Origin and evolution of cultivated *Agrostis* spp.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Keenan L. Amundsen
Master of Science
Michigan State University, 2003

Director: Donald Seto, Associate Professor
Department of Bioinformatics and Computational Biology

Fall Semester 2009
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

I would like to thank my committee members Dr. Don Seto, Dr. Andrea Weeks, Dr. Huzefa Rangwala, and Dr. Scott Warnke for their support and guidance throughout the completion of this project. Thanks are also due to Dr. John Grefenstette, a former committee member, for assistance during the early stages of this study.

I would also like to thank the United States Golf Association and Dr. Scott Warnke of the Floral and Nursery Plants Research Unit, United States Department of Agriculture for providing financial support toward academic and research expenses. Additional thanks are extended to Ms. Vicki Bradley of the National Plant Germplasm System, Dr. Doug Johnson of the Forage and Range Research Laboratory, United States Department of Agriculture, Dr. Geunhwa Jung of the University of Massachusetts, Dr. Faith Belanger of Rutgers, the State University of New Jersey, and Mr. Kevin Morris executive director of the National Turfgrass Evaluation Program for providing research materials invaluable to the completion of this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

ORIGIN AND EVOLUTION OF CULTIVATED *AGROSTIS* SPP.

Keenan L. Amundsen, PhD.

George Mason University, 2009

Dissertation Director: Dr. Donald Seto

Creeping bentgrass (*Agrostis stolonifera* L.) is a highly outcrossing allotetraploid species important to the turfgrass industry because it has unique growth and aesthetic characteristics that make it ideally suited for use in high quality turf stands. There are between 150 and 200 species of *Agrostis* (bentgrasses) and the relationships between species are not clearly understood. Resolving these relationships is complicated in part because many of the species share similar morphological features and interspecific hybridization is common, making taxonomic classification difficult. Knowledge of the evolution of *Agrostis* species would provide turfgrass breeders with information on the origins of species and potential sources of novel germplasm.

In the present study 305 *Agrostis* accessions were examined to explore the evolutionary relationships within the *Agrostis* genus. Flow cytometry experiments were performed on each accession to measure DNA content and make ploidy predictions. MITE display molecular markers were designed and 1,309 were used to assess genetic

diversity within a subset of these *Agrostis* accessions. Fragments of the *trnL-trnF* and *atpI-atpH* intergenic spacer regions of the chloroplast genome were also sequenced and used to infer the phylogeny of the chloroplast genome. This data allowed for the prediction of candidate diploid progenitors of cultivated *Agrostis* species, identification of geographic regions of diverse germplasm, and demonstration of narrowing of the cultivated *Agrostis* gene pool. The data presented here gives *Agrostis* breeders valuable information to incorporate novel germplasm in their programs and a direction for recreating the polyploidization events that have led to the cultivated turf-type *Agrostis* species.

1. INTRODUCTION

Creeping bentgrass (*Agrostis stolonifera* L.) is the most widely utilized cool-season turf species for highly managed playing surfaces such as bowling greens and golf course greens, tees, and fairways (Turgeon, 1996). It outperforms other cool season turfgrass species because of its strong stoloniferous growth, ability to maintain a high level of uniformity after mowing, fine leaf texture, high shoot density, and tolerance of mowing down to heights of 3 mm (Warnke, 2003). *A. stolonifera* is well adapted to periodically flooded, well-drained, fine-textured, fertile soils (Beard, 1973). It is believed to have originated in Eurasia and is found throughout the world in cool, humid temperate climates (Harvey, 2007). While it is one of the most hardy of the cool season grasses used as turf, the quality of *A. stolonifera* is adversely affected by wear, soil compaction and a number of pathogens such as dollar spot (*Sclerotinia homoeocarpa* F.T. Benn.), brown patch (*Rhizoctonia solani* Kühn), and anthracnose (*Colletotrichum cereale* Manns).

There are 26 *Agrostis* (bentgrasses) species known within the North American region north of Mexico and 150 to 200 species worldwide (Harvey, 2007). The five species predominantly used as turf are *A. stolonifera*, *A. capillaris* L. (colonial bentgrass), *A. castellana* Boiss. and Reut. (dryland bentgrass), *A. canina* L. (velvet bentgrass), and *A. gigantea* Roth. (redtop bentgrass) (Warnke, 2003). Species delimitation based on

morphological characters is difficult because of similar features and intermediate types; this confusion has led to many distinct names for a single species (Vergara and Bughrara, 2003; Bonos et al., 2002). For example, creeping bentgrass has been cited as *A. palustris* Huds., *A. stolonifera* var. *palustris* (Huds.) Farw., and *A. stolonifera* L. Colonial bentgrass has been referred to as both *A. tenuis* Sibth. and *A. capillaris* L. Brown bentgrass has been named both *A. canina* subsp. *montana* Hartm. and *A. vinealis* Schreb. Another reason taxonomic classification of *Agrostis* species has been a challenge is due to the formation of hybrids and the presence of intermediate morphologies between taxa (Jones, 1955; Monteith, 1930; Shildrick, 1976). For example *A. stolonifera* has been shown to form interspecific hybrids with *A. canina*, *A. castellana*, *A. capillaris*, *A. vinealis*, *A. gigantea* (Belanger et al., 2003a; Jones, 1956bc), and *A. idahoensis* Nash and intraspecific hybrids with its close relatives *Polypogon monspeliensis* L. Desf., *Polypogon fugax* Nees., and *Polypogon viridis* Gouan Breistr. (Zhao et al., 2007).

A number of *Agrostis* species including those used as turf are polyploids. Polyploidy is common in plants and refers to a genome comprised of multiple chromosome sets or sub-genomes. The combining of unreduced gametes and somatic doubling are the primary mechanisms for polyploid formation (Hancock, 1992; Ramsey and Schemske, 1998). Polyploids are further subdivided into two main categories, allo- or autopolyploids. Ramsey and Schemske distinguish between the two categories by the way in which the species have evolved. They describe autopolyploids as having evolved from within populations of a single species and allopolyploids as having evolved from interspecific hybridization.

Jones (1956abc) directed the first studies to examine the genome constitution of the turfgrass species *A. canina*, *A. vinealis*, *A. capillaris*, *A. stolonifera* and *A. gigantea*. These studies were designed to explore the genome origins and relationships between species based on chromosome pairing behavior during metaphase I of meiosis; constituent genomes are believed to have evolved from a common ancestor if their chromosomes pair. To improve our understanding of species relationships, Jones (1956ab) investigated chromosome pairing and the frequency of chiasmata formation during metaphase I of meiosis of *Agrostis canina*, *A. vinealis*, *A. stolonifera*, *A. capillaris* and *A. gigantea*. The base haploid chromosome number of *Agrostis* is x = 7. *A. canina* (2n=2x=14) routinely formed seven bivalents and is a diploid. *A. vinealis* (2n=4x=28) regularly formed quadrivalents suggesting it is an autotetraploid. *A. stolonifera* (2n=4x=28) displayed bivalent pairing of chromosomes indicating it is a strict allotetraploid. *A. capillaris* (2n=4x=28) was found to be a segmental allotetraploid because of the formation of mostly bivalent chromosome pairing with reduced chiasma formation compared with *A. stolonifera*. The hexaploid *A. gigantea* (2n=6x=42) was found to be an allohexaploid based on bivalent chromosome pairing. To further understand the genomic relationships of the tetraploid species, Jones (1956bc) studied the chromosome pairing of hybrids formed between the species. Jones found evidence for trivalent pairing of chromosomes of a hybrid formed between the diploid *A. canina* and the autotetraploid *A. vinealis*. Jones concluded that the homologous sub-genomes of *A. vinealis* were likely derived from *A. canina*. *A. vinealis* is rhizomatous whereas *A. canina* is not, suggesting *A. vinealis* may have evolved in part from an ecotypically

distinct rhizomatous *A. canina*. The hybrids formed from a cross between *A. vinealis* and *A. capillaris* predominantly formed bivalents and quadrivalents suggesting that these species have one diploid genome ancestor in common. A cross between *A. vinealis* and *A. stolonifera* formed mostly univalents, which showed that the chromosome sets of the two species do not associate with one another. The *A. stolonifera* X *A. capillaris* hybrids had a mix of univalents and bivalents; this data is supportive of the parent species sharing one chromosome set from a common diploid ancestor. Hybrids of *A. gigantea* crossed with *A. capillaris* or *A. stolonifera* exhibited on average seven univalents and 14 bivalents. Jones suggested that the diploid or tetraploid ancestors of *A. capillaris* and *A. stolonifera* may have contributed to the evolution of *A. gigantea*. Jones concluded that the genomic constitution of *A. capillaris*, *A. stolonifera*, and *A. gigantea* could be described as $A_1A_1A_2A_2$, $A_2A_2A_3A_3$, and $A_1A_1A_2A_2A_3A_3$ respectively. It follows that the *A. canina* and *A. vinealis* genome designations would be A_1A_1 and $A_1A_1A_1A_1$ respectively (Brilman, 2001). Jones warned that this naming convention was used to describe the chromosome set interactions of the hybrids, but it should not imply that the A_2 genome is identical between the species because the *A. stolonifera* X *A. capillaris* and *A. stolonifera* X *A. gigantea* hybrids had increased sterility and lower seed set.

Zhao et al. (2007) conducted similar cytology experiments on *Agrostis* hybrids and found chromosome associations between *A. stolonifera* and *A. capillaris* and *A. gigantea*. Zhao et al. also found chromosome associations in hybrids of *A. stolonifera* crossed with *A. idahoensis* (2n=4x=28), *Polypogon monspeliensis* (2n=4x=28), *P. fugax* (2n=6x=42), and *P. viridis* (2n=4x=28). Zhao et al. found on average 9.91 univalents and

5.18 bivalents in the *A. stolonifera* X *A. idahoensis* hybrid, 2.67 univalents and 10.00 bivalents in the *A. stolonifera* X *P. monspeliensis*, 9.33 univalents and 6.67 bivalents in the *A. stolonifera* X *P. fugax* hybrid, and 5.00 univalents and 7.66 bivalents in the *A. stolonifera* X *P. viridis* hybrid. These results supported Jones' findings and suggested that the parental species of each hybrid share either one homologous set of chromosomes or homologous chromosome segments throughout the genome. The observations by Jones on chromosome associations during metaphase I of meiosis are still cited as the most comprehensive investigations into the subject (Warnke, 2003; Belanger et al., 2003a; Zhao et al., 2007).

There is increased interest, as outlined below, in revisiting the early cytology experiments of Jones to distinguish *Agrostis* species with modern molecular marker techniques. Part of the reason for this resurgence is because an understanding of how *Agrostis* species relate to one another will improve our ability to predict successful hybrid formation between species. Generating hybrids from wide crosses to incorporate wild germplasm into cultivated material has been a successful method for introducing novel genes for plant improvement in breeding programs (Stalker, 1980; Kalloo, 1992). Even though many of the species used as turf have evolved from naturally occurring interspecific hybrids (Casler and Duncan, 2003), Belanger et al. (2003b) was the first to investigate interspecific hybridization as a means of improving *Agrostis* species. Belanger et al. demonstrated that successful hybrids were formed between *A. stolonifera* and *A. canina*, *A. capillaris*, *A. gigantea*, and *A. castellana*. In 2004, Belanger et al. generated hybrids between *A. stolonifera* and *A. capillaris* to test if interspecific

hybridization could be used to improve dollar spot resistance in *A. stolonifera*. Dollar spot is an economically important disease to the turfgrass industry and *A. stolonifera* is considered susceptible to the disease while *A. capillaris* is tolerant. Hybrid plants were recovered with significantly more dollar spot resistance than the *A. stolonifera* parent suggesting that interspecific hybridization could be an effective tool to improve *Agrostis* species.

Correctly identifying *Agrostis* specimens at the species level is critical before we can begin to understand the evolutionary history of the genus. Morphological characters can be used to distinguish certain species of *Agrostis*. For example *A. capillaris* can be distinguished from *A. gigantea* by ligule length, rhizomatous growth habit, and inflorescence (Harvey, 2007). The presence of intermediate morphological types and hybrids limit the ability to classify *Agrostis* plants solely on this basis. For example *A. capillaris* and *A. castellana* are difficult to distinguish morphologically because they are both rhizomatous and have similar morphological characters commonly used to distinguish species such as leaf width, sheath type, and ligule shape (Zhao et al., 2006; Harvey, 2007). A first step in differentiating species of *Agrostis* is to determine ploidy. Bonos et al. (2002) describes the difficulties of characterizing highland bentgrass, which is often confused with *A. capillaris* (2n=4x=28) but might be a distinct subspecies of *A. castellana* (2n=6x=42). Performing flow cytometry experiments they measured the mean 2C DNA content (pg) of *A. canina*, *A. stolonifera*, *A. capillaris*, *A. vinealis*, *A. gigantea*, and *A. castellana* at 3.42, 5.27, 5.87, 6.31, 8.18, and 8.71 respectively. These DNA content measures obtained by flow cytometry are consistent with those reported by other

studies that included *Agrostis* species (Hollman et al., 2005; Arumuganathan et al., 1999; Arumuganathan and Earle, 1991). Additionally Bonos et al. counted the chromosomes from metaphase cells of *A. canina* (2n=2x=14), *A. capillaris* (2n=4x=28), *A. stolonifera* (2n=4x=28), *A. castellana* (2n=6x=42), and *A. alba* (2n=6x=42). A strong correlation between 2C DNA content and ploidy level was observed. These measures of DNA content could aid in the classification of *Agrostis* species, such as highland bentgrass, that have morphological characteristics borderline between two species of differing ploidy. Ploidy determination and DNA content are important to *Agrostis* breeders because the more similar these values are between parent plants, the more likely that successful hybridization will be achieved (Bonos et al., 2002; Hollman et al., 2005). DNA content measured by flow cytometry has been successful in predicting the ploidy of a number of other cultivated species including *Vicia villosa* Roth (Yeater et al., 2004), *Oxalis tuberosa* L. (Emshwiller, 2002), *Medicago* spp. (Brummer et al., 1999), and *Bromus* spp. (Tuna et al., 2001). Tuna et al. found variation in 2C DNA content between *Bromus* species from the same ploidy level and observed that the octaploids had 1.2 pg less than twice the DNA content of the tetraploids. This data is evidence of reduced genome size with increased ploidy in *Bromus* species and suggests care must be taken when using DNA content measures as a prediction of ploidy. In cases where DNA content does not correlate with ploidy level, chromosome counts could be done to determine ploidy (Jones, 1956abc; Bonos et al., 2002; Zhao et al., 2006; Ahloowalia, 1965). The advantage of predicting ploidy by flow cytometry over chromosome counts is the speed at which the DNA content can be measured.

