ORIGINAL PAPER

Analysis of EST sequences suggests recent origin of allotetraploid colonial and creeping bentgrasses

David Rotter · Arvind K. Bharti · Huaijun Michael Li · Chongyuan Luo · Stacy A. Bonos · Suleiman Bughrara · Geunhwa Jung · Joachim Messing · William A. Meyer · Stephen Rudd · Scott E. Warnke · Faith C. Belanger

Abstract Advances in plant genomics have permitted the analysis of several members of the grass family, including the major domesticated species, and provided new insights into the evolution of the major crops on earth. Two members, colonial bentgrass (*Agrostis capillaris* L.) and creeping bentgrass (*A. stolonifera* L.) have only recently been domesticated and provide an interesting case of polyploidy and comparison to crops that have undergone human selection for thousands of years. As an initial step of characterizing these genomes, we have sampled roughly 10% of their gene content, thereby also serving as a starting point for the construction of their physical and genetic maps. Sampling mRNA from plants subjected to environmental stress showed a remarkable increase in transcription of transposable elements. Both colonial and creeping bentgrass are allotetraploids and are considered to have one genome in common, designated the A_2 genome. Analysis of conserved genes present among the ESTs suggests the colonial and creeping bentgrass A_2 genomes diverged from a common ancestor approximately 2.2 million years ago (MYA), thereby providing an enhanced evolutionary zoom in respect to the origin of maize, which formed 4.8 MYA, and tetraploid wheat, which formed only 0.5 MYA and is the progenitor of domesticated hexaploid wheat.

Communicated by R. Hagemann.

D. Rotter · H. M. Li · C. Luo · S. A. Bonos ·
W. A. Meyer · F. C. Belanger (✉)
Department of Plant Biology and Pathology
and The Biotechnology Center for Agriculture
& the Environment, School of Environmental
and Biological Sciences, Rutgers, The State University
of New Jersey, New Brunswick, NJ 08903, USA
e-mail: belanger@aesop.rutgers.edu

A. K. Bharti · J. Messing
The Plant Genome Initiative at Rutgers,
Waksman Institute, Rutgers,
The State University of New Jersey, Piscataway,
NJ 08854, USA

S. Bughrara
Department of Crop and Soil Science, Michigan State University,
East Lansing, MI 48824, USA

G. Jung
Department of Plant, Soil, and Insect Sciences,
University of Massachusetts-Amherst, Amherst,
MA 01003, USA

S. Rudd
Turku Center for Biotechnology,
Tykistoekatu 6, Turku, Finland

S. E. Warnke
USDA-ARS, Beltsville, MD 20705, USA

Introduction

The grass family is one of the largest, with 10,000 species, and has arisen only 60 million years ago (MYA), indicating a rapid pace of speciation (Kellogg 2001). It is also the source of modern agriculture. Except for soybean, a legume, the cereals, wheat, rice, maize, and sorghum, are the largest crops in the world and provide our major nutritional resources. Initial analysis of comparative maps of some cereal genomes indicated that conserved genetic markers are preserved in their order (synteny) and that speciation resulted in the breakage and reassortment of chromosomes (Gale and Devos 1998). Because of the c-value paradox among these species it therefore became attractive

to conceptualize an anchor genome for the grasses that is not only an important worldwide crop but also has a small genome (Messing and Llaca 1998). This concept resulted in the international effort to sequence the rice genome, which has a size of 378 Mb (International Rice Genome Sequencing Project 2005).

However, comparisons of the rice genome with other members of the grass family have shown that they not only differ in their expansion of chromosomes but also in their micro-colinearity (Tikhonov et al. 1999). Moreover, it appeared that gene movement has lead to gene insertions that disrupt syntenic alignments of orthologous regions from different species (Song et al. 2002). Furthermore, genome organizations differed in their response to polyploidization. Although in hexaploid wheat homoeologous chromosomes are largely preserved due to the *Ph1* locus (Griffiths et al. 2006), maize underwent significant changes not only in terms of gene loss but also in chromosome expansion (Bruggmann et al. 2006). Therefore, our previous beliefs from comparative genetic maps that grasses represent simple integrative genetic systems has to be reconsidered and understanding the speciation processes are more important steps in investigating the divergence, organization, and functions of plant genes than previously anticipated.

We believe that commonly used turfgrasses could provide such interesting speciation events that could provide additional perspectives to the evolution of the grass family and new approaches to breeding. Colonial and creeping bentgrasses (*A. capillaris* L. 2n = 4x = 28, $A_1A_1A_2A_2$ and A. *stolonifera* L. 2n = 4x = 28, $A_2A_2A_3A_3$, respectively) are closely related commercially important turfgrass species that are used extensively on golf courses in temperate regions (Warnke 2003; Ruemmele 2003). Bentgrasses are commonly used on the greens, tees, and fairways where they form a dense turf that can be maintained at low cutting heights. Currently breeding programs for both creeping and colonial bentgrasses exist in research universities and in the private sector. These programs have been successful in developing and releasing improved cultivars to the public (Meyer and Funk 1989). Approximately £2.8 million of creeping bentgrass seed and approximately £1.5 million of colonial bentgrass seed were produced in Oregon in 2004 (Young 2004).

However, not much is known about the gene content of these species. Still in its infancy, genetic linkage maps based on molecular markers have been developed for creeping bentgrass (Chakraborty et al. 2005; Bonos et al. 2005) and a linkage map of colonial bentgrass is currently being developed as well (Rotter et al. 2006). In the absence of a physical map and as a first step to assess their gene content, we generated 8,470 creeping bentgrass EST sequences and 7,528 colonial bentgrass sequences. We

anticipate that these sequences will serve as an important resource for the development of maps with higher densities of markers and therefore enable marker-assisted breeding and discovery of the genes conferring desirable traits.

For instance, one of the major management problems for creeping bentgrass is the fungal disease dollar spot (Walsh et al. 1999). There are no current cultivars of creeping bentgrass that are completely resistant to dollar spot, but there is a considerable variation in the degree of susceptibility (Bonos et al. 2003; Chakraborty et al. 2006). In contrast, cultivars of colonial bentgrass generally have good resistance to dollar spot (Plumley et al. 2000) and, therefore, may be a source of novel genes or alleles that could be used for genetic improvement of creeping bentgrass (Belanger et al. 2003, 2004). The fungal agent of dollar spot is referred to as *Sclerotinia homoeocarpa* F.T. Bennett. However, it is widely recognized that the taxonomic classification is incorrect and the fungus should be reclassified in one of the genera *Rutstroemia*, *Lanzia*, *Moellerodiscus*, or *Poculum* (Carbone and Kohn 1993; Holst-Jensen et al. 1997). Formal reclassification has not been done due to lack of information on the teleomorph (Viji et al. 2004). Differences in disease resistance between colonial and creeping bentgrass could be analyzed based on genetic maps and by differential gene expression. Both approaches will be facilitated by the availability of the ESTs.

Besides building a resource for gene content and expression, we also used sequence-clustering techniques to identify conserved gene copies to determine the relationships of the progenitors of these genomes. Both colonial and creeping bentgrasses are allotetraploids, having the A_2 genome in common (Jones 1956b). Analysis of conserved genes (Fulton et al. 2002) present among the ESTs allowed us to estimate the age of divergence between the two subgenomes within each species and between the shared colonial and creeping bentgrass A_2 genomes.

