Theor Appl Genet (1992) 83:602–607

THEORETICAL AND APPLIED GENETICS

Organization of repeated sequences in species of the genus *Avena*

R. Solano, G. Hueros, A. Fominaya, and E. Ferrer *

Departamento de Biología Celular y Genética, Universidad de Alcalá de Henares, E-28871 Alcalá de Henares, Madrid, Spain

Received June 13, 1991; Accepted September 3, 1991
Communicated by P.L. Pfahler

Summary. Four repetitive sequences from *Avena murphyi* have been isolated and their genome organization studied in different species of the genus *Avena*. A tandem sequence array was found for the *Avena* species that contain the C genome. Three other dispersed sequences present in the A and C genomes were arranged in a genome-specific manner. The fact that no major differences in the hybridization patterns were found between species with the same basic genome is consistent with the current taxonomy of *Avena* species.

Key words: Repetitive sequences – *Avena* – Oats

Introduction

Repetitive DNA constitutes a large fraction of cereal genomes (Flavell 1980). These reiterated sequences have been distributed into two classes according to their arrangement within genomic DNA. Some sequences are dispersed throughout the genome and interspersed between unrelated, often single copy DNA sequences, whereas others contain a repeat unit arranged in tandemly reiterated clusters (Singer 1982; Bouchard 1982).

Detailed studies on cereal genomes have shown that a substantial proportion of these genomes consist of species-specific repeated sequences (Flavell et al. 1977; Rimpau et al. 1980; Flavell et al. 1980). In the last few years a number of repeated sequences obtained from different species have been cloned and their genome organization studied (Bedbrook et al. 1980; Appels and Moran 1984; Rayburn and Gill 1986; Dvorak et al. 1988; McIntyre et al. 1988; Fabijanski et al. 1990; Hueros et al. 1990).

* To whom correspondence should be addressed

Attempts to assign a functional role to repeated DNA in either cellular and evolutionary processes have not been supported by experimental data (Miklos 1986). However, the organization and amount of the families of repeated sequences that are species specific and/or shared have been useful in studying genome evolution at the molecular level (Gupta et al. 1989; Zhao et al. 1989; Vershinin et al. 1990; Belostotsky and Ananiev 1990).

In this study, molecular differences among species belonging to the genus *Avena* were analyzed with respect to the genomic organization of four repeated sequences isolated from *A. murphyi*. This allotetraploid species (genomes AACC) is strongly implicated in the evolution of the hexaploid oats (genomes AACCDD) based on morphological, biochemical and cytological evidence (Craig et al. 1972; Rajhathy and Thomas 1974; Sánchez de la Hoz and Fominaya 1989). However, the diploid progenitors of the polyploid genomes are still controversial (Legget 1980; Fominaya et al. 1988). The karyotypes of all the diploid species fall into two groups: one group with an A-genome and another with a C-genome. Structural modifications of the chromosomes appear to have accompained speciation within each group, as has been indicated by Rajhathy and Thomas (1974).

We provide further examples of the usefulness of the repeated sequences to show divergence between related species as well as evidence of the differences in the evolutionary trends of the repeated sequences.

Material and methods

Plant material

Eight species of the genus *Avena* were analyzed in this study (Table 1). Nomenclature and genome designation were based on the work of Rajhathy and Thomas (1974). All seeds were obtained from the Canadian Avena Collection except for *A. ventri-*

Table 1. Nomenclature and genome designation of *Avena* species

Species	Genome	Accession
A. canariensis	AcAc	CAV 3873
A. damascena	AdAd	CAV 0258
A. longiglumis	AlAl	CAV 2210
A. strigosa	AsAs	CAV 2836
A. pilosa	CpCp	CAV 0063
A. ventricosa	CvCv	Cc 7064
A. murphyi	AACC	CAV 2832
A. sativa	AACCDD	cv 'Pandora'

cosa, which was provided by M. D. Legget (Welsh Plant Breeding Station, Great Britain) and *A. sativa* cv 'Pandora', which was obtained from the Spanish Instituto Nacional de Semillas y Plantas de Viveros.