By evaluating *Bromus* germplasm in the National Plant Germplasm System (NPGS) and recognizing that the majority of the available germplasm was collected from just a few locations and that a number of species were underrepresented in the system, Tuna et al. made a case for increasing the numbers and genetic diversity of *Bromus* in the repository. Certain species of *Agrostis* are underrepresented in available NPGS germplasm (Appendix A) and the ploidy level of a number of these has not yet been determined.

Molecular markers have been used extensively in studies of species delimitation and population genetic diversity in grasses. Molecular markers are molecules such as DNA, RNA, or proteins that because of their size or other properties can be used to distinguish different genotypes. For example, Warnke et al. (1997) examined isozyme markers to assess the relationships between 18 *Agrostis stolonifera* lines including both cultivated and non-cultivated accessions. Polymorphic allozymes were scored for presence or absence as one or zero respectively and the frequency of each marker within the population was calculated. Nei's distance formula (Nei, 1972) was used to calculate the genetic distances between and within the *Agrostis* plants. The genetic distances were used to cluster the accessions by the unweighted pair group method with arithmetic averages (UPGMA). A plant from Europe (PI251945) was found to be significantly different from the United States cultivated material. Warnke et al. suggests European material might be a source of novel germplasm to improve U.S. cultivated material. Zhao et al. (2006) investigated the genetic diversity among 39 *A. capillaris* accessions by AFLP (Amplified Fragment Length Polymorphism; Vos et al., 1995). They compared

genetic similarity or common AFLP markers between each pair of accessions and found a considerable amount of genetic variation. Vergara and Bughrara (2003) used AFLP to evaluate genetic diversity of 40 *Agrostis* accessions representing 14 different species and in 2004 explored the genetic variation of *A. stolonifera* and *A. gigantea* accessions. Vergara and Bughrara were able to identify distinct groups of markers among the *Agrostis* accessions to discriminate species and found that most of the observed marker variation was within the *A. gigantea* accessions. These AFLP experiments included European material that was found to be genetically different from U.S. cultivated *A. stolonifera*. This is consistent with the findings of Warnke et al. (1997) and suggested that European germplasm might be a source of novel genes for U.S. *Agrostis* breeding programs. Similarly, Karlsen and Steiner (2007) used RAPD (Random Amplified Polymorphic DNA; Williams et al., 1990) markers to describe the genetic diversity of *A. capillaris* plants in the Scandinavian region and found that collected material from the northern Scandinavian region was distinct from accessions present in the NPGS. Hollman et al. (2005) performed RAPD analysis to characterize *Agrostis* germplasm and found markers that were able to differentiate *A. canina*, *A. capillaris* and *A. stolonifera* germplasm. The DNA content and predicted ploidy level of each accession was validated by flow cytometry and supported the characterization of plants at the species level by RAPD markers. Caceres et al. (2000) studied RFLP (Restriction Fragment Length Polymorphism; Girardin et al., 1993) markers and found seven markers that could differentiate five *A. stolonifera* accessions. Sequence characterized amplified region (SCAR; Paran and Michelmore, 1993) markers designed from RAPD fragments by

Scheef et al. (2003) were found to co-segregate with either the A_1A_1 or A_3A_3 genomes of *A. stolonifera* and *A. capillaris* respectively and could therefore be used as sub-genome specific markers to differentiate the species. These studies further our understanding of the relationships among *Agrostis* species and demonstrate that molecular marker techniques can be used to distinguish closely related *Agrostis* species.

Miniature inverted-repeat transposable elements (MITE; Feschotte et al., 2002) have recently been used to assess the genetic diversity of *Zea mays* L. populations (Kavar et al., 2007) and *Oryza* species (Park et al., 2003). MITEs are a class of transposable elements that are characterized by terminal inverted repeats (TIR; 10-15 bp), target site duplications (TSD; 2-3 bp), high copy number, and small size (Casa et al., 2004). MITE-display is a relatively new marker system that employs a modified AFLP protocol to amplify DNA fragments anchored to MITE sequences (Wessler et al., 2001). A typical AFLP reaction involves digesting genomic DNA with two separate endonuclease restriction enzymes, ligating adapters to the cut DNA, performing a pre-selective PCR (polymerase chain reaction) amplification with primers complimentary to the adapters plus one additional nucleotide, and performing a selective PCR reaction with the adapter complimentary primers plus three selective nucleotides. MITE-display differs from AFLP by substitution of one pre-selective and one selective PCR primer for ones that are internal to the MITE sequence. MITE-display targets gene rich regions of the genome because MITEs have a tendency to insert near genes (Casa et al., 2004). Kavar et al. (2007) used MITE-display to examine the genetic diversity within and among 15 landrace populations of *Z. mays*. Selective PCR was performed separately with one

primer designed internal to the *Hbr* (Heartbreaker) class of MITEs (Casa et al., 2000) together with one of three MseI+1 primers (MseI + C, MseI + G, or MseI + T). There were 263 polymorphic markers scored for the set of 150 plants. Kavar et al. observed a low level of marker diversity within populations, a significant amount of diversity between populations, and identified 39 population-specific MITE anchored markers. Park et al. (2003) studied genetic variation of 53 *Oryza* accessions representing 13 different species. The 13 species were comprised of six different genome constitution types including diploid species with genome designations AA, BB, CC, and EE and tetraploids with the BBCC and CCDD designations. The MITE-display markers clearly grouped accessions of the same species and genome types together while separating different species.

Tu (2001) developed a software program (FindMITE) to search DNA sequence data for user defined MITE characteristics including TSD, TIR, and size. Tu identified 8 novel MITE sequences from *Anopheles gambiae* s.s. Giles by searching 17,509 STS (sequence tagged sites; Olson et al., 1989) sequences with FindMITE. Tu verified that these MITEs were once mobilized by analyzing the genomic sequence flanking the MITEs. The genomic DNA varied between sequences for a given class of MITE suggesting it was present in different genomic regions, likely through transposition. Since 2001 FindMITE has been successfully used to identify putative MITE sequences from *Gossypium hirsutum* L. (Grover et al., 2008), *Poncirus trivoliata* L. Raf. (Yang et al., 2003), and *Lotus japonicus* L. (Holligan et al., 2006). Rotter et al. (2007) sequenced DNA libraries of the two *Agrostis* tetraploid species *A. stolonifera* and *A. capillaris*

covering approximately 10% of the predicted gene space. A total of 2,574 *A. capillaris* and 1,329 *A. stolonifera* EST sequences were characterized as unigenes and placed into functional classes. Of these *Agrostis* unigenes, 10 *A. capillaris* and 57 *A. stolonifera* sequences were categorized as transposable elements and none of the sequences were identified as MITEs. MITEs are generally classified based on their structural similarity and TIR and TSD sequence identity to known autonomous classes of transposable elements because the internal MITE sequences are not highly conserved (Feschotte et al., 2002). Special parameters are needed when performing BLAST (Basic Local Alignment Search Tool; Altschul et al., 1990) searches to identify MITE sequences (Yang and Hall, 2003a), so it is not surprising that none were found by the initial sequence analysis performed by Rotter et al. FindMITE could be used to identify new classes of MITEs from this *Agrostis* sequence data.

The *Agrostis* sequencing efforts of Rotter et al. (2007) produced 7528 and 8470 EST (expressed sequence tag) sequences of *A. capillaris* and *A. stolonifera* respectively and make up the majority of the *Agrostis* sequence data available through the National Center for Biotechnology Information (NCBI) databases (16,992 *Agrostis* EST sequences as of Mar. 7th, 2008). Rotter et al. identified 177 and 161 *Oryza* COS (Conserved Ortholog Set; Fulton et al., 2002) genes conserved in the *A. capillaris* and *A. stolonifera* DNA sequence databases respectively. Fulton et al. described COS markers as a set of single or low copy genes that are conserved across species and can be used for comparative purposes. Seven of the 161 *A. stolonifera*, and 11 of the 177 *A. capillaris* *Oryza* COS gene homologues were found to have conserved sequences in other grasses.

These sequences were used in both a neighbor-joining and maximum parsimony phylogenetic analysis with maize as an outgroup. Both phylogenetic analysis methods placed the *Agrostis* species closer to *F. arundinacea* than to *A. sativa*, which is interesting since *Agrostis* is in the same tribe as *A. sativa* while *F. arundinacea* is in the Poeae tribe. *Agrostis* was placed in the Aveneae tribe based on morphological characteristics and enzyme restriction patterns of chloroplast DNA (Watson, 1990; Soreng and Davis, 1998). The two tribes are closely related (Kellog, 1998; Soreng et al., 2007) and share many common chloroplast DNA restriction sites (Soreng et al., 1990).

Another interesting analysis to come from the *Agrostis* sequence data is a predicted time of divergence for the species (Rotter et al., 2007). A number of the *Agrostis* COS genes were identified as having multiple distinct copies in the EST sequence data, which was expected since *A. capillaris* and *A. stolonifera* are tetraploids. Four of the multiple copy COS sequences were found in both the *A. stolonifera* and *A. capillaris* sequence databases. Following the clustering methods of Xu and Messing (2006) for each of the multi-copy COS gene sets, Rotter et al. predicted which of the two copies of *A. stolonifera* was more closely related to which corresponding COS gene copy in *A. capillaris*. The gene pairs sharing the highest degree of similarity between *A. stolonifera* and *A. capillaris* were considered to be from the sub-genome (A_2A_2) in common between the two species. The COS gene pairs were further analyzed by comparing each sequence to the corresponding *Oryza* ortholog and calculating the number of substitutions at synonymous sites (K_s) and non-synonymous sites (K_a). It was determined that these four COS genes were undergoing purifying selection due to the low

K_a/K_s ratio. Purifying selection is when the amount of genetic diversity in a population is reduced over time. The rate of divergence of the sub-genomes for both *A. capillaris* and *A. stolonifera* were estimated by the formula $T = K_s/2r$ where T is the time of divergence and r is the rate of divergence of synonymous sites as described by Senchina et al. (2003). By calibrating to *Oryza* and comparing the rate of divergence of the *Agrostis* sub-genomes, Rotter et al. found the time of divergence for *A. capillaris* A_1A_1 and A_2A_2 sub-genomes, *A. stolonifera* A_2A_2 and A_3A_3 sub-genomes, and *A. capillaris* and *A. stolonifera* A_2A_2 sub-genomes to be 8.9, 10.6, and 2.2 MYA (Million Years Ago) respectively. The age of the A_2A_2 sub-genomes (2.2 MYA) suggests a recent origin of the tetraploid *Agrostis*, well after the diploid progenitors diverged.

Phylogenetic analysis of nuclear DNA sequences is challenging in polyploids because the genome duplication events leading to their formation often results in homoeologous loci, which confound the analysis. Warnke (2003) suggests the evolution of *Agrostis* could be further resolved by performing a phylogenetic analysis of plastid DNA based sequences. The lineage of *Agrostis* is complicated particularly since many of the species are highly outcrossing polyploids that frequently form interspecific hybrids. A phylogenetic analysis of *Agrostis* chloroplast markers could simplify the analysis for maternal relationships since chloroplast DNA is maternally inherited in flowering plants. There are a number of chloroplast DNA based markers frequently used for this type of analysis. Soreng, et al. (2007) performed a phylogenetic analysis of the Poeae tribe *sensu lato* using morphological characteristics and sequence data of three chloroplast-encoded genes, maturase K (*matK*), NADH dehydrogenase subunit F (*ndhF*), and ATP synthase

beta subunit (*atpβ*). The sequence data for this project is available through the NCBI databases and the analysis included the species *A. tenerrima* Trin., *Calamagrostis arundinacea* (L.) Roth., *C. canadensis* (Michx.) P. Beauv., *Polypogon monspeliensis*, *Festuca rubra* L., and *F. subverticillata* (Pers.) E.B. Alexeev. Soreng et al. examined 18 morphological characteristics and sequence variation among the three chloroplast markers. It was found that the molecular character data was better than the morphological data at describing a well-resolved set of relationships between species. There were 754 (14.3%) parsimony informative characters from a total of 5259 sites among the chloroplast sequences examined. Of the two major clades (Aveneae and Poeae types), *Agrostis* is within the Aveneae group and is more closely related to *Polypogon* spp. than to *C. arundinacea* or *C. canadensis* and is more distantly related to *Avena* spp. Similarly Döring et al. (2007) sequenced the *matK* gene from a number of specimens from the Aveneae/Poeae complex including *A. capillaris*, *C. rivalis* (Torges) H. Scholz., *C. macrolepis* Litv., *Apera spica-venti* (L.) P. Beauv., *P. monspeliensis*, *Avena sativa* L., *F. altissima* All., and *F. gigantean* (L.) Vill. *Agrostis* can hybridize with *Polypogon* spp. and shares morphological characteristics in common with *Apera* and *Calamagrostis* spp (Harvey, 2007). The results of Döring et al. (2007) were consistent with Soreng et al. (2007) and it is interesting that both phylogenies of chloroplast encoded gene sequence data placed *Agrostis* more closely to *Avena* than *Festuca* in contrast to the COS phylogenetic analysis performed by Rotter et al. (2007).