Materials and methods

RNA isolation and cDNA library construction

The RNA was isolated from an individual creeping bentgrass plant and an individual colonial bentgrass plant that were part of the 2002 field test (Belanger et al. 2004). Leaf samples were collected 56 days after inoculation with the dollar spot fungus and stored at −80°C. For RNA isolation, leaf samples (1 g) were ground to a fine powder with liquid nitrogen and resuspended in 10 ml Tri-Reagent (Sigma-Aldrich, St. Louis, MO, USA). Debris was removed by centrifugation and the supernatant was extracted twice with chloroform. RNA in the aqueous layer was precipitated with isopropanol, and the RNA pellet was washed once

with ethanol and dissolved in water. Poly(A^+) RNA was isolated from the total RNA using a commercial kit (Oligotex mRNA Purification Midi Kit, Qiagen USA, Valencia, CA, USA).

The cDNA libraries were constructed from the creeping bentgrass and colonial bentgrass poly(A^+) RNA. cDNA synthesis and phage packaging were carried out by using a commercial kit (λ ZAP-Express cDNA Library Construction Kit, Stratagene, La Jolla, CA, USA). The primary creeping and colonial bentgrass cDNA libraries contained 0.4×10^6 and 1×10^6 plaque-forming units, respectively. The primary libraries were amplified and stored at $-80°C$, as described by the manufacturer.

EST sequencing

pBK-CMV phagemids were mass excised from the λ ZAP-Express vector, transformed into bacterial strain XLOLR, and plated onto Luria-Bertani (LB) plus kanamycin medium as described by the manufacturer. Approximately 2,000 phagemids per plate were grown in 243×243 mm^2 BioAssay Dishes (Nalge Nunc International Corp., Rochestser, NY, USA). To allow discrimination between recombinant and nonrecombinant phagemids, a solution containing 1 ml LB, 1 ml of 100 mm isopropylthio-β-D-galactoside (IPTG), and 200 μl of 100 mg ml^{-1} 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (XGAL) was spread on the agar surface of each plate prior to plating the bacterial cultures. White colonies were picked on a 'Q' Pix2 automated colony picker (Genetix USA Inc., Boston, MA, USA) into 96 well plates containing 200 μl per well of LB plus kanamycin freezing medium (Zimmer and Gibbins 1997). The cultures were incubated at $37°C$ for 16 h, duplicated, and stored at $-80°C$.

The glycerol stocks were used to inoculate deep well plates with 1.2 ml of LB plus kanamycin using a 'Q' Bot robot (Genetix USA Inc.). Plasmid DNA was isolated from these overnight cultures using the alkaline lysis method (Birnboim and Doly 1979) in a 96-well format using Whatman filters (Whatman Inc., Clifton, NJ, USA). The DNA was pelleted using isopropanol and dissolved in 1 mM Tris–Cl, pH 8.0. The cDNA inserts were sequenced from the 3′ end with the T7 primer using an ABI PRISM BigDye Terminator v3.1 Cycle Sequencing Ready Reaction kit (Applied BioSystems, Foster City, CA, USA). The unincorporated dyes were removed by solid phase reversible immobilization (SPRI) using magnetic beads (Agencourt Biosciences Corp., Beverly, MA, USA). The reactions were then suspended in 0.1 mM EDTA and run on an ABI 3730xl automated capillary sequencer with 50 cm capillary arrays. Base-calling was done using the ABI KB basecaller with lower quality sequences being trimmed using the "Lucy" software (TIGR) for quality control (.Q16).

Sequences were also screened for vector contamination. All sequences >100 bp were processed for GenBank submission. A total of 15,998 successful sequence reads were obtained and deposited to GenBank.

Computational analysis

ABI trace files were converted to fasta format by using the AutoEditor program (TIGR). Cleaned sequences were assembled into contigs by using the CAP3 program (Huang and Madan 1999). The ESTs were imported into the openSputnik EST database and clustered into unigenes for functional annotation (Rudd 2005). Functional assignments were performed using BLASTX (threshold value of 1×10^{-10}) against the MIPS catalog of functionally assigned proteins (FunCat; Rueppe et al. 2004).

Reverse transcriptase-PCR (RT-PCR) amplification of creeping bentgrass sequences

Two of the creeping bentgrass sequences (accessions EH667188 and EH667189) used for the subgenome comparisons were obtained by RT-PCR to extend the coding sequences from the original EST accessions DV860785 and DV865904, respectively. Total RNA was isolated from the creeping bentgrass plant by using Tri-Reagent, as described earlier. First strand cDNA was synthesized using a polyT primer, the CDSIII/3′ primer from the Smart cDNA Library Construction Kit (Clontech Laboratories, Inc., Mountain View, CA, USA). One microgram of total RNA and 1 μl of 10 mM CDSIII/3′ primer were mixed in a total volume of 5 μl, heated at $70°C$ for 5 min and cooled on ice for 2 min. To this mixture was then added 2 μl of M-MLV RT buffer, 1 μl of 10 mM dNTP mix, 1 μl of M-MLV reverse transcriptase (Fisher Scientific, Pittsburgh, PA, USA), and 0.8 μl of Rnasin RNase inhibitor (Promega Corporation, Madison, WI, USA). The reaction was carried out at $42°C$ for 60 min and then stopped by cooling on ice.

PCR amplification of the entire cDNA product was carried out with gene-specific primers by adding 5 μl of 10X Taq PCR buffer, 4 μl of 2.5 mm dNTP mix, 5 μl of the forward and reverse primers at 10 pmol μl^{-1}, and 1 μl of AmpliTaq DNA Polymerase (Applied Biosystems, Inc.) in a final volume of 50 μl. The 3′ primer for amplification (5′-GCAAGTTAGTGGAACAAGTTAAGAAC-3′) was designed based on the creeping bentgrass EST sequence DV865904. The 5′ primer (5′-GGTGGCTCCCTCTTGTCATA-3′) was designed based on the coding sequence of the colonial bentgrass EST sequence DV855725. The reaction was carried out in a GeneAmp 9700 thermocycler (Applied Biosystems Inc.). The initial denaturation was conducted at $94°C$ for 2 min, followed by 40 cycles of 30 s denaturation at $94°C$, 30 s annealing at $56°C$, and 1 min extension at

72°C, followed by a final extension at 72°C for 7 min. The PCR products were purified with a QIAquick PCR purification Kit (Qiagen). The purified PCR products were ligated into the plasmid pGEM-T-Easy (Promega Corporation) and transformed by electroporation into XL1-Blue MRF' cells (Stratagene). Plasmids from several colonies were purified with a QIAprep Spin Miniprep kit (Qiagen) and sequenced using the forward PCR primer (GeneWiz, Inc., North Brunswick, NJ, USA).

Phylogenetic analysis

The CLUSTAL-X (Thompson et al. 1997) program was used to align DNA sequences. Phylogenetic analysis was performed with the PAUP program (version 4.0b10 for Macintosh; Swofford 2002). Phylogenetic trees were generated using the maximum parsimony (1,000 bootstrap replications) and the neighbor-joining methods. The K_s and K_a values were determined using the MEGA3.1 program (Kumar et al. 2004).

Results

EST generation

As a first step to gain insight into the gene content and the evolution of colonial and creeping bentgrass, we constructed cDNA libraries for both. To simultaneously facilitate discovery of differential gene expression in response to environmental stress, we took advantage of the fact that colonial bentgrass generally has good resistance to dollar spot while creeping bentgrass does not. Therefore, leaf tissue from creeping and colonial bentgrass plants from the 2002 field test was taken after inoculation with the dollar spot fungus, at a time when the creeping bentgrass plant was showing symptoms of dollar spot disease (Belanger et al. 2004). The colonial bentgrass plant was not exhibiting symptoms of the disease. The creeping and colonial bentgrass cDNA libraries are therefore expected to include clones for genes that were induced in response to environmental stress and in particular to the fungal inoculum.