DNA isolation and molecular cloning

Large-scale preparation of total DNA was carried out with young green leaf material (usually from 1-month-old plants) following the procedure described by Sharp et al. (1988).

The clones used in this study were isolated from a genomic library of *A. murphyi* DNA constructed in pUC19. The library was produced by MboI digestion of *A. murphyi* DNA under conditions that generated a majority of fragments with a length ranging from 1,000 to 2,000 bp. The fragments were cloned into the BamHI site of pUC19. *E. coli* HB101 was used as a bacterial host. DNAs of each recombinant plasmid were transferred to Zeta Probe membranes (BioRad), and those containing repetitive sequences were identified by the intensity of their signal after hybridization with biotinylated *A. murphyi* DNA. Ligation, bacterial transformation, preparation of plasmid DNA, DNA restriction and agarose gel electrophoresis were carried out using standard techniques (Maniatis et al. 1982).

Nick-translation, filter hybridization and colorimetric detection of biotinylated probes

These operations were performed as described in Hueros et al. (1990). The probes were labelled with either biotin 11-dUTP or biotin 7-dATP.

Results

Four clones containing recombinant plasmids which strongly hybridized to genomic *A. murphyi* DNA were selected from the 100 clones probed since only those containing repetitive DNA give a strong hybridization signal with this system (Saul and Potrykus 1984). The clones were designated pAm1, pAm2, pAm3 and pAm4. The recombinant plasmids were double digested with SmaI and XbaI to estimate the total length of each insert. Clones pAm1 though pAm4 contained 464, 771, 2488 and 2898 bp inserts, respectively. The four inserts were presumed to belong to different repeated families since none showed cross hybridization (data not show).

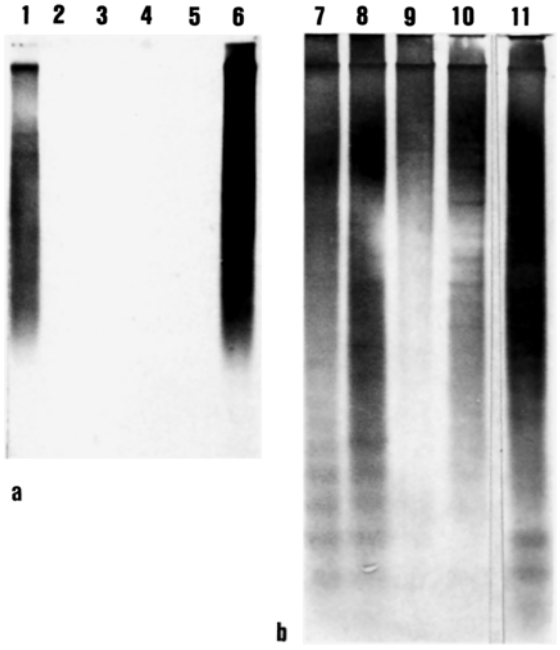

Fig. 1 a, b. Southern blot analysis of *Avena* DNA using biotinylated pAm1. **a** Five micrograms DNA from each of *A. pilosa* (*lane 1*), *A. strigosa* (*lane 2*), *A. canariensis* (*lane 3*), *A. longiglumis* (*lane 4*), *A. damascena* (*lane 5*) and *A. ventricosa* (*lane 6*) digested with HaeIII. **b** Five micrograms DNA from each of *A. murphyi* (*lanes 7 and 8*), *A. sativa* (*lanes 9 and 10*) and *A. pilosa* (*lane 11*) digested with HindIII (*lanes 7 and 9*) and HaeIII (*lanes 8, 10 and 11*). Digests were fractionated in 0.8% agarose gels

Genomic organization of the cloned sequences

Repetitive sequences occur in two prevalent patterns: either dispersed throughout the genome or clustered in tandem arrays. In order to determine the organization of the four repeated families in the *A. murphyi* DNA and to know how widespread these families were in other *Avena* species, total DNA from each of the eight species (Table 1) was digested with several different restriction enzymes and hybridized with biotinylated probes.