Reichman et al. (2006) examined the flow of pollen from transgenic *Agrostis stolonifera* to wild grown *Agrostis* species. As part of their study, they conducted one of

the more detailed phylogenetic studies of *Agrostis* based on DNA sequence data. The transgenic *A. stolonifera* plants contained a glyphosate-resistance selectable marker allowing for easy selection of hybrid plants. DNA was extracted from wild *A. gigantea*, *A. stolonifera* and *A. exarata* plants that were glyphosate resistant and both the nuclear ribosomal ITS (Internal Transcribed Spacers) and chloroplast encoded *matK* DNA were sequenced from the hybrids and compared to sequences from 13 *Agrostis* species and three *Polypogon* species that were used as an outgroup. The sequence data was used in a phylogenetic analysis to determine the parentage of the hybrids. For both the ITS and *matK* phylogenetic trees the transgenic plants recovered in the 2004-2005 study fell within the *A. stolonifera* clade providing evidence that the transgenic *A. stolonifera* plants were the pollen donors of the glyphosate resistant wild grown plants. The main point of this study was to demonstrate that transgenic *Agrostis* could pollinate wild relatives grown nearby. The phylogenetic trees built separately from the *matK* and ITS sequence data are interesting because the two phylogenetic trees presented incongruent information and trees of different topologies. For example *P. viridis* USHN11 is placed near *A. exarata* USU43 and *A. exarata* NGC based on the ITS sequence data, but the *matK* data suggests that *P. viridis* USHN11 is not closely related to the *A. exarata* accessions. Similarly the *matK* data suggests that *A. idahoensis* USU13 is closely related to the *A. exarata* accessions, while the ITS data place *A. idahoensis* sister to *A. scabra* OSU40 and A. *vinealis* GRIN3.

Different homologous DNA segments from related taxa often have inconsistent evolutionary histories and analyzing the DNA segments separately will often yield trees

of differing topologies (McBreen and Lockhart, 2006). One reason these inconsistencies arise is from reticulation events such as horizontal gene transfer, recombination, and hybridization. Reticulation events are not well represented by bifurcating phylogenetic trees (Linder and Rieseberg, 2004; Huson et al., 2005; Huson and Bryant, 2006). A phylogenetic network is a useful tool to examine the evolution of complex lineages and is often used to determine when hybridization or horizontal gene transfer has played a role in the evolution of the species under investigation. A number of *Agrostis* species are polyploids and thought to have evolved from naturally occurring interspecific hybrids (Casler and Duncan, 2003). A phylogenetic network might be the best representation of the evolution of *Agrostis*. Huson and Bryant (2006) defined a phylogenetic network as a network composed of lines and nodes where lines represent evolutionary relationships and nodes represent taxa. Phylogenetic trees, split networks, and reticulate networks are examples of phylogenetic networks (Huson and Bryant, 2006). Phylogenetic trees are often depicted as rooted or unrooted bifurcating graphs that represent the evolutionary history of a species. Split networks are a means of showing inconsistencies between sets of data. Reticulate networks are effective at explaining evolutionary history of taxa when reticulation events have occurred, such as the events that are thought to have lead to the evolution of *Agrostis* species. Huson and Bryant (2006) developed the software progam SplitsTree4 that can perform a number of phylogenetic analyses including building reticulate networks from independent gene trees. The first step in building a reticulate network in SplitsTree4 is to build or import separate phylogenetic trees of genes from independent loci across the species under study; the individual gene trees can be

constructed by standard methods such as maximum likelihood or maximum parsimony. SplitsTree4 generates a splits network to explain the incongruent nodes between individual gene trees. Finally the program tries to replace the nodes explaining incongruent points by reticulation events such as hybridization. Hybridization networks are reticulate networks used to explain hybridization events and typically include an outgroup in the analysis, whereas split networks generally do not include an outgroup.

An outgroup is a sequence or group of sequences that are homologous to the gene under investigation, similar enough to provide resolution of the differences, and distinct enough to be separated from the genes of interest and root the tree (Huson and Kloepper, 2005). *Agrostis* is a member of the Aveneae tribe, of which *Avena sativa* L. (oat) is the most recognized agronomic species. *Agrostis* is closely related to *Polypogon* species and has morphological characters in common with *Calamagrostis* and *Apera* species (Harvey, 2007). Reichman et al. (2006) used *Polypogon* species as an outgroup in their analysis of *matK* and ITS sequences, and if there are enough nucleotide changes between *Polypogon* and *Agrostis* it might be a good outgroup. However, *Calamagrostis* has been hypothesized (Kuoh, 2003) and *Polypogon* is known (Zhao et al., 2007; Barkworth, 2007) to form hybrids with *Agrostis* species and therefore might not be distantly related enough from *Agrostis* to be effective outgroups. Rotter et al. (2007) found *Agrostis* to be more closely related to *Festuca arundinacea* than to *Avena sativa* based on nuclear COS genes, while an analysis of chloroplast encoded genes by Döring et al. (2007) and Soreng et al. (2007) showed *Agrostis* to be more closely related to *Avena sativa*. Therefore

Avena, *Festuca* or another species similar to *Agrostis* such as *Apera* would be good candidates to meet the outgroup criteria in a phylogenetic analysis of *Agrostis*.

Molecular markers targeted to chloroplast, nuclear ribosomal, and single or low-copy genes are frequently used for reconstructing evolutionary histories (Raymond et al., 2002). Chloroplast markers such as *matk* are maternally inherited and are effective for understanding maternal inheritance but alone cannot describe the evolutionary history of reticulate events (Feliner et al., 2007) because they do not explain the paternal lineage. The ITS region is a high copy nuclear ribosomal sequence that is frequently used for phylogenetic analysis but is not effective for assessing phylogenies of outcrossing polyploids. These sequences (ITS) are in high copy number and there is a tendency for them to become homogenous by the mechanisms of concerted evolution. This process is problematic for phylogenetic analyses in polyploids because it leads to the loss of sub-genome specific ITS sequence types. Additionally since allotetraploids for example have two independent sub-genomes, the rate of homogenization between sub-genomes varies between lineages and in turn makes ITS unreliable for phylogenetic analysis. Feliner et al. (2007) and Raymond et al. (2002) suggested that single or low-copy nuclear genes might be a better choice for phylogenetic analysis of complex lineages. The difficulty associated with isolating single or low-copy genes conserved across species was pointed out as one of the main disadvantages of using these sequences in the analysis. Following the methods of Fulton et al. (2002), Rotter et al. (2007) were able to identify conserved single or low copy orthologous genes from *Agrostis* species and use them in a phylogenetic analysis.

Sang et al. (1995, 1997ab) and Sang and Zhang (1999) conducted one of the more thorough investigations of reticulate evolution in an allotetraploid. Sang et al. (1995) examined the phylogeny of 27 different *Paeonia* species by ITS. There was evidence of nucleotide additivity among 15 of the species, so the remaining 12 were used for parsimony analysis to build a phylogenetic tree. The 15 species not included in the initial analysis were believed to be of hybrid origin and were added to the tree by adding branches between them and the species for which nucleotide additivity was observed. In 1997(a), Sang et al. sequenced *matK* and two intergenic spacers, *psbA-trnH* and *trnL-trnF* from 32 *Paeonia* species. A comparison of the ITS phylogenetic analysis (Sang et al., 1995) with that of the chloroplast sequences revealed many discordances. The reticulate phylogeny produced from the combination of these analyses suggested more hybridization events than were predicted from the ITS analysis alone. Sang and Zhang (1999) sequenced the two nuclear encoded *Adh1* and *Adh2* genes from five *Paeonia* species. They cloned the genes and sequenced one to three distinct copies from each *Paeonia* species. The distinct *Adh1* or *Adh2* sequences from each species were found to be divergent (i.e. *Adh1* Type I, Type II, or Type III) and were used in parsimony analysis. The phylogenetic analysis of distinct *AdhI* and *AdhII* alleles from each species along with the ITS and chloroplast DNA allowed for interpretation of the hybrid origin of *Paeonia* species. Similarly, incongruent phylogenies were derived from nuclear and chloroplast DNA sequences in the evolutionary history studies of *Glycine* species, which suggested hybridization events have occurred in the evolution of the species (Doyle et al., 2003). Comparisons of single or low-copy nuclear DNA can also provide evidence towards the

origins of allopolyploid taxa (Sang, 2002). Ge et al. (1999), sequenced *Adh1*, *Adh2*, and *matK* from 23 *Oryza* species. Nine different genome types were present among the *Oryza* accessions. The *mat*K chloroplast encoded gene sequence was used to explain the maternal parentage of the species and the allotetraploid origins were inferred from the *Adh1* and *Adh2* gene sequences by comparing sequence polymorphisms of the *Adh1* and *Adh2* genes of the tetraploids with the diploids. From these studies it is clear that because of the unpredictable nature of allotetraploid evolution, where gene duplication, loss and silencing are common, it is important to analyze multiple gene sequences (Raymond et al., 2002).

Since the 1956 cytological experiments of Jones, there have been a number of advancements for investigating relationships between *Agrostis* species, particularly with the use of molecular markers, sequence data, and phylogenetic analyses. It is suggested that *A. stolonifera* evolved from the hybridization of two unknown diploid species (Jones, 1956b), although hybridization between ancient allotetraploids is another possibility. The study of nuclear DNA sequences in conjunction with chloroplast DNA sequences would improve our understanding of the *Agrostis* phylogeny, as was the case for *Paeonia*, *Glycine*, and *Oryza*. An *Agrostis* hybridization network could be built from nuclear DNA sequences in combination with chloroplast DNA sequences. The evolutionary history of *Agrostis* is still unresolved but if better understood would aid *Agrostis* breeding programs (Warnke, 2003) by providing information on parentage of the species, the formation of polyploids and interspecies relationships.

The present study is designed to explore species relationships and assess genetic diversity within a set of publicly available *Agrostis* germplasm. In chapter 2 the DNA content was measured and used to predict the ploidy level of *Agrostis* germplasm available through the National Plant Germplasm System (NPGS). Accessions of *Agrostis* that are genetically distinct from modern cultivated species may be a source of unique stress tolerance genes. Molecular markers based on Miniature Inverted-repeat Transposable Elements (MITEs) were developed in chapter 3 to assess the genetic diversity within a set of NPGS *Agrostis* germplasm. The diversity analysis data is presented in chapter 4. *Agrostis* species relationships were explored in chapter 5 by performing phylogenetic analyses of two independent chloroplast DNA sequence regions. Chapter 6 is a conclusion of the project and summarizes the important findings from this study.

2. PLOIDY ANALYSIS OF NPGS *AGROSTIS* ACCESSIONS

Species within the *Agrostis* genus (the bentgrasses) represent a ploidy series with a base haploid chromosome number of x=7. Creeping bentgrass (*A. stolonifera* L.), colonial bentgrass (*A. capillaris* L.), dryland bentgrass (*A. castellana* Boiss. and Reut.), velvet bentgrass (*A. canina* L.), and redtop bentgrass (*A. gigantea* Roth.) are the dominant turf-type species (Warnke, 2003; Harvey, 2007). The turf-type *Agrostis* species range in ploidy from diploid, 2n=2x=14, to hexaploid, 2n=6x=48. *A. stolonifera*, *A. capillaris*, and *A. canina* are well adapted for use in high quality intensively managed turf environments such as golf course putting greens, tees, and fairways (Turgeon, 1996). The bentgrasses are considered hardy, but are susceptible to a number of biotic and abiotic stresses. Traditional plant breeding techniques are the primary method for improving stress tolerance in *Agrostis* (Bonos et al., 2006; Phillips, 2007; Warnke, 2003).

The National Plant Germplasm System (NPGS) houses the largest collection of publicly available *Agrostis* germplasm within the United States. The collection has more than 300 accessions representing more than 25 distinct species, accounting for 12-16% of the known 150 to 200 species of *Agrostis* (Harvey, 2007). The commonly used turf-type *Agrostis* species are diploid (*A. canina*) or tetraploid (*A. stolonifera* and *A. capillaris*). Other compatible diploid and tetraploid germplasm could be used in wide-cross breeding to improve cultivated *Agrostis*.