The creeping and colonial bentgrass cDNAs were cloned unidirectionally into the λ ZAP Express vector. The cDNA clones were sequenced from the 3′ ends to generate sequence data from the untranslated region of the transcripts. The 3′ untranslated regions of genes are generally more polymorphic than the coding sequences, and therefore more useful for molecular marker development (Bhattramakki et al. 2002; Brady et al. 1997). The characteristics of the colonial and creeping bentgrass ESTs are summarized in Table 1. A total of 10,944 randomly chosen cDNA clones from each library were single pass sequenced from

the 3′ end. From the creeping bentgrass cDNAs, 8,470 usable sequences were obtained with an average read length of 567 bases after vector trimming. From the colonial bentgrass cDNAs, 7,528 usable sequences were obtained with an average read length of 745 bases after vector trimming. These sequences have been deposited in the NCBI dbEST (colonial bentgrass accession nos: DV852741–DV860268; creeping bentgrass accession nos: DV860269–DV868738) and to the openSputnik EST database (http://www.sputnik.btk.fi/ests; Rudd 2005). The sequences can also be searched at http://www.aesop.rutgers.edu/~belangerlab/research.htm.

Chloroplast and microbial sequences

To determine the level of chloroplast contamination among the ESTs, the colonial and creeping bentgrass sequences were compared with the wheat chloroplast genome sequence (Ogihara et al. 2002) by a BLASTN search. Some sequences nearly identical to the wheat chloroplast genome were present among the colonial and creeping bentgrass ESTs (Table 1). The contamination level was 0.5 and 1.4% for the colonial and creeping bentgrass ESTs, respectively.

Since the cDNA libraries were constructed from field grown plants that were inoculated with the dollar spot fungus, some microbial sequences would be expected among the ESTs. To estimate the level of microbial ESTs in the libraries, the colonial and creeping bentgrass ESTs were compared to the 14,522 protein sequences of the plant pathogen *Sclerotinia sclerotiorum* (*Sclerotinia sclerotiorum* Sequencing Project, Broad Institute of Harvard and MIT, http://www.broad.mit.edu) using a BLASTX search. Although the dollar spot fungus should not be classified in

Table 1 Characteristics of the colonial and creeping bentgrass EST sequences

	Colonial ESTs	Creeping ESTs
EST count	7,528	8,470
Total nucleotides	5,611,738	4,809,144
Average EST length (nt)	745	567
Unigenes	4,425	4,747
Multimember	1,095	980
Singleton	3,330	3,767
Average unigene length (nt)	795	555
Apparent coding potential	70.48%	54.53%
Colonial/creeping shared ESTs	5,754 (76.4%)	6,911 (81.6%)
Colonial/creeping unique ESTs	1,774 (23.6%)	1,559 (18.4%)
BlastX match	6,548 (87.0%)	6,190 (73.1%)
No BlastX match	980 (13.0%)	2,280 (26.9%)
Chloroplast sequences	38 (0.5%)	123 (1.4%)
Microbial sequences	184 (2.4%)	99 (1.2%)

the genus *Sclerotinia* (Holst-Jensen et al. 1997), *S. sclerotiorum* is the most closely related species for which there is a whole genome sequence available. For the bentgrass ESTs that had a match to a *S. sclerotiorum* protein sequence, the expected value was compared with the expected value to the best match in a BLASTX search of the NCBI database. Those bentgrass ESTs for which the expected value was lower with the *S. sclerotiorum* match than with the NCBI match were considered to be likely of microbial origin. This comparison was done because most of the *S. sclerotiorum* protein sequences were not included in the NCBI database. For the colonial and creeping bentgrass ESTs, 184 and 99 sequences, respectively, met these criteria. Sequences of microbial origin in the libraries are therefore estimated to be 2.4% for colonial bentgrass and 1.2% for creeping bentgrass.

Unigene analysis

The sequences from each library were imported into the openSputnik EST database and clustered into unigenes (Rudd 2005). Analysis of the colonial bentgrass sequences resulted in 1,095 multimember unigenes and 3,579 singletons. Analysis of the creeping bentgrass sequences resulted in 980 multimember unigenes and 3,901 singletons. Recent estimates of the gene content of another ancient tetraploid grass species, maize, have placed its total gene number between 42,000 and 56,000

genes (Haberer et al. 2005), whereas the total predicted gene number for the sequenced rice genome was 37,544 (International Rice Genome Sequencing Project 2005). Therefore, we estimated that we had obtained sequences from roughly 10% of the total gene content of the two turfgrass species.

The unigene sequences were assigned to functional categories based on the MIPS catalog of functionally assigned proteins (Ruepp et al. 2004) (Table 2). Of the 4,425 colonial bentgrass unigenes, 1,843 could be classified. Of the 4,747 creeping bentgrass unigenes, 1,048 could be classified. There were more creeping bentgrass unigenes with no match. This may be a reflection of the shorter average length of the creeping bentgrass sequences and the creeping bentgrass unigenes. Since the sequences were from the 3′ ends, there may be inadequate coding sequence information in some of the creeping bentgrass ESTs on which to base a match for functional classification. Of those unigenes that could be assigned a functional classification, the percentage distribution between colonial and creeping bentgrass was similar, with one exception. There was a 10-fold higher representation of creeping bentgrass unigenes in the category of transposable elements, relative to that of colonial bentgrass. The difference in transposable elements between the creeping and colonial ESTs is discussed later. Annotation of the unigenes within each functional category is available on the openSputnik website (http://www.sputnik. btk.fi/ests).

Table 2 Distribution of the functional classes of colonial and creeping bentgrass unigenes

Category	Colonial ESTs		Creeping ESTs	
Number of unigenes represented in the FunCat analysis	1,843	42%	1,048	28%
Number of unigenes with no match	2,582	58%	3,699	78%
Metabolism	500	19.4%	244	18.3%
Energy	256	9.9%	134	10.1%
Cell growth, cell division, DNA synthesis	62	2.4%	33	2.5%
Transcription	172	6.7%	63	4.7%
Protein synthesis	122	4.7%	120	9.0%
Protein destination	151	5.9%	102	7.7%
Transport facilitation	159	6.2%	65	4.9%
Cellular transport and transport mechanisms	86	3.3%	40	3.0%
Cellular biogenesis	119	4.6%	48	3.6%
Cellular communication/signal transduction	264	10.2%	129	9.7%
Cell rescue, defense, cell death and ageing	262	10.2%	119	8.9%
Ionic homeostasis	2	0.1%	1	0.1%
Cellular organization	358	13%	155	11.7%
Motility	1	0.04%	0	0%
Tissue specificity	0	0%	0	0%
Development	47	1.8%	19	1.4%
Transposable elements, viral and plasmid proteins	10	0.4%	57	4.2%
Organism specific proteins	3	0.1%	0	0%
Number of matches[a]	2,574		1,329	

[a] Number of matches exceeds the number of unigenes represented because the same unigene may have more than one match from different role categories

Transposable element sequence representation in the colonial and creeping bentgrass ESTs