pAm1: a genome-specific repetitive sequence

When DNA from the various *Avena* species was digested with HaeIII and hybridized with the pAm1 probe, only the species that contained the C-genome in their karyotype showed a very strong hybridization signal, whereas none of the DNA samples from the A-genome diploid species exhibited detectable hybridization (Fig. 1). There is, therefore, a clear distinction at the DNA sequence level between the pAm1 repetitive sequence and any other repetitive family present in the A-genome and, consequently, the pAm1 family can be considered to be C-genome specific.

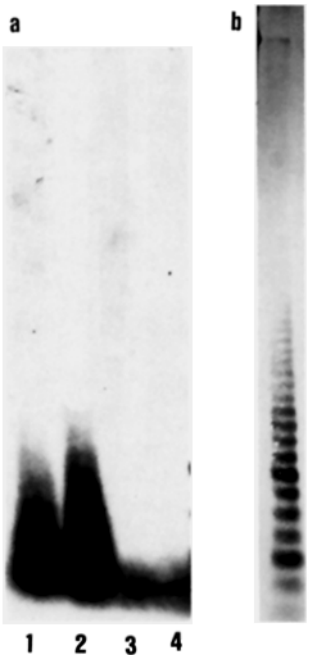

Fig. 2a, b. Southern blot analysis of DNA of *Avena* species which contain the C genome probed with biotinylated pAm1. **a** Five micrograms DNA from each of *A. murphyi* (*lanes 1 and 3*) and *A. pilosa* (*lanes 2 and 4*) were digested with BamHI (*lanes 1 and 2*) and Sau3A (*lanes 3 and 4*) and fractionated in a 0.8% agarose gel. **b** Five micrograms *A. murphyi* DNA were digested with BamH1 and fractionated in a 2% agarose gel

The organization of the pAm1 sequence family members within the DNA of the C-genome-containing species was revealed after digestion of their DNAs with BamH1 and Sau3A (Fig. 2). Instead of the characteristic smear of molecules hybridizing to the probe with occasional bands of low molecular weight, as was seen with HaeIII and HindIII digestions (Fig. 1), BamHI digestion generated a ladder of numerous bands fractionated on 2% gels (Fig. 2). This hybridization pattern did not change with larger amounts of BamHI restriction enzyme or with longer incubation times. The basic monomer unit had a length of 58 bp, and the major fragments corresponded in size to multimers of that basic unit. On the other hand, Sau3A digestion produced a unique band of 58 bp fractionated on 0.8% gel (Fig. 2). No differences have been found in these hybridization patterns among the C-genome-containing species so far investigated (*A. pilosa*, *A. ventricosa*, *A. murphyi* and *A. sativa*).

The hybridization pattern for pAm1 with the BamHI digest of the DNA from the above-mentioned species suggested that the members of the pAm1 family were organized in tandem arrays. The observed repetitive ladder was probably the result of either mutation or of methylation both of which might have altered the BamHI recognition sequence in random members of the repeated family. BamHI is reported to cleave the GGAT$\overset{*}{C}$C sequence when the marked cytosine is unmethylated, thus the ladder would be generated if unmethylated units flanked other methylated monomers. However, the Sau3A digestion pattern of the DNA from the C-genome-containing species does not support this hypothesis. Sau3A is also sensitive to methylation at the cytosine residue of its target site (GATC), however only a band of 58 bp appeared in the corresponding Southern blots (Fig. 2), indicating that no methylation occurred at the Sau3A sites of the monomers. Therefore, we can conclude that the ladder pattern from BamHI digestion is due to random point mutations in the flanking residues of its recognition sequence.