Interspecific hybridization is known to occur in natural populations of *Agrostis* (Belanger et al., 2003b) and intergeneric crosses also occur between *Agrostis* and *Polypogon* species, a close relative of *Agrostis* (Zhao et al., 2007; Barkworth, 2007). Belanger et al. (2003b) proposed using interspecific hybridization as a means of improving *A. stolonifera*. Belanger et al. (2003a) successfully generated hybrids between *A. stolonifera* and *A. canina*, *A. capillaris*, *A. gigantea*, and *A. castellana*. The most important fungal disease of *Agrostis stolonifera* is dollar spot caused by the fungal pathogen *Sclerotinia homeocarpa* F.T. Bennett (Vargas, 1994). Dollar spot can cause significant damage to *Agrostis stolonifera* during the warm humid summer months and more money spent in the United States on managing this disease than on all of the other turf diseases. *A. capillaris* is more tolerant of dollar spot than is *A. stolonifera* (Belanger et al., 2004). Belanger et al. generated hybrids between *A. stolonifera* and *A. capillaris* and assessed dollar spot sensitivity within the progeny. Hybrid progeny were identified that had significantly higher dollar spot tolerance than the *A. stolonifera* parent. Belanger et al. demonstrated that dollar spot resistance could be introduced into *A. stolonifera* through interspecific hybridization. Wide cross breeding has been used to improve other turfgrass species such as *Poa* spp. (Bashaw and Funk, 1987; Goldman and Sims, 2005; Van Dijk and Winkelhorst, 1981) and the *Lolium*/*Festuca* complex (Thomas and Humphreys, 1991; Kopecky et al., 2005).

Hybridization of parental plants of differing ploidy levels often results in the formation of sterile aneuploid progeny, therefore ploidy determinations of parental material is needed. Chromosome counts are the most accurate method for predicting

ploidy, but the process is labor intensive and time consuming. There is a correlation between ploidy level and nuclear DNA content that can be measured by flow cytometry. Flow cytometry is faster than chromosome counting and can be used to efficiently screen large sets of germplasm to predict ploidy levels. Flow cytometry has been used to determine the ploidy level of various plants, such as *Vicia villosa* Roth (Yeater et al., 2004), *Oxalis tuberosa* L. (Emshwiller, 2002), *Medicago* spp. (Brummer et al., 1999), and *Bromus* spp. (Tuna et al., 2001).

Bonos et al. (2002) measured DNA content by flow cytometry to predict ploidy level and differentiate certain species of *Agrostis*. Bonos et al. measured the mean 2C DNA content of *A. canina* (3.42 pg), *A. stolonifera* (5.27 pg), *A. capillaris* (5.87 pg), *A. vinealis* (6.31 pg), *A. gigantea* (8.18 pg), and *A. castellana* (8.71 pg). These DNA content measures were consistent with previous reports by Hollman et al. (2005), Arumuganathan et al. (1999), and Arumuganathan and Earle (1991). Ploidy levels determined by chromosome counts and DNA content measures were correlated, allowing for differentiation of species. Similarly, Hollman et al. (2005) identified putative *A. canina* accessions based on random amplified polymorphic DNA (RAPD) markers and flow cytometry.

Tuna et al. (2001) found variation in 2C DNA content between *Bromus* species from the same ploidy level. Tuna et al. also found that the octaploid *Bromus* accessions had 1.2 pg less than twice the DNA content of the tetraploids suggesting reduced genome size with increased ploidy. These observations are consistent with the DNA content measures in *Agrostis*. For example, Bonos et al. found that the DNA content of the

tetraploid species varied by 1 pg per 2C and the DNA content of the tetraploid was less than twice the DNA content of the diploid. Chromosome counts could be performed in instances where DNA content cannot be used to clearly predict the ploidy level (Jones, 1956abc; Bonos et al., 2002; Zhao et al., 2006; Ahloowalia, 1965).

Understanding ploidy level is also important when examining the evolution of *Agrostis* species. For example, Vergara and Bughrara (2003) counted chromosomes from meiotic cells and determined that *Agrostis transcaspica* (PI283174) is a diploid, *A. munroana* (PI230236) and *A. lachnantha* (PI195917 and PI299461) accessions are triploid, and *A. mongolica* (PI362190) and *A. hygrometrica* (PI477045) are tetraploid. Neither of the sub-genomes ($A_2A_2A_3A_3$) of *A. stolonifera* are known, but based on ploidy and clustering of AFLP markers, Vergara and Bughrara hypothesize that *A. transcaspica* may be one of the diploid progenitors (A_3A_3 genome) of *A. stolonifera* and *A. gigantea.*

This study is designed to determine the ploidy level, predicted by flow cytometry, of the available NPGS *Agrostis* germplasm. Novel diploid and tetraploid accessions with important agronomic characteristics could be incorporated into breeding programs to improve cultivated germplasm.

MATERIALS AND METHODS

Plant Material

Agrostis germplasm (Appendix A) was obtained from NPGS (232 accessions, representing 21 species), Dr. Doug Johnson, research plant physiologist with the Forage and Range Research Laboratory of the United States Department of Agriculture (48

accessions, representing two species), or from Mr. Kevin Morris, executive director of the National Turfgrass Evaluation Program (25 accessions, representing two species). Two to five seeds were planted separately into Premier Pro-Mix BX soil (Premier Horticulture Inc., Quakertown, PA) in five 2.5 cm diameter Ray Leach Cone-tainers (Stuewe and Sons, Tangent, Oregon) and later culled such that no more than one seedling was present in each pot. They were grown in a temperature-controlled greenhouse with 75°F day/70°F night temperatures and a 12 hr photoperiod.

Flow Cytometry

Plants were grown to the third leaf stage and three to five one-inch pieces of leaf tissue were chopped finely with a razor blade in chopping buffer (Cystain UV Precise P, Partec, Munster, Germany), and allowed to incubate at room temperature for two minutes. Separately 5 μL of chicken red blood cells (CRBC, Pocono Rabbit Farm, Canadensis, PA) were added to 400 μL of the chopping buffer and allowed to incubate for at least one minute at room temperature. Three volumes of the Cystain UV Precise P, 4'-6-diamidino-2-phenylindole (DAPI) staining buffer was then added separately to the CRBC tube and to the chopped leaf tissue. The stained tissue and 1 ml of the CRBC mixture were passed through a 30 μM CellTrics filter (Partec) into the same sample tube and incubated on ice for 1 minute. The solution was then loaded directly onto the PA-II ploidy analyzer (Partec). Data was collected until clear peaks were formed, generally between 5,000 and 10,000 cell counts. The sample and CRBC peak data were analyzed with the Partec software to determine the best-fit Gaussian peaks. The 2C DNA content (pg) was determined by comparing the position of the sample flow cytometry peak to the

peak of the internal CRBC control with a known DNA content of 2.33 pg/2C (Bonos et al., 2002).

Chromosome Counts

Chromosome counts were performed on a select set of accessions to validate the flow cytometry data. Briefly, actively growing root tips were harvested into 0° C water, held for at least 24 hrs at 4° C, and transferred to Carnoy fixative (3:1 ethanol to acetic acid). The root tips were hydrolyzed in 1 N HCl for 10 minutes at 60° C. The acid was replaced by Feulgen stain and the root tip was stained at 22° C for 10 minutes. The root cap was removed from the root and placed into a drop of 45% acetic acid onto a glass slide. The root tip was squashed with a number 1 cover slip under moderate force. The chromosomes were visualized at 1200x magnification.

RESULTS

Figure 1 shows the smoothed cell count data from the ploidy analyzer plotted against 2C DNA content. The first of the four major peaks represents the CRBC control and the next three major peaks correspond with diploid, tetraploid, and hexaploid *Agrostis* accessions respectively. These three ploidy levels are clearly differentiated by the flow cytometry data. The 2C DNA contents of the available NPGS *Agrostis* accessions were measured by flow cytometry (Appendix A) and ranged from 3.64 to 21.66 pg with a mean of 11.03 +/- 2.92 pg.

The mean 2C DNA content (5.29 +/- 0.39 pg) was measured for the cultivated diploid *Agrostis canina* accessions Bavaria, Vesper, and SR7200 from the 1998 NTEP

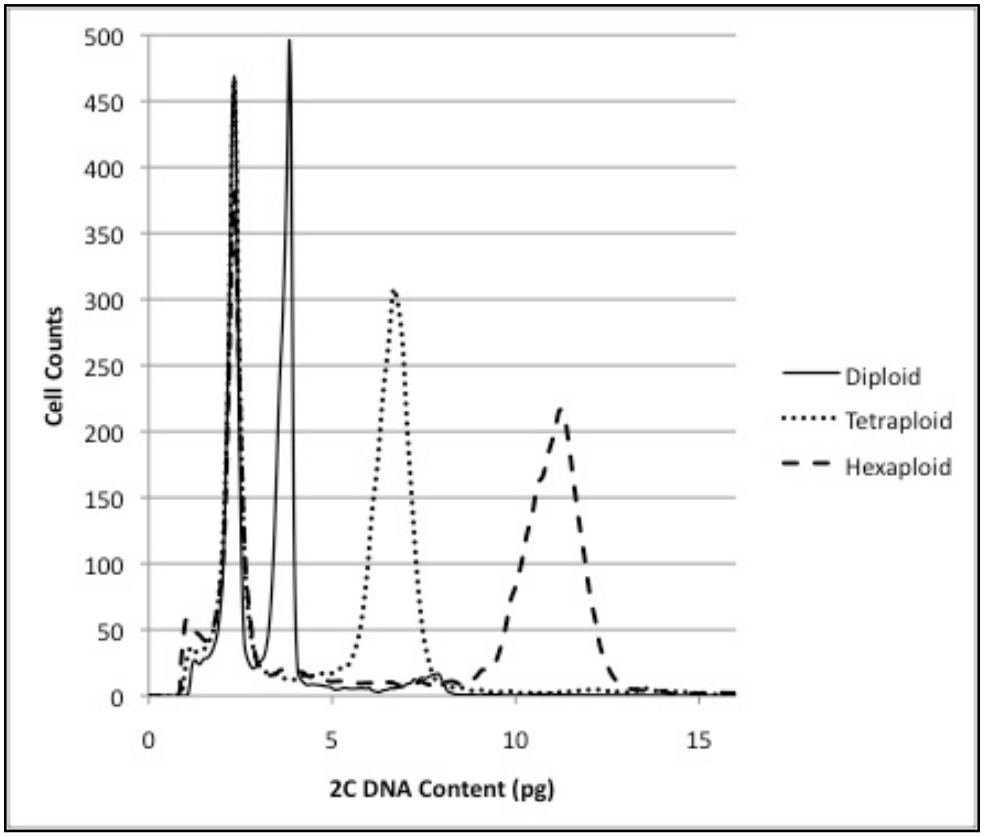


Figure 1. Ploidy analyzer trace data. Nuclei counts of an internal CRBC control and three *Agrostis* accessions representing three different ploidy levels.

bentgrass study, and SR7200, Legendary, Villa, Venus, Vesper, Greenwich from the 2003 bentgrass study. The 2C DNA content (8.69 +/- 0.45 pg) was measured from the cultivated *A. stolonifera* entries Penn G-6, Penn G-1, Penn A-4, and Penn A-2 of the NTEP 1998 bentgrass study and Penn A-1, Alpha, T-1, Declaration, Independence, Bengal, Kingpin, 007, Shark, MacKenzie, Tyee, and Pennlinks II entries of the 2003 bentgrass study. Candidate diploids were predicted as being less than the mean 2C DNA content of the NTEP cultivated *A. canina* accessions plus 2 standard deviations, or less than 6.07 pg/2C. Table 1 shows the NPGS accessions predicted to be diploid along with their DNA content measures. Tetraploids were predicted as being within +/- 2 sd of the mean 2C DNA content of the NTEP *A. stolonifera* accessions, or between 7.79 and 9.59 pg. The NPGS predicted tetraploid germplasm is listed in Table 2.

The 2C DNA contents of accessions used in the Vergara and Bughrara (2003) study were measured. *Agrostis munroana* (PI230236), *A. transcaspica* (PI283174), *A. lachnantha* (PI195917 and PI299461), *A. mongolica* (PI362190), and *A. hygrometrica* (PI477045) had 2C DNA content measures of 9.04, 10.22, 12.37, 12.77, 13.33, and 21.66 pg respectively.

Figure 2 shows chromosomes from the diploid *A. canina* cv. Bavaria (5.31 +/- 0.21 pg/2C), the tetraploid *A. stolonifera* cv. Penn A-2 (8.69 +/- 1.21 pg/2C), and the hexaploid *A. gingantea* cv. Reton (11.37 +/- 0.98 pg/2C). The counted number of chromosomes for these accessions was consistent with the ploidy level predicted by the flow cytometry data. The predicted ploidy level based on the flow cytometry data was supported by chromosome counts for several accessions (table 3).

Table 1. Predicted diploids based on 2C DNA content

Study ID	Accession	Taxon	2C DNA
acan-10	PI189141	*Agrostis canina*	3.64±0.18
acan-11	PI290707	*Agrostis canina*	3.66±0.13
acap-7	PI237717	*Agrostis capillaris*	5.74±1.04
acla-2	PI632584	*Agrostis clavata*	4.93±0.18
acla-5	W621240	*Agrostis clavata*	4.54±0.35
agig-1	RUS-07-01-002	*Agrostis gigantea*	5.12±0
amon-3	PI632549	*Agrostis mongolica*	5.2±0.24
amon-5	W619706	*Agrostis mongolica*	5.33±0.36
**apal-1*	PI238226	*Agrostis pallida*	6.09±0.9
asp-4	PI502279	*Agrostis sp.*	5.18±0.21
asto-40	PI632687	*Agrostis stolonifera*	6.02±1.83
atri-2	PI632559	*Agrostis trinii*	4.19±0.66
atri-4	W619572	*Agrostis trinii*	4.98±0.24
atri-5	W621432	*Agrostis trinii*	4.9±0.45
avin-1	PI440110	*Agrostis vinealis*	5.25±0.24

The DNA content of **apal-1* is slightly higher than the diploid cutoff value of 6.07 pg/2C.