In the FunCat analysis of the unigenes, transposable element sequences represented a 10-fold higher percentage of the creeping bentgrass unigenes, relative to the colonial bentgrass unigenes. The distribution of transposable element ESTs between the two species was, therefore, further investigated. Transposable elements are found in all plant genomes and are classified into two groups, the Class I retroelements (retrotransposons with LTRs and retrotransposons without LTRs) and the Class II DNA transposons (hAT, CACTA, *Mutator*, MITEs, Helitrons, etc.). Many grass genomes are largely composed of retrotransposon sequences (Bennetzen 2000; Feschotte et al. 2002; Messing et al. 2004; Haberer et al. 2005). An analysis of 7.8×10^5 EST sequences from numerous plant species revealed that retrotransposons were active and represented 0.12% of the ESTs (Vicient et al. 2001). Some of the colonial and creeping bentgrass EST sequences were most similar to retrotransposons described from other species, particularly rice. The colonial and creeping bentgrass ESTs were systematically searched with the coding sequences of numerous rice retrotransposons of different classes and the best matches were tabulated. The results are summarized in Table 3. The expected values of the matches ranged from 8×10^{-5} to 8×10^{-128}. For all the ESTs included in the data summarized in Table 3, a retrotransposon sequence was the best match. There is a striking difference in the representation of retrotransposons among the colonial bentgrass ESTs, relative to that among the creeping bentgrass ESTs. In the colonial bentgrass library, retrotransposon ESTs were 0.18% of the total, similar to the 0.12% reported from the survey of ESTs from many plant species (Vicient et al. 2001). In contrast, in the creeping bentgrass EST library retrotransposon sequences represented 1.4% of the total, an 8-fold higher representation than in the colonial bentgrass library.

The Class II DNA elements are found in lower copy numbers than the retrotransposons. The bentgrass ESTs

were searched with several Class II DNA element coding sequences and some significant matches were found (Table 3). Similar to the case with the retrotransposons, there was a 16-fold higher representation of the Class II DNA elements among the creeping bentgrass ESTs, relative to the colonial bentgrass ESTs.

This difference in the transposable element representation between the creeping bentgrass and colonial bentgrass ESTs shows an interesting correlation between expressed transposable elements and response to environmental stress. Stress was phenotypically measurable because the creeping bentgrass sample was exhibiting symptoms of dollar spot disease, whereas the colonial bentgrass sample was not. Further studies of this aspect would call for microarray or 454 sequencing analysis of different mRNA sampling from different types of stress conditions. Nevertheless, it has been shown that transposable elements can be activated by various abiotic and biotic stresses, including pathogen attack (McClintock 1984; Casacuberta and Santiago 2003; Grandbastien 1998; Kimura et al. 2002; Wessler 1996).

Phylogenetic analysis

Taxonomically, the genus *Agrostis* is placed in the subfamily Pooideae of the grass family Poaceae. *Agrostis* is assigned to the tribe Aveneae based on morphological characters and chloroplast DNA restriction site characters (Renvoize and Clayton 1992; Watson 1990; Soreng and Davis 1998). *Agrostis* species have not previously been included in any phylogenetic comparisons with other grass species based on DNA sequence data. To better establish the relationship of *Agrostis* to the cereal grasses, we generated phylogenetic trees based on our EST sequences (Fig. 1; Tables 4, 5).

For this type of analysis, it is important to select putative orthologous sequences from all the species involved. Identification of orthologous genes usually requires positional information so that their colinearity and descent from a common ancestral chromosome can be established. Such information cannot be obtained from cDNA sequences. However, clustering methods have shown that orthologous and paralogous sequences can be distinguished by their divergence rates (Xu and Messing 2006). Fulton et al. (2002) first proposed the concept of conserved ortholog set (COS) markers for comparative genomics. COS markers are single copy genes that have been conserved throughout evolution, and can therefore be considered as orthologs (Fulton et al. 2002). The creeping and colonial bentgrass ESTs were therefore systematically searched for COS genes. COS genes have been identified in rice (1,290) in the Michelmore lab (http://www.cgpdb.ucdavis.edu/COS Arabidopsis/). Here, creeping and colonial bentgrass EST sequences were compared to the rice COS genes and orthologs

Table 3 Colonial and creeping bentgrass transposable element ESTs

	# Colonial ESTs	# Creeping ESTs
Class I—retroelements		
Ty1-copia-like elements	7	16
Ty3-gypsy-like elements	3	88
Unclassified retrotransposon	4	16
Total	14 (0.18%)	120 (1.4%)
Class II—DNA transposons		
hAT Superfamily	0	2
CACTA En/Spm	1	12
Total	1 (0.01%)	14 (0.16%)

Fig. 1 Rooted maximum parsimony phylogenetic trees comparing COS sequences from colonial bentgrass and creeping bentgrass with orthologous sequences from cereal grasses. The gene identifications and the GenBank accession numbers for the sequences used in the analysis are presented in Tables 4 and 5. The maize sequence was designated as the outgroup for rooting the trees. The *numbers* at the nodes are the bootstrap percentages based on 1,000 replications
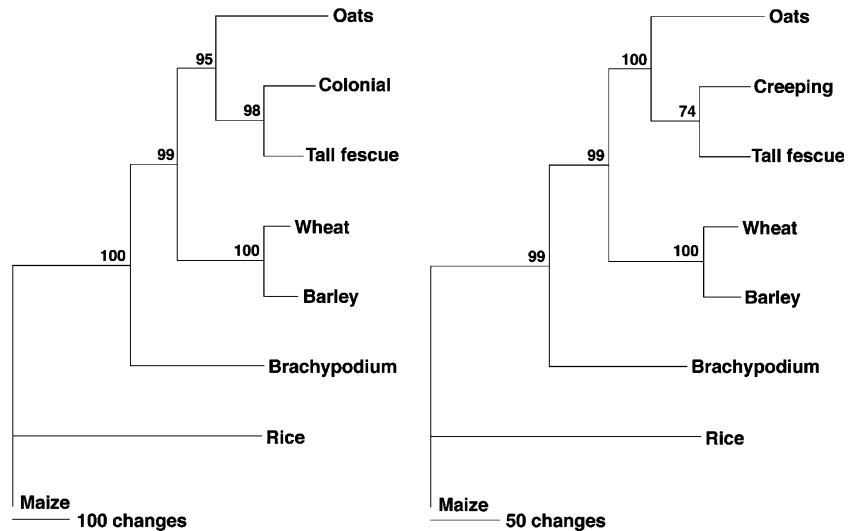
Left tree: Oats, Colonial, Tall fescue (95, 98), 99, Wheat, Barley (100), Brachypodium, 100, Rice, Maize, 100 changes

Right tree: Oats, Creeping, Tall fescue (100, 74), 99, Wheat, Barley (100), Brachypodium, 99, Rice, Maize, 50 changes

Table 4 Accession numbers of sequences used for the colonial bentgrass phylogenetic analysis presented in Fig. 1