The hybridization pattern of pAm1 with DNAs from other *Avena* species digested with HaeIII and HindIII indicated that the recognition sequences for these enzymes were absent in most of the repeats. Small differences in the amount of the lower molecular weight fragments that hybridized to pAm1 were found depending on both the species DNA and the enzymes used (Fig. 1). These results would indicate that there were specific subsets of tandem arrays in the pAm1 family that were present in varying frequencies in the DNA of the different C-genome-containing species.

pAm2, pAm3 and pAm4: repetitive families from the A- and C-genomes

When DNA belonging to species with the A-genome (*A. strigosa*, *A. canariensis*, *A. longiglumis* and *A. damascena*), C-genome (*A. pilosa*) and with both the A- and C-genomes (*A. murphyi*) were digested with HaeIII and HindIII and hybridized with biotinylated pAm2, pAm3 or pAm4 separately, hybridization signals were detected in the Southern blots from all species (Figs. 3–5), indicating that repeated families with homology to pAm2, pAm3 and pAm4 are widespread in the A- and C-genomes of the *Avena* genus.

The dispersed nature of the repeated sequences related to pAm2, pAm3 and pAm4 in the genome of *Avena* species was deduced from their patterns of hybridizing fragments (Figs. 3–5), which showed a variety of lengths. However, there appeared to be no simple relationship between the sizes of the major bands that were generated after digestion with either HaeIII or HindIII and then probed with any of the clones. For instance, there was no evidence of bands that were multimers of other bands, as had been the case with the DNA sequences that were homologous to pAm1.

Interspecies differences in the genomic arrangement of the repeated families represented by pAm2, pAm3 and pAm4 were deduced after comparing their characteristic fragment patterns. All of the probes and enzymes used in the present study revealed two patterns of hybridization corresponding either to DNA from an A-genome diploid

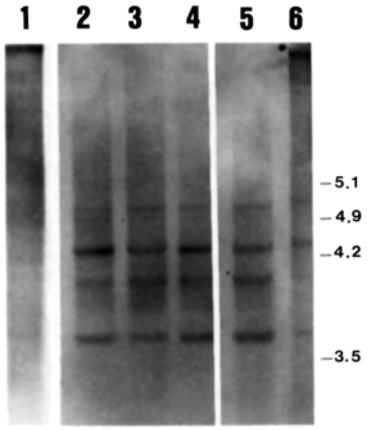

Fig. 3. HindIII-digested DNAs (5 µg per lane) of *A. pilosa* (*lane 1*), *A. canariensis* (*lane 2*), *A. longiglumis* (*lane 3*), *A. damascena* (*lane 4*), *A. strigosa* (*lane 5*) and *A. murphyi* (*lane 6*) probed with biotinylated pAm2. *Numbers* refer to molecular weight (in kb) of EcoRI-HindIII-digested lambda DNA. Digests were fractionated in a 0.8% agarose gel

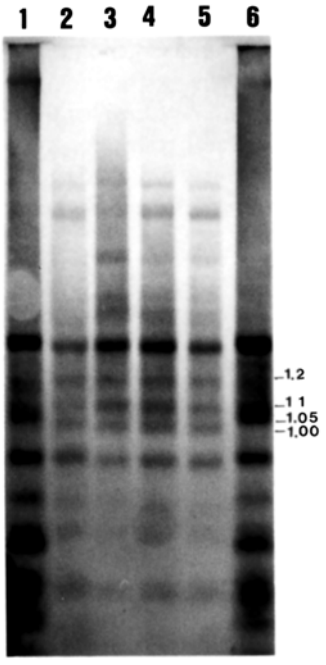

Fig. 5. HaeIII-digested DNAs (5 µg per lane) of *A. murphyi* (lane 1), *A. strigosa* (*lane 2*), *A. canariensis* (*lane 3*), *A. longiglumis* (*lane 4*), *A. damascena* (*lane 5*) and *A. pilosa* (*lane 6*) probed with biotinylated pAm4. *Numbers* refer to molecular weight (in kb) of characteristic bands (see text). Digests were fractionated in a 0.8% agarose gel