Table 2. Predicted tetraploids based on 2C DNA content

Study ID	Accession	2C DNA
acap-1	PI171470	8.06±0.87
acap-13	PI325194	8.72±1.06
acap-14	PI392338	9.22±0.01
acap-15	PI420235	9.56±1.27
acap-2	PI172698	8.39±0.36
acap-21	PI494121	9.25±3.26
acap-24	PI578527	9.42±0.53
acap-9	PI283173	8.13±1.21
acas-11	PI240132	9.22±0.81
agig-10	RUS-07-27-103	7.95±0.33
agig-102	PI383584	8.37±1.59
agig-109	PI387903	9.46±1.97
agig-11	RUS-07-29-116	8.48±0.88
agig-12	RUS-07-32-131	8.92±0.53
agig-123	PI443051	8.91±0.66
agig-13	RUS-07-36-142	8.53±0.58
agig-137	PI538784	9.35±1.13
agig-18	KGZ-06-04-019	8.49±0.21
agig-3	RUS-07-04-016	7.84±0.19
agig-5	RUS-07-12-041	8.01±0.89
agig-57	PI206880	9.18±1.04
agig-76	PI251098	8.54±0.94
agig-77	PI251099	9.06±0.49
agig-79	PI251524	9.04±0.52
agig-85	PI267051	9.37±0.4
agig-86	PI272111	9.25±1.05
agig-87	PI274601	9.52±0.59
agig-92	PI298085	9.34±1.63
agig-98	PI311009	9.35±0.76
amun-6	PI229720	9.4±0.31
amun-7	PI230236	9.04±0.33
aneb-1	PI196319	9.53±1.52
asp-10	PI618782	9.1±0.74
asp-2	PI478596	8.59±0.1
asp-3	PI486302	7.94±0.39
asp-9	PI595078	9.06±0.78
asto-1	KGZ-06-08-051	8.19±0.5
asto-18	PI578530	9.49±0.54
asto-19	PI204390	8.94±0.67
asto-21	PI235440	9.43±0.47
asto-22	PI235541	8.75±1.29
asto-23	PI251945	9.04±0.84
asto-24	PI600776	9.17±0.67
asto-25	W66606	8.86±0.83
asto-28	PI269838	9.54±0.55
asto-29	PI302902	8.57±0.75
asto-30	PI318934	8.73±0.78
asto-32	PI494118	8.33±0.82
asto-35	W66602	8.19±0.27
asto-36	PI578529	8.52±0.62
atri-1	PI598462	7.9±0.44
atri-3	PI636572	9.46±0.52

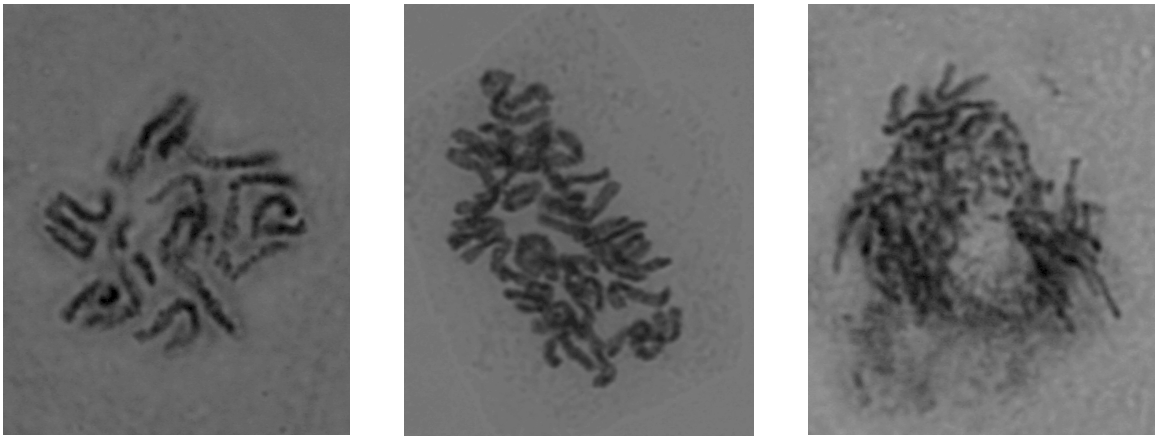


Figure 2. Feulgen Stained Chromosomes. Chromosomes of root tip squashes from *A. canina* cv. Bavaria (*acan-1*; 2n=2x=14), *A. stolonifera* cv. Penn A-2 (*asto-5*; 2n=4x=28), and *A. gingantea* cv. Reton (*agig114*; 2n=6x=42).

DISCUSSION

The commonly used turf-type *Agrostis* species are either diploid or tetraploid and therefore accessions at these ploidy levels are of most interest to the turfgrass industry. Bonos et al. (2002) reported the 2C DNA content of the diploid *A. canina* as 3.42 pg and *A. stolonifera* as 5.27 pg. In the present study, NTEP *A. canina* diploids had a mean 2C DNA content of 5.29 pg and the cultivated *A. stolonifera* had a mean 2C DNA content of 8.69 pg. The DNA content measures presented here were not consistent with previous studies, but they were very reproducible and allowed for the classification of *Agrostis* accessions into ploidy groups. Some of the differences in DNA content measures from this study in comparison to previous studies may be attributed to different DNA stains used, different equipment, or variation in the genome size of the CRBC control.

The DNA content measures in the present study of several NPGS accessions used in the Vergara and Bughrara (2003) study were not consistent with the reported ploidy levels. For example, the 2C DNA content of *A. munroana* (PI230236 - 9.04 pg/2C) suggests that it is a tetraploid but it was reported to be triploid. *A. lachnantha* accessions (PI195917 - 12.37 pg/2C; PI299461 - 12.77 pg/2C) were also reported to be triploid, but the flow cytometry data suggests they are hexaploid. The accessions *A. mongolica* (PI362190 - 13.33 pg/2C) and *A. hygrometrica* (PI477045 - 21.66 pg/2C) were reported to be tetraploid, but the flow cytometry data suggests they are hexaploid and higher ploidy respectively. Vergara and Bughrara also hypothesize that *A. transcaspica* (PI283174 - 10.22 pg/2C) is a diploid progenitor of *A. stolonifera*, but the flow cytometry data suggests it is a tetraploid.

Table 3. Validation of flow cytometry ploidy predictions

Study ID	Accession	Species	2C DNA	Ploidy	Count
acla-5	W621240	*Agrostis clavata*	4.54	2x	14
amon-3	PI632549	*Agrostis mongolica*	5.20	2x	14
amon-5	W619706	*Agrostis mongolica*	5.33	2x	14
asto-40	PI632687	*Agrostis stolonifera*	6.02	2x	14
atri-2	PI632559	*Agrostis trinii*	4.19	2x	14
atri-4	W619572	*Agrostis trinii*	4.98	2x	14
atri-5	W621432	*Agrostis trinii*	4.90	2x	14
acan-1	NT-1998-2	*Agrostis canina*	5.32	2x	14
acan-2	NT-1998-11	*Agrostis canina*	4.97	2x	14
acan-9	NT-2003-22	*Agrostis canina*	6.10	*2x	14
atra-1	PI283174	*Agrostis transcaspica*	10.22	*4x	28
agig-9	RUS-07-24-089	*Agrostis gigantea*	7.04	*4x	28
asto-4	NT-1998-27	*Agrostis stolonifera*	9.18	4x	28
asto-5	NT-1998-29	*Agrostis stolonifera*	8.69	4x	28

Accessions outside of the predicted ploidy level DNA content ranges are indicated (*). Ploidy - the predicted ploidy level based on the 2C DNA content measures. Count - the number of chromosomes counted from root tip squashes.

Cytogenetic experiments by Jones (1956abc) demonstrated the genome relationships between certain *Agrostis* species. The diploid *A. canina*, with the genome designation A_1A_1 or a close diploid ancestor is believed to have contributed one of the sub-genomes of *A. capillaris* ($A_1A_1A_2A_2$), while the other sub-genome is of unknown origin. Neither of the sub-genomes of creeping bentgrass ($A_2A_2A_3A_3$) is known. *Agrostis* accessions *acan-10, acan-11, acap-7, acla-2, acla-5, agig-1, amon-3, amon-5, *apal-1, asp-4, asto-40, atri-2, atri-4, atri-5,* and *avin-1* were predicted to be diploid and several accessions were also predicted to be tetraploid (tables 1 and 2). Further investigation of these accessions could provide insights into the evolution of cultivated *Agrostis* species and provide *Agrostis* breeders with a resource for understanding species relationships. This would greatly benefit breeding programs since a knowledge of crossability and speciation events would allow breeders to introduce novel genes through specific hybridization events or bridge species. For example, germplasm improvement breeding could be done, at the simpler diploid level followed by genome doubling and hybridization to further improve the cultivated tetraploids.

3. MITE-DISPLAY MARKER DEVELOPMENT IN *AGROSTIS*

Modern cultivars of *Agrostis stolonifera* have superior quality than those used just a few decades ago (2003 National Bentgrass Test, National Turfgrass Evaluation Program, http://www.ntep.org/). *Agrostis* species are adversely affected by biotic and abiotic stresses such as wear, heat, dollar spot (*Sclerotinia homoeocarpa* F.T. Benn.), and brown patch (*Rhizoctonia solani* Kühn). *Agrostis* breeding programs work to develop improved biotic and abiotic stress tolerant germplasm (Bonos et al, 2006) while making selections based on morphological characters that contribute to high turf quality at low mowing heights (Phillips, 2007; Warnke, 2003). The efficiency of new cultivar development could be improved by identifying molecular markers linked to traits of interest and making selections based on those molecular markers (Warnke, 2003), a process known as marker assisted selection.

Several molecular marker types have been implemented in the study of *Agrostis* species such as isozymes (Warnke et al., 1997), amplified fragment length polymorphism (AFLP; Vergara and Bughrara, 2003), random amplified polymorphic DNA (RAPD; Chakraborty et al., 2005), restriction fragment length polymorphism (RFLP; Chakraborty et al., 2005), sequence characterized amplified region (SCAR; Scheef et al, 2003) and simple sequence repeat (SSR; Jensen, et al., 2007). Multi-allelic markers such as SSRs can be difficult to interpret in outcrossing polyploids, such as *A. stolonifera*. As a result

marker types such as AFLPs are commonly used because they generate large numbers of single-dose dominant markers.

Clustering of AFLP markers has been observed in several mapping projects (Wang et al., 2005; Castiglioni et al., 1999; Kang et al., 2001) and thus the genome coverage provided by these markers may be limited. Combining multiple marker types or using different restriction enzymes in the early stages of AFLP is a common practice in genetic linkage map development and provides greater genome coverage (Zhao et al., 2006). Chakraborty et al. (2005) reported the first genetic linkage map of *Agrostis stolonifera.* This map includes 169 RAPD, 180 AFLP, and 75 cDNA RFLP markers for a total of 424 mapped loci covering 1,110 cM. The development of additional genetic markers would provide a more detailed linkage map of *A. stolonifera* and would be a resource for marker assisted selection based breeding strategies.

MITE-display is a modified AFLP technique that anchors the amplified fragments to miniature inverted-repeat transposable elements (MITE). McClintock (1950) was the first to identify transposable elements while investigating mutable loci in maize. McClintock found activator (Ac) and dissociation (Ds) elements that induced chromosome breakage and were responsible for the observed mutations. Transposable elements are present in high copy numbers in grass species (Turcotte et al., 2001; Meyers et al., 2001) and may have contributed to the expansion of certain genomes (Vicient et al., 1999; Sanmiguel and Bennetzen, 1998, Wessler et al., 1995). Transposable elements are divided into two main classes based on their mode of transposition; elements that transpose via an RNA intermediate are type I elements and those that transpose via a

DNA intermediate are type II. Each type is further divided into autonomous and non-autonomous elements based on if the element encodes the machinery needed for its own mobilization. MITEs are a class of non-autonomous type II elements, first described by Bureau and Wessler (1992). MITEs are characterized by their small size, flanking target site duplications (TSD), terminal inverted repeats (TIR) and ability to achieve high copy numbers. An example sequence alignment of three MITEs from the *MDM*-2 class is shown in figure 3 (Yang and Hall, 2003b). Certain classes of MITEs are prevalent in non-coding regions of genes (Bureau and Wessler, 1992, 1994ab; Bureau et al., 1996) and therefore MITE-display could target transcriptionally active regions of the genome in contrast to AFLP markers that are targeted to non-coding regions of the genome (Meudt and Clarke, 2007). The primary means of identifying new MITEs is by searching sequence databases and by PCR-based approaches (Tu, 2001; Yang and Hall, 2003a; Lyons et al., 2008).

There is limited *Agrostis* sequence data available in publicly available databases. As of March 25th, 2009 the National Center for Biotechnology Information (NCBI) had 16,992 *Agrostis* EST sequences. The majority of this sequence data was part of a collaborative EST sequencing project (Rotter et al., 2007). The objective of the present study is to determine if MITEs can be identified within this *Agrostis* sequence library and to ascertain if a set of molecular markers can be developed from any identified MITEs. MITE-based molecular markers may target new regions of the *Agrostis* genome and provide more detailed genetic linkage maps. An improved genetic linkage map of *Agrostis* would help breeders to identify agronomically important loci and to implement

Figure 3. Sequence alignment of *MDM*-2 MITEs from *Oryza sativa*. GenBank accession numbers of MITE sequences are given. Conserved residues are boxed in gray. The target site duplications are shown by the arrow and the terminal inverted repeats are indicated by the solid black bar.

marker assisted selection for the genetic improvement of *Agrostis*.