Gene identification	Length, nt	Colonial bentgrass	Tall fescue	Oats	Wheat	Barley	Brachypodium	Rice	Maize
Photosystem I, subunit N	331	DV857165	DT699693	CN817892	DR740275	X66428	DV483607	NM_001072866	DQ244519
Chloroplast inositol phosphatase	364	DV855844	DT681526	CN817763	AY736122	AY934272	DV475324	CT833053	AY104484
Photosystem I, subunit psaK	424	DV853423	DT696505	CN819088	CV759509	L12707	DV473885	NM_001065430	AY103750
Unknown protein	222	DV855067	DT690743	CN815299	CJ581175	CV063834	DV470209	NM_001065286	AY107008
Unknown protein	297	DV855223	DT680969	CN817514	CJ703194	BJ453469	DV483494	CT829478	AY104509
Ribosomal protein P1	194	DV857454	DT695295	CN817208	AY290723	DN186825	DV469825	CT830647	DQ244520
Photosystem II, 23 kDa subunit	452	DV853316	DT714267	CN818843	X57407	BF266992	DV479214	NM_001065398	DQ245893
Membrane protein	279	DV853516	DT682707	CN819120	CJ714688	BQ459226	DV473901	CI029831	CD998231
Mg-protoporphyrin IX monomethyl ester cyclase	393	DV854211	DT700471	CN819692	CJ662371	AV938975	DV477204	CX099627	DT944112
50S Ribosomal protein L18	277	DV856175	DT687644	CN820108	CV778756	BM370151	DV474922	CI034243	DV549463
Mitochondrial ATP synthase D chain	442	DV858277	DT701407	CN820057	CJ535602	CX630548	DV471336	CX109377	EE168369

of 177 and 161 rice COS genes were found among the colonial and creeping bentgrass ESTs, respectively. To maximize the total number of sequences that could be used for the phylogenetic analysis, separate trees for colonial bentgrass and creeping bentgrass COS sequences were generated because there was not much overlap between the two bentgrass COS datasets. These colonial and creeping bentgrass COS sequences were used to identify similar sequences from *Avena sativa* (oats), *Festuca arundinacea* (tall fescue), *Hordeum vulgare* (barley), *Triticum aestivum* (wheat), *Brachypodium distachyon*, *Oryza sativa* (rice) and *Zea mays* (maize). Overall, for 11 and 7 of the colonial bentgrass and creeping bentgrass COS sequences, respectively,

similar sequences were available for all the other species. These genes were confirmed to be single copy genes in rice by BLASTX searches in Gramene (http://www.gramene.org/). For each gene, the sequences were aligned and trimmed to include only the region of coding sequence overlap for all the species. The trimmed sequences for all the genes were combined and a phylogenetic tree generated from a maximum parsimony analysis (Fig. 1). The maximum parsimony analysis of the combined dataset for the colonial bentgrass tree was based upon 3,675 total characters, of which 2,820 were constant, 483 variable characters were parsimony uninformative, and 372 characters were parsimony informative. The creeping bentgrass tree was

Table 5 Accession numbers of sequences used for the creeping bentgrass phylogenetic analysis presented in Fig. 1

Gene Identification	Length, nt	Creeping bentgrass	Tall fescue	Oats	Wheat	Barley	Brachypodium	Rice	Maize
Photosystem I, subunit psaK	261	DV861783	DT682777	CN819088	CV781092	L12707	DV473885	CT830590	AY103750
Porphobilinogen deaminase	350	DV861883	DT715353	CN820915	AY052154	BQ663234	DV480877	NM_001052552	AY107718
Photosystem I, subunit N	260	DV861960	DT698093	CN818960	CJ731306	X66428	DV483607	NM_001072866	DQ244519
Ferredoxin-thioredoxin reductase	102	DV863300	DT703681	CN818623	CD931726	CB870362	DV475078	NM_001069210	DQ246053
CBS domain containing protein	239	DV863405	DT687222	CN818511	CJ522253	BM816271	DV478349	NM_001057733	AY104325
Hypothetical protein	160	DV864281	DT690743	CN814648	CV778458	AL450589	DV470209	NM_001065286	AY107008
Chloroplast inositol phosphatase-like protein	382	DV866247	DT681526	CN817763	AY377991	BF262745	DV475324	NM_001066511	AY104484

based on 1,754 total characters, of which 1,361 were constant, 234 variable characters were parsimony uninformative, and 159 characters were parsimony informative. The maize sequence was designated as the outgroup for rooting the trees. Neighbor-joining analysis of the data generated trees of identical topology (not shown). In both the maximum parsimony and neighbor-joining analyses colonial bentgrass and creeping bentgrass were closer to tall fescue (tribe Poeae) than to oats (tribe Aveneae).

## Estimation of the divergence time of the bentgrass subgenomes

An extensive cytological investigation into some *Agrostis* spp. and their interspecific hybrids was reported by Jones (1956a, b, c). Based on the cytological results, Jones (1956b) proposed a model for genome organization of creeping and colonial bentgrass in which both species were considered to be allotetraploids, having one ancestral genome in common. Both species were found to have 14 chromosome pairs. The genome organization of colonial bentgrass was designated as $A_1A_1A_2A_2$ and that of creeping bentgrass $A_2A_2A_3A_3$ (Jones 1956b). The A_1 subgenome of colonial bentgrass was considered to be related to the diploid species *A. canina* (velvet bentgrass). The diploid origins of the A_2 and A_3 genomes are unknown. Analysis of marker segregation in creeping bentgrass demonstrated that inheritance is strictly disomic (Chakraborty et al. 2005) so distinct subgenomes are expected.

As allotetraploids, both creeping and colonial bentgrass would be expected to have homoeologous COS genes. Sequence comparisons of orthologous creeping and colonial bentgrass genes will be helpful in the identification of the subgenomes of both species. For some of the COS genes represented in the bentgrass ESTs, multiple sequences for each species were obtained. There were four such cases where two distinct sequence types were found among each of the colonial and the creeping COS ESTs.

Maximum parsimony phylogenetic analyses of these homoeologous sequence sets revealed that some creeping and colonial sequences were more similar to each other than they were to the other similar sequences from within each species (Fig. 2). These results are consistent with expectations based on the previous cytological work. Since the A_2 genome is shared between colonial and creeping bentgrass (Jones 1956b), the A_2 genome sequences are expected to be more similar to each other than to the A_1 or A_3 genome sequences. Subgenome assignments could therefore be made from the phylogenetic analyses of the sequences. The orthologous rice sequence was used to root the trees. The tree topologies were the same using either the maximum parsimony or neighbor-joining methods (data not shown). For the phylogenetic analyses presented in Fig. 2, both the available coding sequences and the 3′ untranslated regions of the EST sequences were used. The sequences of the creeping bentgrass accessions EH667188 and EH667189 were generated by RT-PCR because the original EST sequences did not include much coding sequence.

The divergence times of the subgenome pairs were estimated by using an approach similar to that used to estimate the age of polyploidy in *Gossypium* (cotton) (Senchina et al. 2003). First, the average divergence rate for each gene was calculated by calibration to rice using the formula $T = K_s/2r$ (Senchina et al. 2003). The divergence time (T) between rice and the Aveneae has been estimated to be 46 MYA (Gaut 2002). The average K_s (substitutions at synonymous sites) and K_a (substitutions at nonsynonomous sites) for the coding sequence of each gene relative to rice was determined by comparing the colonial and creeping bentgrass coding sequences to the orthologous rice sequence. The low K_a/K_s ratios for the four COS genes used for the divergence time estimates indicate these genes are under purifying selection. From the average K_s for each gene, the rate of divergence of synonymous sites (r) for each gene could therefore be calculated (Table 6). The