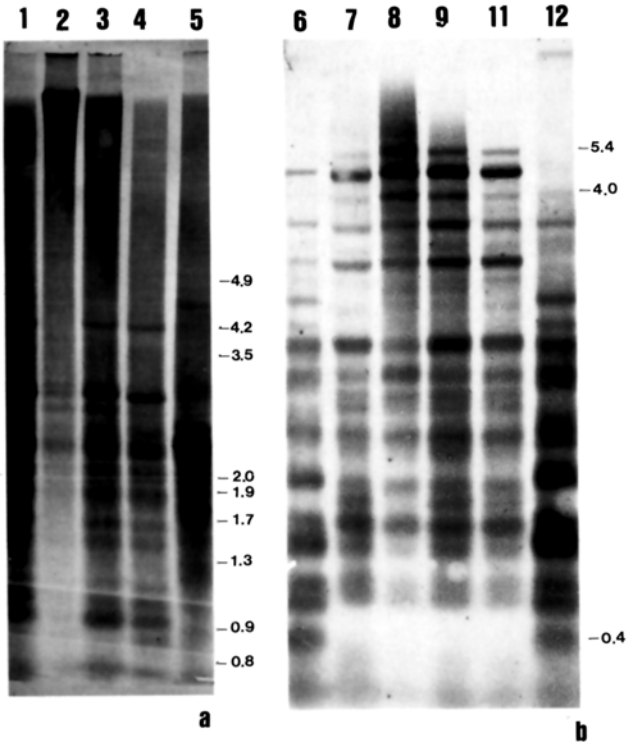

Fig. 4a, b. Southern blot analysis of *Avena* DNA using biotinylated pAm3. Five micrograms DNA from each of *A. strigosa* (*lanes 1 and 7*), *A. canariensis* (*lanes 2 and 8*), *A. longiglumis* (*lanes 3 and 9*), *A. damascena* (*lanes 4 and 11*), *A. pilosa* (*lanes 5 and 6*) and *A. murphyi* (*lane 12*) were used. **a** HindIII digestions. The *numbers* refer to molecular weight (in kb) of EcoR1-HindIII-digested lambda DNA. **b** HaeIII digestions. The *numbers* indicate molecular weight (in kb) of characteristic bands (see text). Digests were fractionated in 0.8% agarose gels

species or to DNA from a C-genome diploid species. These patterns were:

I. pAM2. Both HindIII (Fig. 3) and HaeIII digestions of C-genome DNA generated a smear of fragments over a wide range of sizes without prominent bands. On the other hand, a few well-defined bands of characteristic sizes appeared in the A-genome DNA digestions.

II. pAm3. The patterns for the HindIII and HaeIII digestions of the A-genome diploid species were different from those of the C-genome diploid species (Fig. 4). In the HaeIII digestion (Fig. 4b) two bands with high molecular weight (5.4 and 4 kb) were characteristic of the A-genome species. One band with low molecular weight (0.4 kb) was only found in the C-genome species.

III. pAm4. Most of the hybridizing HindIII and HaeIII fragments that appeared after probing with pAm4 were common to all of the species tested. Nevertheless, differences in minor bands were still detected; three bands of 1.2, 1.1 and 1 kb in the HaeIII digest (Fig. 5) were characteristic of the A-genome species.

The organization of the repeated DNA sequences throughout the genome therefore probably reflects the evolution of the *Avena* species: the conserved hybridization fragments would indicate a common origin. In contrast, the patterns that are specific to each genome would be the result of sequence rearrangements that have taken place since the evolutionary divergence of each species.

The hybridization patterns of *A. murphyi* DNA were not always additive of the two A- and C-genome patterns. This was most obvious for the restriction fragments probed with pAm4, which closely resembled those of *A. pilosa* (Fig. 5.). Similarly, several restriction fragments characteristic of the A-genome diploid species, when probed with pAm3, were not present in the *A. murphyi* DNA (Fig. 4b).

Discussion

We have studied four repeated sequence families isolated from *A. murphyi* and demonstrated that sequences related to them are present in other *Avena* species.