MATERIALS AND METHODS

DNA Libraries and MITE Identification

Random sheared genomic DNA (RS) and EST DNA sequences from both *Agrostis stolonifera* and *A. capillaris* were obtained from Dr. Faith Belanger (Rutgers, The State University of New Jersey). MITEs were identified from the *Agrostis* sequence libraries by the software program FindMITE (Tu, 2001) using the following criteria, 11bp TSD, a minimum length of 50 bp, and TIRs with the sequences TA, TAA, TAT, TTA, NNNN, or NNNNNNNN. These settings were chosen because they have been used previously to identify both novel and previously reported MITE classes (Tu, 2001).

MITE-insertional polymorhpism

Evidence of MITE transposition was tested by MITE insertional polymorphism (MIP; Lyons et al., 2008) on a diverse set of *Agrostis* germplasm. Briefly, MITE flanking PCR primers were designed with the software program primer3 (Rozen and Skaletsky, 2000) and obtained from Integrated DNA Technologies (IDT; Coralville, IA). Standard CTAB DNA extraction was performed to isolate genomic DNA (Doyle and Doyle, 1987). Ten ng of genomic DNA, 1 μL each of 5 μM MITE-flanking primers, and 1x GoTaq (Promega, Madison, WI) was used to amplify MITE target regions. All PCR reactions were carried out on a GeneAmp 2700 thermalcycler (Applied Biosystems, Foster City, CA) with an initial denaturation step of 94° C for 5 min, then 94° C for 1 min, 56° C for 1 min, and 72° C for 1 min for 40 cycles, 72° C for 5 min, and finally held

at 4° C until they were analyzed by gel electrophoresis. The PCR products were run on a 1% agarose 1x TAE.

MITE-display

MITE-display was carried out as described by Lyons et al. (2008). MITE-int and MITE-ext nested primers (table 4) were designed by primer3 and obtained from IDT. The MITE-int primers were labeled with 5'-6-carboxyfluorescein (6-FAM). A CTAB method (Doyle and Doyle, 1987) was used to extract total DNA from the parents and 94 progeny of an *Agrostis* experimental mapping population. Genomic DNA (500 ng) was digested with 2 Units DpnII (New England Biolabs, Beverly, MA) at 37° C for 3 hrs followed by heat inactivation of the enzyme (65° C for 20 minutes). The digested DNA was purified using the QIAquick 96 PCR Purification Kit (Qiagen, Valencia, CA) and eluted in 40 μL 10 mM Tris-HCl. Adaptors were prepared by adding equal amounts of the single stranded adaptors (25 μM), TDADA1 and TDADA2 (table 4), heating them to 80° C for 2 min and then allowing to cool to room temperature. Adaptors were ligated to the digested DNA using 200 Units of T4 DNA Ligase (New England Biolabs) and 20 Units of BamHI (New England Biolabs) was added to inhibit dimer formation. The ligation reaction involved incubating the sample for 30 min at 16° C followed by 10 min of 37° C for 12 cycles. Pre-selective and selective amplifications were performed as described by Lyons et al. Amplified products were run on an Applied Biosystems 3730 genetic analyzer (Applied Biosystems, Foster City, CA) and visualized in the software program genographer (Amundsen and Warnke, 2005; Benham et al., 1999).

MITE-display genetic markers were scored as 1 for present and 0 for absent. The

Table 4. MITE-display primer sequences

Clone ID	Nested Internal MITE Primer	External MITE Primer
CRP289T_E09	*ACGAACGGGTAGGTCTTGAGG	TTCCTTGACGAACGGGTAGGT
COEST06T_B08	*ACTTTGGGGACGCCGAC	CGTCAGCGACGACTTTG
CRP238T_C07	*CAAAACCTACATCACACCCTC	GTACACAAAACCTTCGGTCAG
Colonial2T_H05	*CACCGCATCATCCAATCT	TTATTCCTCCACCGCATCATC
CRP239T_G02	*CGAGCCTGAAGGTTCTCTCG	TCGTCGATAAGGCGAGC
CRP263T_C09	*CGTCCAACTGGAGATTGAAGC	CATCGAAATGGAAGATTATCG
COL403T_H08	*CTGATGGCATTTCTAGCGGAG	CTCTGATGGCATTTCTAGCGG
COL404T_F01	*CTGCCCATGTTCGGGTG	GGCAGGTACTGGACAATGTGG
Colonial7T_H07	*TGGCCCGGGACTAATGG	CATTGGTCCCGGTTCAT
COL383T_E03	*TTACTTCTGGTGCCTCAACTC	GCCTGTCACAATTGGAGTTAC
TDADA1 - adapter 1	GACAGTTGTGTACCTCGAATG	
TDADA2 - adapter 2	GATCCATTCGAGGTACACAACTG	
TDPre - preselective	GTTGTGTACCTCGAATGGATCA	
TDSelC - selective	GTTGTGTACCTCGAATGGATCAC	
TDSelG - selective	GTTGTGTACCTCGAATGGATCAG	

* Primers labeled with 5'-6-carboxyfluorescein (6-FAM).

polymorphic information content (PIC) was calculated for each marker as:

$$\mathrm{PIC_i} = 1 - \sum_{j=1}^{n} p_{ij}^2$$

where p_{jj} is the frequency of the jth allele for the ith marker (Anderson et al., 1993).

RESULTS

FindMITE was used to identify MITEs from 7,024 EST and 666 RS sequences from *Agrostis capillaris* and 7,772 EST and 660 RS sequences from *A. stolonifera*. Candidate MITEs were found in 231 EST and 66 RS sequences from the *A. capillaris* libraries and 100 EST and 98 RS sequences from the *A. stolonifera* libraries. A higher percentage (12.4%) of MITEs was found in the RS sequence libraries than in the EST sequence libraries (2.2%). The FindMITE data is summarized in table 5.

Primer3 was able to design 158 MIP primer pairs (Appendix B) from the 495 MITE-containing sequences. A panel of tetraploid *Agrostis* germplasm was screened with a subset of the MIP primer pairs. The tetraploids had complex MIP banding profiles making it difficult to determine if PCR product size variation was the result of MITE activity (data not shown). To avoid this complication, a diverse panel of *Agrostis* diploid germplasm, identified in chapter 2, was tested with each of the 158 MIP primer pairs. A total of 146 primer pairs amplified a product and 79 exhibited polymorphic banding pattern consistent with MITE transposition activity. Figure 4 shows a MIP banding profile. In this example, primers were designed to flank the candidate MITE from the *A. capillaris* clone, Colonial2T_H05 with the TSD sequence CAAC, TIR sequence

Table 5. Number of candidate MITEs from *Agrostis* DNA sequence libraries

TSD	*A. capillaris*	*A. stolonifera*	Total
TA	110	110	220
TAA	2	4	6
TAT	2	5	7
TTA	5	3	8
NNNN	174	73	247
NNNNNN	4	3	7
Total	297	198	495

Number of candidate MITEs predicted by the software program FindMITE based on the specified TSD (target site duplications).

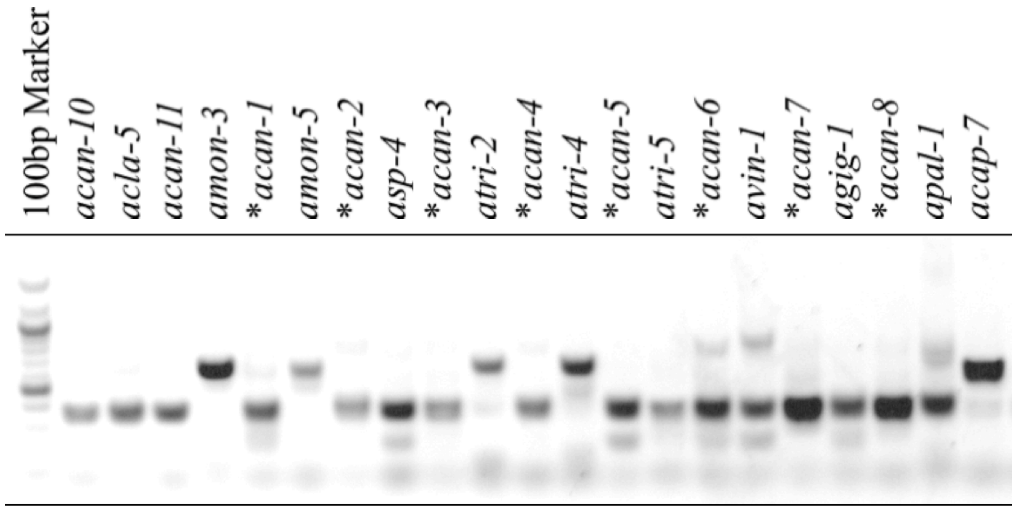


Figure 4. *Agrostis* diploid MIP screen. MIP band profile generated from the primer pair designed to flank the predicted MITE in the random sheared clone Colonial2T_H05. NTEP *A. canina* accessions are identified with a star (*).

ATTGGTG[AG]ATG, and a predicted size of 188 bp. The size differences of the primary fragments between species correspond with the predicted MITE size plus the neighboring sequence flanked by and including the PCR primers.

Parents of an *Agrostis* experimental linkage mapping population were tested with 10 MITE-display primer pairs (Table 4) amplified separately with the selective TDSelC primer (table 4; Lyons et al., 2008). Each of the tested primer combinations resulted in MITE-display banding profiles (Figure 5). Two of the MITE-display primers along with the selective primers TDSelC and TDSelG separately were screened on the parents and progeny of an experimental *Agrostis* mapping population. The four selective MITE-display reactions amplified 141 polymorphic markers with a mean PIC of 0.89. There were 26 markers that fit expected segregation ratios (1:1, 3:1) within this mapping population.

DISCUSSION

The majority of the *Agrostis* DNA sequence data available through NCBI databases originated from a single sequencing project (Rotter et al., 2007). A FindMITE search of these sequences effectively identified candidate MITEs. Certain classes of MITEs are associated with genes and have been found in introns and other nontranscribed gene regions. It was surprising to find 331 MITEs in EST sequences since transposition into an exon would likely disrupt gene function. Sequence analysis would reveal if these MITEs were located near the 3' end of the transcripts where they might not impact gene function.

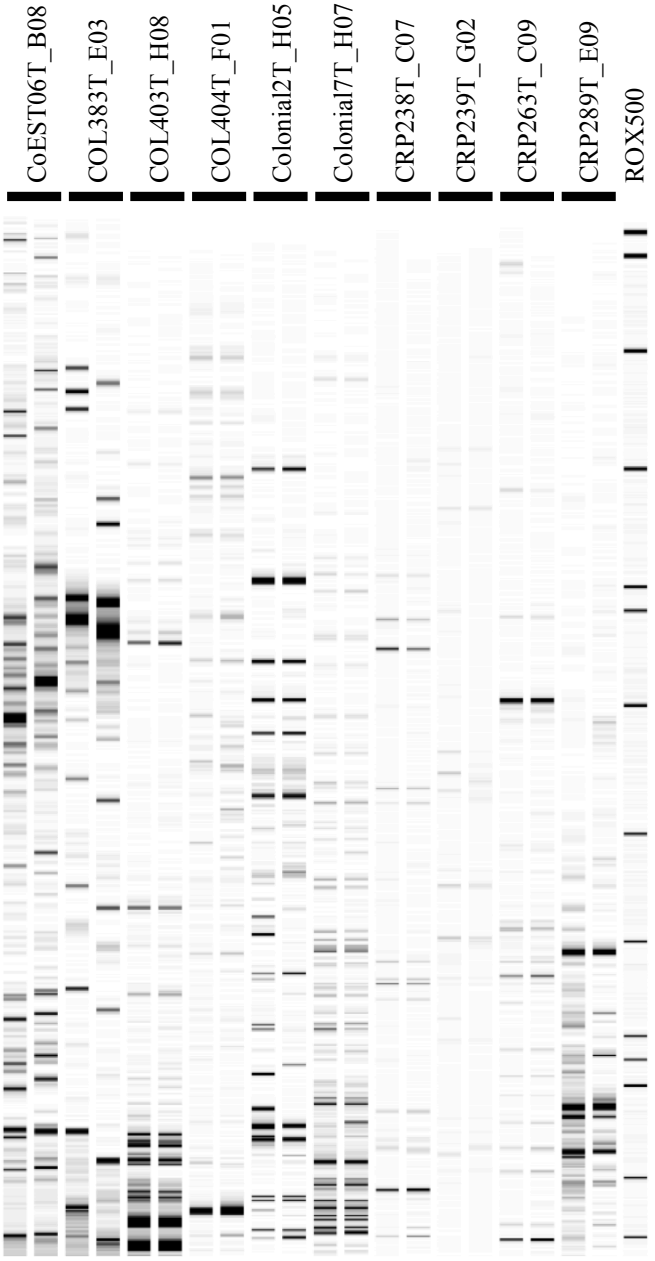


Figure 5. MITE-display Test. MITE-Display banding profile of 10 primer combinations on *Agrostis stolonifera* parents of an experimental mapping population visualized with the software program Genographer (Benham et al., 1999; Amundsen and Warnke, 2005)

The software program FindMITE searches for sequences with MITE characteristics, so to ensure that predicted MITEs are not sequence artifacts, MIP was used to demonstrate transposition. Transposition of the candidate MITE is implied if the predicted MITE is present in some *Agrostis* germplasm and absent from others. Polymorphic banding profiles were observed in 54% of the 146 MIP primer pairs. A more diverse *Agrostis* panel may exhibit MIP banding patterns with the remaining 67 MIP primer pairs.