Fig. 2 Rooted maximum parsimony phylogenetic tree comparing homoeologous colonial and creeping bentgrass COS sequences. The orthologous rice sequence was designated as the outgroup for rooting the trees. The *numbers* at the nodes are the bootstrap percentages based on 1,000 replications. The total characters/constant characters/parsimony uninformative variable characters/parsimony informative characters for the analyses presented in **a–d** are **a** (508/400/91/17), **b** (547/412/126/9), **c** (703/528/151/24), and **d** (603/461/130/12)
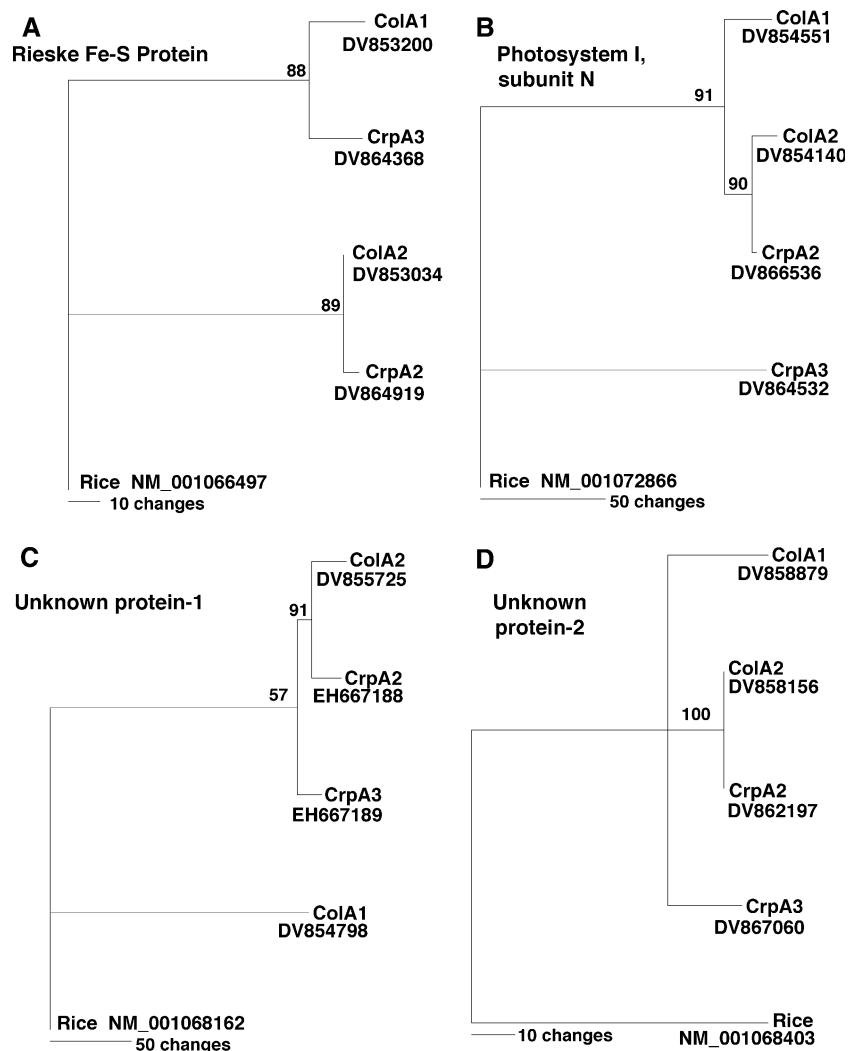
calculated rates varied from 4.3×10^{-9} to 6.7×10^{-9} substitutions per synonymous site per year, which are similar to rates calculated for other grass species (Gaut et al. 1996; Swigonova et al. 2004; Devos et al. 2005).

The calculated rates of divergence of each gene were then used to calculate the time of divergence between the colonial A_1/colonial A_2, creeping A_2/creeping A_3, and colonial A_2/creeping A_2 subgenome pairs, based on the K_s value determined for each pair (Table 7). The mean divergence times of the colonial and creeping bentgrass subgenomes were similar, at 8.9 and 10.6 MYA, respectively. The mean divergence time of the colonial and creeping bentgrass A_2 subgenomes was 2.2 MYA, suggesting a recent origin of these allotetraploid species.

Discussion

The generation of colonial and creeping bentgrass ESTs presented here is a first step in applying a genomics approach to the analysis of the gene content and the evolution of these commercially important turfgrasses. We have made available 7,528 and 8,470 colonial and creeping bentgrass EST sequences, respectively. Based on the unigene clustering, the redundancy of the EST sequences was 56% for both colonial and creeping bentgrass. This relatively low redundancy level suggests a high gene number for these species and that additional cDNA sequencing would be a cost-effective way to obtain additional gene content information.

Analysis of COS sequences found among the ESTs was used to place *Agrostis* in a phylogeny of other Poaceae. Although *Agrostis* is traditionally considered a member of the tribe Aveneae (Renvoize and Clayton 1992; Watson 1990; Soreng and Davis 1998), both the maximum parsimony and neighbor-joining analyses placed it closer to tall fescue, a member of the tribe Poeae, rather than to oats, also in the tribe Aveneae. The tribes Aveneae and Poeae are closely related (Kellogg 1998) and in some phylogenies have been combined, although this point remains unresolved (Grass Phylogeny Working Group 2001). The

Table 6 Estimation of the divergence rate of colonial and creeping bentgrass sequences relative to rice

Gene identification	Length, nt	Subgenome	K_s	K_a	Average K_s	Average divergence rate
Rieske Fe-S protein	303				0.592	6.43×10^{-9}
		ColA1	0.667 ± 0.169	0.05 ± 0.017		
		ColA2	0.503 ± 0.119	0.05 ± 0.017		
		CrpA2	0.554 ± 0.131	0.05 ± 0.017		
		CrpA3	0.645 ± 0.158	0.05 ± 0.017		
Photosystem I, subunit N						
	399				0.396	4.30×10^{-9}
		ColA1	0.417 ± 0.087	0.068 ± 0.017		
		ColA2	0.4 ± 0.085	0.08 ± 0.018		
		CrpA2	0.379 ± 0.08	0.065 ± 0.016		
		CrpA3	0.388 ± 0.078	0.078 ± 0.019		
Unknown protein-1						
	527				0.462	5.02×10^{-9}
		ColA1	0.431 ± 0.075	0.027 ± 0.008		
		ColA2	0.482 ± 0.081	0.029 ± 0.009		
		CrpA2	0.474 ± 0.08	0.027 ± 0.008		
		CrpA3	0.518 ± 0.091	0.027 ± 0.008		
Unknown protein-2						
	396				0.621	6.75×10^{-9}
		ColA1	0.621 ± 0.121	0.094 ± 0.021		
		ColA2	0.611 ± 0.12	0.092 ± 0.021		
		CrpA2	0.611 ± 0.12	0.092 ± 0.021		
		CrpA3	0.644 ± 0.126	0.091 ± 0.02		

The accession numbers of the sequences used in these analyses are given in Fig. 2

phylogenetic analysis presented here is consistent with the close relationship of the Aveneae and Poeae tribes.

To examine the relationships of the bentgrass subgenomes to each other we estimated the divergence times of the subgenome pairs based on comparisons of four COS genes for which all sequence types were present among the ESTs. The average divergence times for the colonial A_1/ colonial A_2, creeping A_2/creeping A_3, and the colonial A_2/ creeping A_2 subgenome pairs were 8.9, 10.6, and 2.2 MYA, respectively. Based on a cytological analysis, Jones (1956b) concluded that creeping and colonial bentgrass had the A_2 genome in common. The analysis presented here based on the EST sequence data supports his conclusion. The colonial and creeping bentgrass A_2 subgenome sequences are considerably more similar to each other than to the orthologous A_1 or A_3 sequences, respectively. The estimated divergence of the colonial and creeping bentgrass A_2 subgenomes from a common ancestor 2.2 MYA suggests this as the approximate time of the polyploidy events.