The C-genome-specific repetitive sequence (pAm1) that has been characterized can be used as a molecular marker in simple dot-blot experiments to distinguish whether or not an *Avena* species possesses a C-genome. Features such as the tandem arrangement of the repeating units and the length of their monomers indicate that this sequence is a satellite DNA sequence. The characteristic ladders produced by restriction digests indicate the existence of sequence heterogeneity among the members of the pAm1 family (Fig. 2). The accumulation of mutations within the monomers following by amplification of the variant repeats and deletion of older members have been proposed to explain the evolution of different eukaryotic satellites (Smith 1976; Dover 1982; Singer 1982). The absence of hybridization of the pAm1 sequence with the A-genome DNA suggests that this sequence was either eliminated from the A-genome or was extensively modified during the speciation process that originated sequences present in both genomes but with little homology between them.

The other three repeated families we have described are characterized by their dispersed nature in the genome of all the *Avena* species tested and by their genome-specific distribution patterns. Although this kind of sequence is more conserved, in contrast to satellite DNA, major differences are apparent even though the repeated units have some striking homologies. Features such as the similarity in the pattern of hybridization of the different *Avena* species and in the intensity of the hybridization signal of the repeated families reflect their different evolutionary patterns. Thus, the pAm2 family shows the most differences in the restriction profiles between genomes (Fig. 3). These differences reveal variations in the nucleotide sequence of the pAm2 members, suggesting that distinct sets of sequences have been amplified in each genome. On the other hand, the hybridization patterns showed by pAm3 and pAm4 (Figs. 4 and 5) indicate that the nucleotide sequences of each family have been fairly well conserved during the speciation of the genus,

The genome organization of the four repetitive sequences separates the *Avena* species tested into two groups according to their genome designation, thus confirming the basic pattern of speciation for diploids reported by Rajhathy and Thomas (1974).

Recently, several studies have shown that closely related species present variation in the genome organization of several repeated families. *Triticum* species (Dvorak et al. 1988), *Hordeum* species (Gupta et al. 1989; Vershinin et al. 1990), *Secale* species (Bedbrook et al. 1980) and different grasses of the *Triticeae* tribe (McIntyre et al. 1988) show characteristic hybridization patterns for at least one of the repeated sequences probed, whereas in the genus *Avena*, the repeated sequences analyzed so far (the four described in this paper and the two studied by Fabijansky et al. 1990) have failed to show differences between species belonging to the same genome group. Therefore, with the restriction enzymes used, these *Avena* probes are not informative in the search for the ancestor of the tetraploid species.

The complex organization of the repeated sequences in the allopolyploid *A. murphyi* deserves additional comment: the repeated subfamilies characteristic of each genome are unevenly represented in the DNA of this species. This is particularly evident for the pAm4 family (Fig. 5) and also for ribosomal DNA (Fabijansky et al. 1990). In both cases the distinctive fragments of the A-genome are absent in the allotetraploid. Furthermore, no additional fragment has been detected in the tetraploid with any probe. This seems to indicate that a deletion of the A-genome-specific sequence arrangement has accompained or followed the origin of the alloploid and that amplifications of new sequence arrangements have not occurred. This would confirm the essentially conserved nature of most of the sequences studied. However, since only one accession has been examined for each species, the existence of polymorphism in the ancestor of *A. murphyi* or in the tetraploid can not be ruled out.

In summary, we have characterized *Avena* repetitive DNA sequences that are either genome specific or shared by the A- and C-genomes. Their genome arrangements reflect their different evolutionary patterns, although all of them confirm the current taxonomy of *Avena* species.

Acknowledgements. The authors thank Ms. C.F. Warren of the ICE at the University of Alcalá de Henares for linguistic assistance.