Six MITE-display primer pairs were screened on the full mapping population and 141 polymorphic markers were identified. These data demonstrate the effectiveness of MITE-display as a marker tool for linkage map development in *Agrostis*. This is the first report of MITEs and use of MITE-display to develop genetic markers in *Agrostis*. MITE-display is no more complicated in practice than AFLP, but upfront experimentation is needed to develop effective MITE-display markers. Future experiments could determine if MITE-display genetic markers complement AFLP markers and provide more detailed and extended genetic linkage maps.

MITEs have been found to transpose near genes (Bureau and Wessler, 1992; Bureau and Wessler, 1994ab; Bureau et al., 1996). MITEs have also been found to mobilize under stress conditions (Jiang et al., 2003; Kikuchi et al., 2003) and may provide a source of genetic variation in perennial species. These observations make MITE-based markers an attractive tool for genetic linkage map development in *Agrostis*.

4. MITE-DISPLAY DIVERSITY ANALYSIS IN *AGROSTIS*

The five *Agrostis* species commonly used for turf are *A. stolonifera* L., *A. capillaris* L., *A. castellana* Boiss. and Reut., *A. canina* L., and *A. gigantea* Roth. (Warnke, 2003). Novel stress tolerance genes from compatible extant *Agrostis* germplasm can be introduced into cultivated *Agrostis* through wide-cross breeding (Belanger et al., 2003ab). The National Plant Germplasm System (NPGS) has 368 available *Agrostis* accessions (as of April 19, 2009) representing more than 25 distinct species and it is the largest source of publicly available *Agrostis* germplasm within the United States. This collection accounts for less than 17% of the between 150 and 200 known *Agrostis* species (Harvey, 2007) but may be a source of novel stress tolerance genes.

Interspecific hybridization occurs between *Agrostis* species in natural populations and it has been shown to be effective as a method for improving *A. stolonifera* (Belanger et al., 2003ab). Belanger et al. (2003b) generated successful hybrids between wide crosses of *A. stolonifera* and *A. canina*, *A. capillaris*, *A. gigantea*, and *A. castellana*. *A. capillaris* is generally considered to be more tolerant of dollar spot than *A. stolonifera* (Belanger et al., 2004). Dollar spot is caused by the fungal pathogen *Sclerotinia homoeocarpa* F.T. Bennett and causes significant damage to *A. stolonifera* golf course putting greens (Walsh et al, 1999). Belanger et al. (2004) generated hybrids between *A.*

stolonifera and *A. capillaris* and identified hybrid progeny with reduced dollar spot severity compared to the *A. stolonifera* parent; dollar spot resistance was introduced into *A. stolonifera* from an *A. capillaris* donor.

Germplasm diversity is a major consideration for wide-cross breeding strategies since distinct germplasm is a potential source for novel stress tolerance genes. Several genetic marker types have been used to assess genetic diversity in *Agrostis* collections. Isozyme markers were used to determine the amount of genetic variation within a group of 18 *A. stolonifera* accessions (Warnke et al., 1997). A European plant introduction was found to be genetically distinct from the United States accessions. Similarly, amplified fragment length polymorphic (AFLP) markers were used to explore the genetic diversity between 39 *A. capillaris* accessions (Zhao et al., 2006) and random amplified polymorphic DNA (RAPD) markers were used to assess the genetic diversity in a set of 18 *A. capillaris* accessions (Karlsen and Steiner, 2007). These studies support the finding of Warnke et al. (1997) that European germplasm is genetically distinct from United States germplasm. Vergara and Bughrara (2003) conducted the most comprehensive study of *Agrostis* interspecies genetic diversity by using AFLP markers. Their study included a set of 40 *Agrostis* accessions representing 14 species from the USDA, regional plant introduction station (Pullman, WA); the accessions were from different geographic regions and ploidy levels. The UPGMA clustering of AFLP markers divided the accessions into seven groups, three of which were composed of single accessions and the remaining four clustered closely with species designations. Based on chromosome counts, Vergara and Bughrara reported that PI383584 (*A. transcaspica*; *atra-1*) is a

diploid and based on clustering of AFLP markers they hypothesize that *A. transcaspica* may be one of the unknown diploid progenitors of *A. stolonifera* and *A. gigantea*.

Genetic markers are helpful in understanding the diversity and evolutionary history of closely related species. For example, MITE-display molecular markers were used to examine the relationships between a polyploid series of *Oryza* species (Park et al., 2003). Miniature inverted-repeat terminal elements (MITEs) are a class of non-autonomous type II transposable elements (Bureau and Wessler, 1992). They are found in high copy number relative to other type II, DNA based elements and are characterized by their small size (less than 500 bp), target site duplications (TSD), and terminal inverted repeats (TIR). MITE-display is a modified AFLP technique that anchors the amplified fragments to MITEs (Casa et al., 2004). Park et al. found several MITE-display markers to be genome specific, allowing for a means of exploring polyploid origins and sub-genome relationships between accessions of higher ploidy. MITE-like sequences are prevalent in *Agrostis* and were demonstrated to be useful molecular markers (chapter 3). *Agrostis* species relationships and the amount of genetic diversity within and between accessions could be resolved with MITE-display markers.

In the present study MITE-display markers developed in chapter 3 will be used to assess the genetic diversity within and between cultivated and novel NPGS *Agrostis* germplasm. Cluster analysis will be used to infer species relationships.

MATERIALS AND METHODS

Plant Material

Seeds of *Agrostis* germplasm (Appendix A) were obtained from the National Plant Germplasm System, Dr. Doug Johnson (research plant physiologist with the Forage and Range Research Laboratory of the United States Department of Agriculture), or from Mr. Kevin Morris (executive director of the National Turfgrass Evaluation Program). Two to five seeds were planted separately into Premier Pro-Mix BX soil (Premier Horticulture Inc., Quakertown, PA) in five 2.5 cm diameter Ray Leach Cone-tainers (Stuewe and Sons, Tangent, Oregon) and later culled such that no more than one seedling was present in a single pot. They were grown at 75°F day/70°F night temperatures on a 12 hr photoperiod in a climate-controlled greenhouse. Total DNA was extracted by a standard CTAB extraction procedure (Doyle and Doyle, 1987).

MITE-display

MITE-display was performed as described in chapter 3 on 383 DNA samples extracted separately from three to five replicates of 81 *Agrostis* accessions. GeneMapper 3.7 (Applied Biosystems, Foster City, CA) was used to automatically create a scoring bin set. Individual bins were defined if any sample had a peak at a given bp size with relative fluorescence units greater than 1000. All samples were then analyzed with the defined bin allowing GeneMapper3.7 to automatically define the peaks. A string of scores, one if a band was present and zero if absent, was generated for each sample. A bulked sample table was made by enforcing a majority rule to the scoring of replicates for a given accession. For example, if three or more of five replicates of an accession contained a peak at a given bin location, the bulk was assigned a one and if three or more replicates did not have a peak at that bin location the bulk was assigned a zero.

Diversity Analysis

The binary matrix of scored MITE-display markers was analyzed with NTSYSpc version 2.20e. Briefly SimQual was used to construct a pair-wise genetic similarity matrix using the sequential agglomerative hierarchical nested (SAHN) cluster analysis method (Sneath and Sokal, 1973) implementing the Dice coefficient (Dice, 1945). Unweighted pair group method with arithmetic mean (UPGMA) clustering was performed on the similarity matrix and represented as a dendrogram. The quality of the clustering was determined using COPH and MXCOMP. The DCENTER and EIGEN procedures were used to perform a three dimensional principle component analysis.

The software program Structure 2.2 searched the bulked binary data for K=2 to 12 discrete subpopulations using correlated gene frequencies with admixture models. A total of 10,000 iterations were performed following 10,000 burn-in iterations.

RESULTS

Six MITE-display primer combinations (table 6) were run on a panel of DNA from *Apera*, *Polypogon*, and *Agrostis* accessions (table 7). GeneMapper3.7 generated 1,356 bins of which 1,309 were polymorphic and subsequently scored on the *Apera*, *Polypogon*, and *Agrostis* accessions. The mean genetic similarity (GS) of the 81 accessions was 0.43. Accessions *aint-1* and *agig-1* were the least similar (GS=0.16) and accessions *asto-3* and *asto-6* were the most similar (GS=0.73). The cophenetic correlation showed a good correlation of the clustering of MITE-display markers to the data, r=0.970. Eight groups were identified by the UPGMA cluster analysis (figure 6).

Table 6. Number of MITE-display markers generated with each primer pair

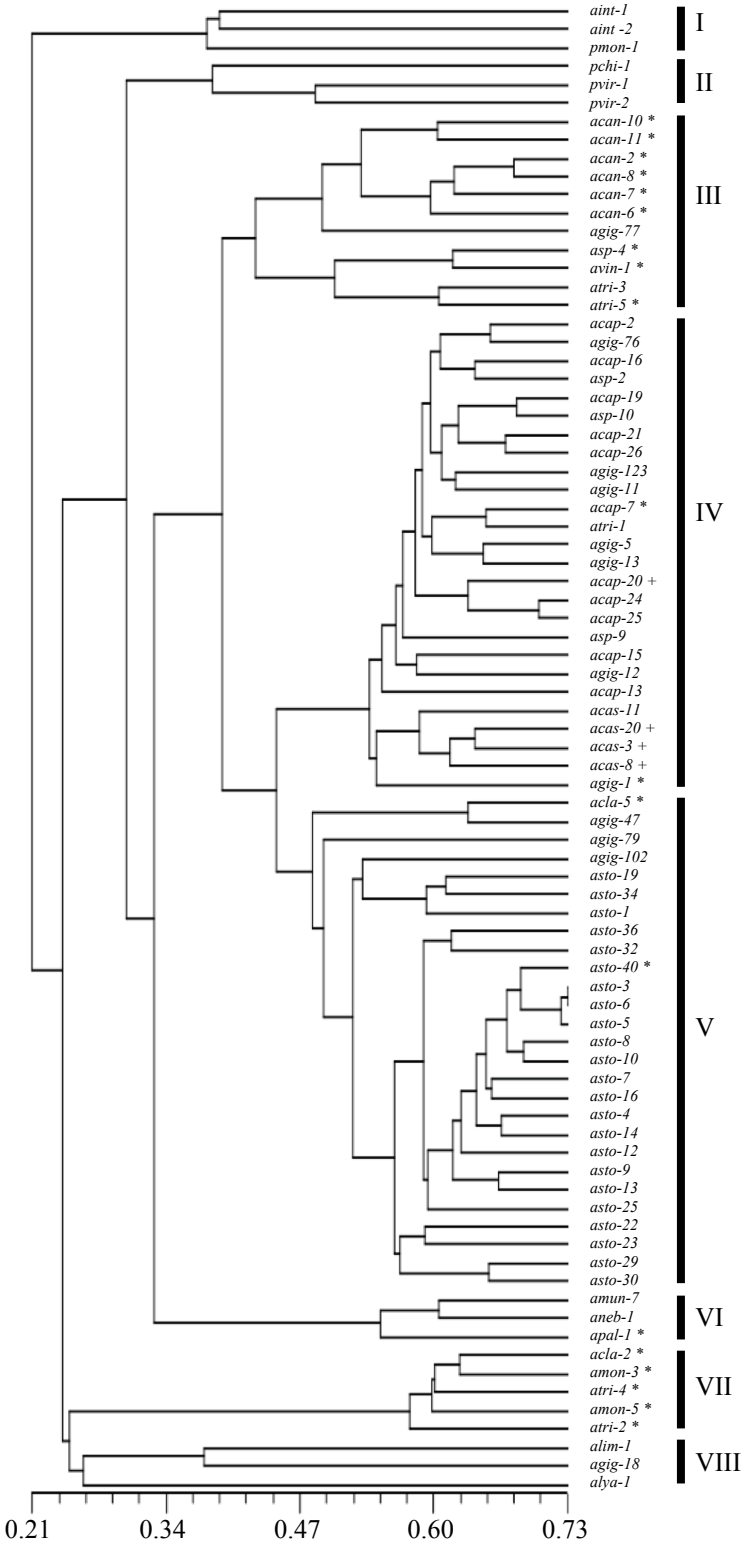
MITE Specific Primer	MITE-display Primer	# Markers
Colonial7T_H07	TDSelC	262
	TDSelG	273
Colonial2T_H05	TDSelC	301
	TDSelG	176
CRP263T_C09	TDSelC	162
	TDSelG	182
Total		1,356