Recent analysis of homoeologous regions of the maize genome with orthologous regions of the sorghum and rice genomes showed that sorghum and maize separated from rice 50 MYA, whereas the progenitors of maize and sorghum split nearly instantaneously 11.9 MYA (Swigonova et al. 2004). Maize then arose from the hybridization of two progenitors as recently as 4.8 MYA. Since maize chromosomes expanded in size after this event in bursts (Du et al. 2006), it will be interesting to analyze the repeat elements of the bentgrass subgenomes as well. This would be particularly interesting in light of their possible differential roles in response to environmental stress. Colonial bentgrass and the A_1 diploid species velvet bentgrass both have good resistance to dollar spot whereas creeping bentgrass does not. Therefore, the genetic basis of the dollar spot resistance in colonial bentgrass possibly originates from the A_1 subgenome.

Colonial and creeping bentgrass provide examples of polyploid grass genomes that are intermediate in age between that of maize, which has become diploidized, and hexaploid wheat. The tetraploid ancestor of wheat (*T. turgidum*) is believed to have formed 0.5 MYA and its hybridization with the diploid species *Aegilops tauschii* producing hexaploid *T. aestivum* is believed to have been as recently as 8,000 years ago (Huang et al. 2002).

Although rice has a diploid genome, it underwent many segmental duplications, the most recent one 7.7 MYA (Rice Chromosomes 11 and 12 Sequencing Consortia 2005).

Table 7 Estimation of divergence times of the bentgrass subgenome pairs

Gene identification	Length, nt	K_s	K_a	Divergence time, MYA
Rieske Fe-S protein	303			
ColA1/ColA2		0.117 ± 0.044	0 ± 0	9.1
CrpA2/CrpA3		0.101 ± 0.040	0 ± 0	7.8
ColA2/CrpA2		0.028 ± 0.019	0 ± 0	2.2
Photosystem I, subunit N	399			
ColA1/ColA2		0.125 ± 0.041	0.014 ± 0.007	14.5
CrpA2/CrpA3		0.182 ± 0.053	0.01 ± 0.006	21.1
ColA2/CrpA2		0.036 ± 0.019	0.014 ± 0.007	4.2
Unknown protein-1	527			
ColA1/ColA2		0.09 ± 0.026	0 ± 0	8.9
CrpA2/CrpA3		0.081 ± 0.027	0 ± 0	8
ColA2/CrpA2		0.024 ± 0.013	0 ± 0	2.4
Unknown protein-2	396			
ColA1/ColA2		0.043 ± 0.022	0.007 ± 0.005	3.2
CrpA2/CrpA3		0.077 ± 0.030	0.01 ± 0.006	5.7
ColA2/CrpA2		0.000 ± 0.000	0 ± 0	0
Combined				Mean divergence time, MYA
ColA1/ColA2				8.9 ± 4.6
CrpA2/CrpA3				10.6 ± 7.0
ColA2/CrpA2				2.2 ± 1.7

The accession numbers of the sequences used in these analyses are given in Fig. 2

Given the recent polyploidization of colonial and creeping bentgrass, it will also be interesting to see whether any of the ancestral rice segmental duplications are also detectable in the four bentgrass homoeologous subgenomes.

The age of polyploidy in tetraploid cotton (*Gossypium*) was estimated to be about 1.5 MYA by comparing the divergence times of the tetraploid subgenomes with the corresponding diploid species (Senchina et al. 2003). For tetraploid colonial and creeping bentgrasses, the diploid origin of the shared A_2 genome is unknown. We therefore directly compared the divergence times of four colonial bentgrass A_2 and creeping bentgrass A_2 gene sequences to estimate the age of polyploidy. Velvet bentgrass was considered to be the diploid origin of the A_1 genome present in colonial bentgrass (Jones 1956b). Our efforts to use RT-PCR using primers designed from colonial bentgrass to obtain sequence data from velvet bentgrass for inclusion in the analysis were unsuccessful. As more sequence data becomes available from velvet bentgrass and if the diploid A_2 and A_3 *Agrostis* spp. can be identified, a more extensive analysis of the age of polyploidy of colonial and creeping bentgrass will be possible. Nevertheless, even sampling the gene content of genomes at roughly the 10% level provides important insights into the effects of polyploidization during speciation, and a solid foundation to build a resource for the studies of the evolution and gene pool of today's domesticated plant species.

References

Belanger FC, Plumley K, Day PR, Meyer WA (2003) Interspecific hybridization as a potential method for improvement of *Agrostis* species. Crop Sci 43:2172–2176

Belanger FC, Bonos S, Meyer WA (2004) Dollar spot resistant hybrids between creeping bentgrass and colonial bentgrass. Crop Sci 44:581–586

Bennetzen JL (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol 42:251–269

Bhattramakki D, Dolan M, Hanafey M, Wineland R, Vaske D, Register III JC, Tingey SV, Rafalski A (2002) Insertion-deletion polymorphisms in 3′ regions of maize genes occur frequently and can be used as highly informative genetic markers. Plant Mol Biol 48:539–547

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Bonos SA, Casler MD, Meyer WA (2003) Inheritance of dollar spot resistance in creeping bentgrass. Crop Sci 43:2189–2196

Bonos SA, Honig J, Kubik C (2005) The identification of quantitative trait loci for dollar spot resistance in creeping bentgrass using simple sequence repeats. In: The ASA-CSSA-SSSA international annual meeting, Salt Lake City

Brady KP, Rowe LB, Her H, Stevens TJ, Eppig J, Sussman DJ, Sikela J, Beier DR (1997) Genetic mapping of 262 loci derived from expressed sequences in a murine interspecific cross using single-strand conformational polymorphism analysis. Genome Res 7:1085–1093

Bruggmann R, Bharti AK, Gundlach H, Lai J, Young S, Pontaroli AC, Wei F, Haberer G, Fuks G, Du C, Raymond C, Estep MC, Liu R, Bennetzen JL, Chan AP, Rabinowicz PD, Quackenbush J, Barbazuk WB, Wing RA, Birren B, Nusbaum C, Rouonsley S, Mayer KF, Messing J (2006) Uneven chromosome contraction and expansion in the maize genome. Genome Res 16:1241–1251

Carbone I, Kohn LM (1993) Ribosomal DNA sequence divergence within internal transcribed spacer 1 of the Sclerotiniaceae. Mycologia 85:415–427

Casacuberta JM, Santiago N (2003) Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 311:1–11

Chakraborty N, Bae J, Warnke S, Chang T, Jung G (2005) Linkage map construction in allotetraploid creeping bentgrass (*Agrostis stolonifera* L.). Theor Appl Genet 111:795–803

Chakraborty N, Chang T, Casler MD, Jung G (2006) Response of bentgrass cultivars to *Sclerotinia homoeocarpa* isolates representing ten vegetative compatibility groups. Crop Sci 46:1237–1244

Devos KM, Beales J, Ogihara Y, Doust AN (2005) Comparative sequence analysis of the *Phytochrome C* gene and its upstream region in allohexaploid wheat reveals new data on the evolution of its three constituent genomes. Plant Mol Biol 58:625–641

Du C, Swigonova Z, Messing J (2006) Retrotranspositions in orthologouous regions of closely related grass species. BMC Evol Biol 6:62

Feschotte C, Jiang N, Wessler SR (2002) Plant transposable elements: where genetics meets genomics. Nat Rev Genet 3:329–341

Fulton TM, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14:1457–1467

Gale MD, Devos KM (1998) Comparative genetics in the grasses. Proc Natl Acad Sci USA 95:1971–1974

