (Poaceae, Pooideae, Aveneae). Taxon 41:60–61

Sauer W, Heubl GR (1984) Beiträge zur Kenntnis ausdauernder Wildhafer: 2. Karyotyp-Analysen an west- und ost-europäischen sowie an alpinen Wildhaferarten der Gattung *Avenula* (Dumort.). Phyton 24:193–223

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. Theor Appl Genet 85: 801–808

Wu K-K, Chung M-C, Wu T, Ning C-N, Wu R (1991) Localisation of specific repetitive DNA sequences in individual rice chromosomes. Chromosoma 100:330–338

Wu T, Wu R (1987) A new rice repetitive DNA shows sequence homology to both 5*S* RNA and tRNA. Nucleic Acids Res 15: 5913–5923

Xin Z-Y, Appels R (1988) Occurrence of rye (*Secale cereale*) 350-family DNA sequences in *Agropyron* and other Triticeae. Pl Syst Evol 160:65–76

Zentgraf U, King K, Hemleben V (1992) Repetitive sequences are valuable molecular markers in studies of phylogenetic relationships within the genus *Cucumis*. Acta Bot Neer 41:397–406

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