References

Appels R, Moran LB (1984) Molecular analysis of alien chromatin introduced into wheat. In: Gustafson JP (ed) Gene manipulation in plant improvement. Sixteenth, Stadler Genet Symp, Columbia, pp 559–557

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

Belostotsky DA, Ananiev EV (1990) Characterization of relic DNA from barley genome. Theor Appl Genet 80:374–380
Bouchard RA (1982) Moderately repetitive DNA in evolution. Int Rev Cytol 76:113–193
Craig IL, Murray BE, Rajhathy T (1972) Leaf esterase isozymes in *Avena* and their relationship to the genomes. Can J Genet Cytol 14:581–589
Dover G (1982) Molecular drive: A cohesive mode of species evolution. Nature 299:111–117
Dvorak J, McGuire PE, Cassidy B (1988) Apparent sources of the A genomes of wheats inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. Genome 30:680–689
Fabijansky S, Fedak G, Armstrong K, Altoosar I (1990) A repeated sequence probe for the C genome in *Avena* (oats). Theor Appl Genet 79:1–7
Flavell RB (1980) The molecular characterization and organization of plant chromosomal DNA sequences. Annu Rev Plant Physiol Physiol 31:569–596
Flavell RB, Rimpau J, Smith DB (1977) Repeated sequence DNA relationships in four cereal genomes. Chromosomes 63:205–222
Flavell RB, O'Dell M, Hutchinson J (1980) Nucleotide sequence organization in plant chromosomes and evidence for sequence translocations during evolution. Cold Spring Harbor Symp Quant Biol 45:501–508
Fominaya A, Vega C, Ferrer E (1988) C-banding and nucleolar activity of tetraploid *Avena* species. Genome 30:633–638
Gupta PK, Fedak G, Molnar SJ, Wheatcroft R (1989) Distribution on a *Secale cereale* DNA repeat sequence among 25 Hordeum species. Genome 32:383–388
Hueros G, Monte JV, Ferrer E (1990) *Hordeum chilense* repetitive sequences. Genome characterization using biotinylated probes. Theor Appl Genet 80:24–32
Legget JM (1980) Chromosome relationships and morphological comparisons between the diploid oats *Avena prostrata*, *A. canariensis* and the tetraploid *A. macroccana*. Can J Genet Cytol 22:287–294
Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
McIntyre CL, Clarke BC, Appels R (1988) Amplification and dispersion of repeated DNA sequences in the *Triticeae*. Plant Syst Evol 160:39–59
Miklos GLC (1986) Localized highly repetitive DNA sequences in vertebrate and invertebrate genomes. In: MacIntyre RJ (ed) Molecular evolutionary genetics. Plenum Press, New York London, pp 241–322
Rayburn AL, Gill BJ (1986) Molecular identification of the D-genome chromosomes of wheat. J Hered 77:253–255
Rajhathy T, Thomas H (1974) Cytogenetics of oats (*Avena* L.). Misc Publ Genet Can 2:1–90
Rimpau J, Smith DB, Flavell RB (1980) Sequence organisation in barley and oats chromosomes revealed by interspecies DNA/DNA hybridization. Heredity 44:131–149
Sańchez de la Hoz P, Fominaya A (1989) Studies of isozymes in oat species. Theor Appl Genet 77:735–741
Saul M, Potrykus I (1984) Species-specific repetitive DNA used to identify interspecific somatic hybrids. Plant Cell Rep 3:65–67
Sharp PJ, Desai S, Gale MD̀ (1988) Isozyme variation and RFLPs at the β-amylase loci in wheat. Theor Appl Genet 76:691–699
Singer MF (1982) Highly repeated sequences in mammalian genomes. Int Rev Cytol 76:67–112
Smith GP (1976) Evolution of repeated DNA sequences by unequal crossover. Science 191:528–535
Vershinin AV, Salina EA, Solovyov VV, Timofeyeva LL (1990) Genomic organization, evolution and structural peculiarities of highly repetitive DNA of *Hordeum vulgare*. Genome 33:441–449
Zhao A, Wu T, Xie Y, Wu R (1989) Genome-specific repetitive sequences in the genus *Oryza*. Theor Appl Genet 78:201–209