Gaut BS (2002) Evolutionary dynamics of grass genomes. New Phytol 154:15–28

Gaut BS, Morton BR, McCaig BC, Clegg MT (1996) Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. Proc Natl Acad Sci USA 93:10274–10279

Grandbastien M-A (1998) Activation of plant retrotransposons under stress conditions. Trends Plant Sci 3:181–187

Grass Phylogeny Working Group (2001) Phylogeny and subfamilial classification of the grasses (Poaceae). Ann Mo Bot Gard 88:373–457

Griffiths S, Sharp R, Foote TN, Bertin I, Wanous M, Reader S, Colas I, Moore G (2006) Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. Nature 439:749–752

Haberer G, Young SK, Bharti AK, Gundlach H, Raymond C, Fuks G, Butler E, Wing RA, Rounsley S, Nusbaum C, Birren B, Mayer KFX, Messing J (2005) Structure and architecture of the maize genome. Plant Physiol 139:1612–1624

Holst-Jensen A, Kohn LM, Schumacher T (1997) Nuclear rDNA phylogeny of the Sclerotiniaceae. Mycologia 89:885–899

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002) Genes encoding plastid acetyl-CoA carboxylae and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploidy wheat. Proc Natl Acad Sci USA 99:8133–8138

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–877

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800

Jones K (1956a) Species determination in *Agrostis*. Part I. Cytological relationships in *Agrostis canina* L. J Genet 54:370–376

Jones K. (1956b) Species differentiation in *Agrostis*. Part II. The significance of chromosome pairing in the tetraploid hybrids of *Agrostis canina* subsp. *montana* Hartmn., *A. tenuis* Sibth. and *A. stolonifera* L. J Genet 54:377–393

Jones K (1956c) Species determination in *Agrostis*. Part III. *Agrostis gigantea* Roth. and its hybrids with *A. tenuis* Sibth. and *A. stolonifera* L. J Genet 54:394–399

Kellogg EA (1998) Relationships of cereal crops and other grasses. Proc Natl Acad Sci USA 95:2005–2010

Kellogg EA (2001) Evolutionary history of the grasses. Plant Physiol 125:1198–1205

Kimura Y, Tosa Y, Betsuyaku S, Sasabe Y, Tomita R, Murakami J, Nakayashiki H, Mayama S (2002) Oat retrotransposon OARE-1 is activated in both compatible and incompatible interactions with pathogenic fungi. J Gen Plant Pathol 68:8–14

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163

McClintock B (1984) The significance of responses of the genome to challenge. Science 226:792–801

Messing J, Llaca V (1998) Importance of anchor genomes for any plant genome project. Proc Natl Acad Sci USA 95:2017–2020

Messing J, Bharti AK, Karlowski WM, Gundlach H, Kim HR, Yu Y, Wei F, Fuks G, Soderlund CA, Mayer KFX, Wing RA (2004) Sequence composition and genome organization of maize. Proc Natl Acad Sci USA 101:14349–14354

Meyer WA, Funk CR (1989) Progress and benefits to humanity from breeding cool-season grasses for turf. In: Sleeper DA, Asay KH, Pederson JF (eds) Contributions from breeding forage and turfgrasses. Crop Science Society of America Special Publication Number 15, Madison, WI, USA

Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K (2002) Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. Mol Genet Genomics 266:740–746

Plumley KA, Meyer WA, Murphy JA, Clarke BB, Bonos SA, Dickson WK, Clark JB, Smith DA (2000) Performance of bentgrass cultivars and selections in New Jersey turf trials. Rutgers Turfgrass Proc 32:1–21

Renvoize SA, Clayton WD (1992) Classification and evolution of the grasses. In: Chapman GP (ed) Grass evolution and domestication. Cambridge University Press, Cambridge, pp 3–37

Rice Chromosomes 11 and 12 Sequencing Consortia (2005) The sequence of rice chromosomes 11 and 12, rich in disease resistance genes and recent gene duplications. BMC Biol 3:20

Rotter D, Warnke S, Bonos SA, Meyer W, Belanger FD (2006) Colonial bentgrass genetic linkage mapping. In: The ASA-CSSA-SSSA international annual meeting, Indianapolis

Rudd S (2005) openSputnik—a database to ESTablish comparative plant genomics using unsaturated sequence collections. Nucleic Acids Res 33:D622–D627

Ruemmele BA (2003) *Agrostis capillaris* (*Agrostis tenuis* Sibth.) colonial bentgrass. In: Casler MD, Duncan RR (eds) Turfgrass biology, genetics, and breeding. Wiley, Hoboken, pp 187–200

Ruepp A, Zollner A, Maier D, Albermann K, Hani J, Mokrejs M, Tetko I, Guldener U, Mannhaupt G, Munsterkotter M, Mewes HW (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res 32:5539–5545

Senchina DS, Alvarez I, Cronn RC, Liu B, Rong J, Noyes RD, Paterson AH, Wing RA, Wilkins TA, Wendel JF (2003) Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. Mol Biol Evol 20:633–643

Song R, Llaca V, Messing J (2002) Mosaic organization of orthologous sequences in grass genomes. Genome Res 12:1549–1555

Soreng RJ, Davis JI (1998) Phylogenetics and character evolution in the grass family (Poaceae): simultaneous analysis of morphological and chloroplast DNA restriction site character sets. Bot Rev 64:1–85

Swigonova Z, Lai J, Ma J, Ramakrishna W, Llaca V, Bennetzen JL, Messing J (2004) Close split of sorghum and maize genome progenitors. Genome Res 14:1916–1923

Swofford L (2002) PAUP*: phylogenetic analysis using parsimony (*and other methods), Version 4. Sinauer Associates, Sunderland

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Tikhonov AP, SanMiguel PJ, Nakajima Y, Gorenstein NM, Bennetzen JL, Avramova Z (1999) Colinearity and its exceptions in orthologous adh regions of maize and sorghum. Proc Natl Acad Sci USA 96:7409–7414

Vicient CM, Jaaskelainen MJ, Kalendar R, Schulman AH (2001) Active retrotransposons are a common feature of grass genomes. Plant Physiol 125:1283–1292

Viji G, Uddin W, O'Neill NR, Mischke S, Saunders JA (2004) Genetic diversity of *Sclerotinia homoeocarpa* isolates from turfgrasses from various regions in North America. Plant Dis 88:1269–1276

Walsh B, Ikeda SS, Boland GJ (1999) Biology and management of dollar spot (*Sclerotinia homoeocarpa*); an important disease of turfgrass. HortScience 34:13–21

Warnke SE (2003) Creeping bentgrass (*Agrostis stolonifera* L.). In: Casler MD, Duncan RR (eds) Turfgrass biology, genetics, and breeding. Wiley, Hoboken, pp 175–185

Watson L (1990) The grass family, Poaceae. In: Chapman GP (ed) Reproductive versatility in the grasses. Cambridge University Press, Cambridge, pp 1–31

Wessler SR (1996) Plant retrotransposons: turned on by stress. Curr Biol 6:959–961

Xu JH, Messing J (2006) Maize haplotype with a helitron-amplified cytidine deaminase gene copy. BMC Genet 7:52

Young WC (2004) Extension estimates for Oregon forage and turf grass seed crop production, 2004. http://www.cropandsoil.oregonstate.edu/seed-ext/Agronomy/04ftprod.html

Zimmer R, Gibbins AMV (1997) Construction and characterization of a large-fragment chicken bacterial artificial chromosome library. Genomics 42:217–226