Table 7. MITE-Display Diversity Study Accession List

	Accession	Taxon	2C DNA
aint-1	PI20344	*Apera interrupta*	nd
aint-2	PI20438	*Apera interrupta*	nd
pchi-1	PI37222	*Polypogon chilensis*	nd
pmon-1	PI31746	*P. monspeliensis*	nd
pvir-1	PI20439	*Polypogon viridis*	nd
pvir-2	PI22324	*Polypogon viridis*	nd
acan-10	PI18914	*Agrostis canina*	3.64
acan-11	PI29070	*Agrostis canina*	3.66
acap-2	PI17269	*Agrostis capillaris*	8.39
acap-7	PI23771	*Agrostis capillaris*	5.74
acap-13	PI32519	*Agrostis capillaris*	8.72
acap-15	PI42023	*Agrostis capillaris*	9.56
acap-16	PI42023	*Agrostis capillaris*	9.66
acap-19	PI49126	*Agrostis capillaris*	7.68
acap-20	PI49412	*Agrostis capillaris*	11.02
acap-21	PI49412	*Agrostis capillaris*	9.25
acap-24	PI57852	*Agrostis capillaris*	9.42
acap-25	PI57852	*Agrostis capillaris*	9.91
acap-26	PI60093	*Agrostis capillaris*	9.70
acas-11	PI24013	*Agrostis castellana*	9.22
acas-20	PI24014	*Agrostis castellana*	10.34
acas-3	PI31892	*Agrostis castellana*	10.42
acas-8	PI46921	*Agrostis castellana*	13.32
acla-2	PI63258	*Agrostis clavata*	4.93
acla-5	W6212	*Agrostis clavata*	4.54
agig-76	PI25109	*Agrostis gigantea*	8.54
agig-77	PI25109	*Agrostis gigantea*	9.06
agig-79	PI25152	*Agrostis gigantea*	9.04
agig-102	PI38358	*Agrostis gigantea*	8.37
agig-123	PI44305	*Agrostis gigantea*	8.91
alim-1	W6236	*Agrostis limprichtii*	9.64
alya-1	PI63665	*Agrostis lyallii*	6.86
amon-3	PI63254	*Agrostis mongolica*	5.20
amon-5	W6197	*Agrostis mongolica*	5.33
amun-7	PI23023	*Agrostis munroana*	9.04
aneb-1	PI19631	*Agrostis nebulosa*	9.53
apal-1	PI23822	*Agrostis pallida*	6.09
asp-2	PI47859	*Agrostis sp.*	8.59
asp-4	PI50227	*Agrostis sp.*	5.18
asp-9	PI59507	*Agrostis sp.*	9.06
asp-10	PI61878	*Agrostis sp.*	9.10
asto-36	PI57852	*Agrostis stolonifera*	8.52

ID	Accession	Taxon	2C DNA
asto-40	PI632687	*Agrostis stolonifera*	6.02
asto-19	PI204390	*Agrostis stolonifera*	8.94
asto-22	PI235541	*Agrostis stolonifera*	8.75
asto-23	PI251945	*Agrostis stolonifera*	9.04
asto-25	W66606	*Agrostis stolonifera*	8.86
asto-29	PI302902	*Agrostis stolonifera*	8.57
asto-30	PI318934	*Agrostis stolonifera*	8.73
asto-32	PI494118	*Agrostis stolonifera*	8.33
asto-34	PI531251	*Agrostis stolonifera*	7.53
atri-1	PI598462	*Agrostis trinii*	7.90
atri-2	PI632559	*Agrostis trinii*	4.19
atri-3	PI636572	*Agrostis trinii*	9.46
atri-4	W619572	*Agrostis trinii*	4.98
atri-5	W621432	*Agrostis trinii*	4.90
avin-1	PI440110	*Agrostis vinealis*	5.25
agig-1	RUS-07-01-002	*Agrostis gigantea*	5.12
agig-5	RUS-07-12-041	*Agrostis gigantea*	8.01
agig-11	RUS-07-29-116	*Agrostis gigantea*	8.48
agig-12	RUS-07-32-131	*Agrostis gigantea*	8.92
agig-13	RUS-07-36-142	*Agrostis gigantea*	8.53
agig-18	KGZ-06-04-019	*Agrostis gigantea*	8.49
asto-1	KGZ-06-08-051	*Agrostis stolonifera*	8.19
agig-47	PRC06IN/11840	*Agrostis gigantea*	9.84
asto-3	NT-1998-26	*Agrostis stolonifera*	8.93
asto-4	NT-1998-27	*Agrostis stolonifera*	9.11
asto-5	NT-1998-29	*Agrostis stolonifera*	8.69
asto-6	NT-2003-2	*Agrostis stolonifera*	8.76
asto-7	NT-2003-6	*Agrostis stolonifera*	8.76
asto-8	NT-2003-7	*Agrostis stolonifera*	8.81
asto-9	NT-2003-10	*Agrostis stolonifera*	9.41
asto-10	NT-2003-11	*Agrostis stolonifera*	8.83
asto-12	NT-2003-15	*Agrostis stolonifera*	7.93
asto-13	NT-2003-21	*Agrostis stolonifera*	8.56
asto-14	NT-2003-23	*Agrostis stolonifera*	8.51
asto-16	NT-2003-25	*Agrostis stolonifera*	8.74
acan-2	NT-1998-11	*Agrostis canina*	4.97
acan-6	NT-2003-16	*Agrostis canina*	5.48
acan-7	NT-2003-18	*Agrostis canina*	5.15
acan-8	NT-2003-19	*Agrostis canina*	5.61

Figure 6. Cluster Analysis of 81 *Agrostis* Accessions. UPGMA clustering of the similarity values between accessions calculated using the Dice (1945) coefficient. Accessions predicted to be diploid (*) or hexaploid (+) based on the flow cytometry data are indicated. The remaining accessions are predicted to be tetraploid.

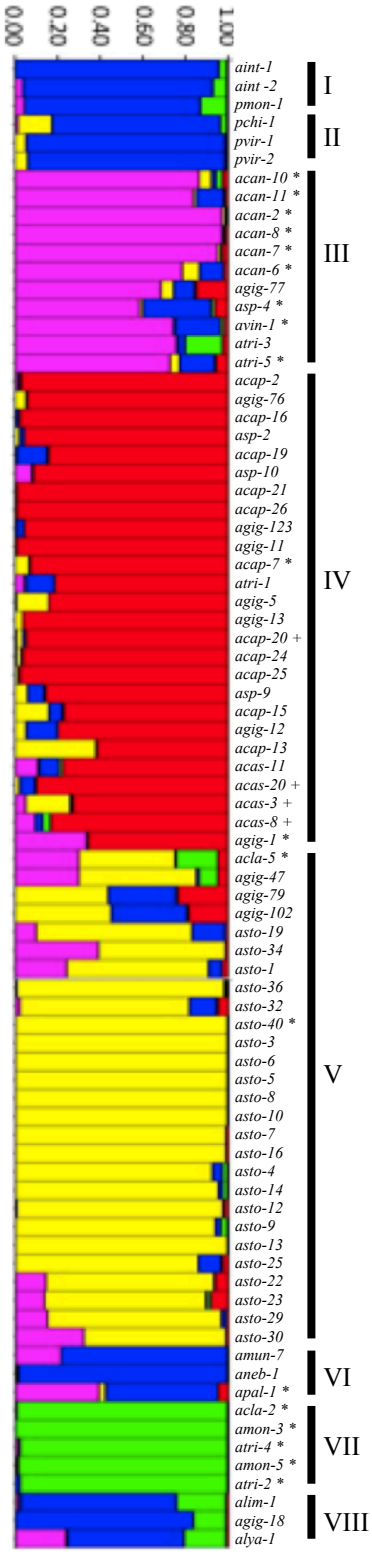


The software program Structure 2.2 was used to test for substructure within the MITE-display marker data. Several *k* values were tested (*k*=2 to 12) and dividing the data into *k*=5 distinct subgroups was sufficient to explain the distribution of MITE-display markers within these 81 accessions (figure 7). The *k*=5 groups preserve the 8 groups identified by the UPGMA analysis with the exception of UPGMA groups I, II, VI, and VIII which formed a single group in the Structure 2.2 analysis. Principal component analysis (figure 8) also supported the groupings defined by the UPGMA analysis, although groups I, II, VI, VII, and VIII were not as well defined.

Group I (GS=0.38) is made up of the *Apera* and *Polypogon* accessions *aint-1*, *aint-2*, and *pmon-1*. This group was the least similar to the rest of the *Agrostis* accessions (GS=0.22). Group II (GS=0.42) is composed of the *Polypogon* accessions *pchi-1*, *pvir-1*, and *pvir-2*. The GS between these accessions and the *Agrostis* accessions was 0.298. Group III (GS=0.49) contains the *Agrostis canina* accessions *acan-2*, *acan*-[*6* through *8*], *acan-10*, and *acan-11*, *A. gigantea* accession *agig-77*, undefined species accession *asp-4*, and *A.vinealis* accessions *avin-1*, *atri-3* and *atri-5*. Within group III *acan*-[*1* through *9*] accessions represent cultivated *A. canina* accessions supplied by NTEP and are very similar (GS=0.62).

Group IV (group GS=0.57) is made up of *Agrostis capillaris* accessions *acap-2*, *acap*-7, *acap-13*, *acap-15*, *acap-16*, *acap*-[*19* through *21*], *acap*-[*24* through *26*], *A. castellana* accessions *acas-3*, *acas-8, acas-11*, and *acas-20*, *A. gigantea* accessions *agig-1*, *agig-5*, *agig*-[*11* through *13*], *agig-76*, and *agig-123*, undefined species accessions *asp-2*, *asp-9*, and *asp-10*, and *A. vinealis* accession *atri-1*. The *A. capillaris* accessions have a

Figure 7. Structure Analysis of 81 *Agrostis* accessions. The length of the colored bars indicates the probability of each accession being grouped into the 5 defined groups (indicated by the different colors). Accessions predicted to be diploid (*) or hexaploid (+) based on the flow cytometry data are indicated. The remaining accessions are predicted to be tetraploid.



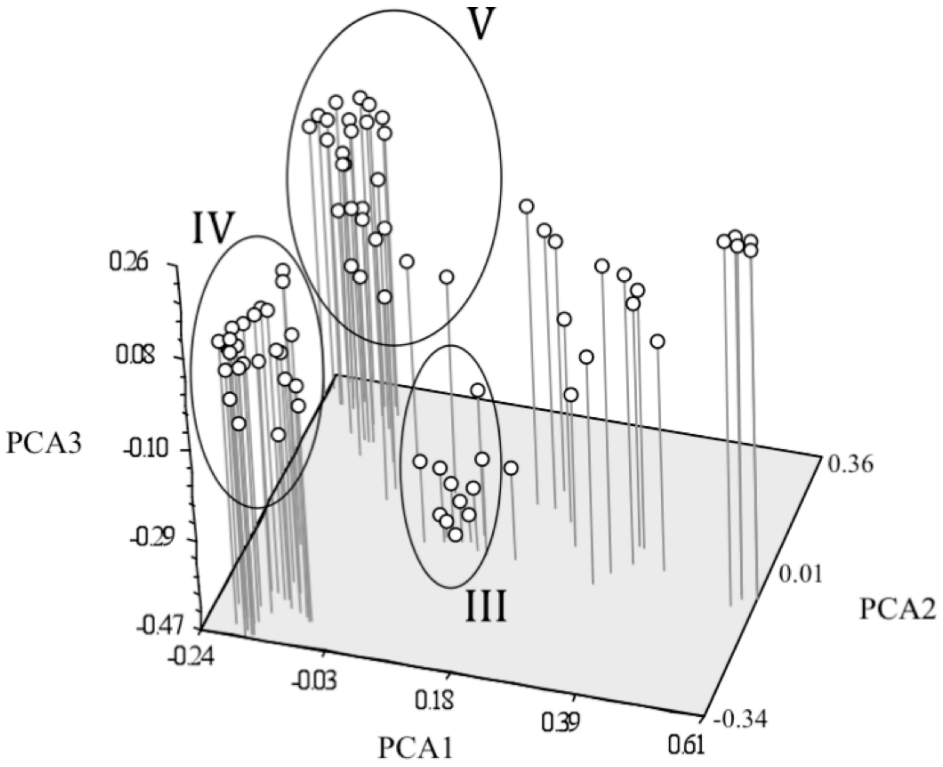


Figure 8. Principal Component Analysis of 81 *Agrostis* Accessions. The cluster analysis groups III, IV, and V representing the turf-type accessions are indicated. The remaining five groups from the cluster analysis were also preserved in this analysis.

GS of 0.49 and the *A. castellana* accessions have a GS of 0.61.

Group V consists of *Agrostis stolonifera* accessions *asto-1*, *asto*-[*3* through *10*], *asto*-[*12* through *14*], *asto-16*, *asto-19*, *asto-22*, *asto-23*, *asto-25*, *asto-29*, and *asto-30*, *asto-32*, *asto-34*, *asto-36*, *asto-40*, *A. clavata* (*acla-5*), and *A. gigantea* accessions *agig-47*, *agig-79*, and *agig-102*. Within the *A. stolonifera* accessions, the improved cultivated varieties supplied by NTEP (*asto*-[*2* through *17*]) had a GS of 0.64 and the *A. stolonifera* accessions from NPGS (*asto-19*, *asto-22*, *asto-23*, *asto-25*, *asto-29*, *asto-30*, *asto-32*, *asto-34*, asto-*36*, and *asto-40*,) had a GS of 0.57.

Group VI (GS=0.57) is composed of the accessions *Agrostis munroana* (*amun-7*), *A. nebulosa* (*aneb-1*), and *A. pallida* (*apal-1*). Groups VII and VIII are less similar to the remaining *Agrostis* accessions than are the *Polypogon* accessions from group II. Group VII (GS=0.60) consists of *A. clavata* (*acla-2*), *A. mongolica* (*amon-3* and *amon-5*), and *A. vinealis* (*atri-2* and *atri-4*). Group VIII is composed of the accessions *A. limprichtii* (*alim-1*), *A. lyallii* (*alya-1*), and *A. gigantea* (*agig-18*). Group VIII accessions are grouped together out of convenience since they are positioned near each other on the phylogenetic tree but they do not share a high degree of genetic similarity (GS=0.30).

DISCUSSION

MITE-display molecular markers can be effective for species differentiation, genetic diversity assessment, resolving species relationships, and understanding polyploidization events (Park et al., 2003). These attributes of MITE-display and its

novelty in *Agrostis* make it an attractive marker type to study diversity and species relationships within this genus.

Groups I and II identified in the UPGMA analysis consist of all non-*Agrostis* species. *Polypogon* species are morphologically similar to and are known to hybridize with *Agrostis* species (Barkworth, 2007; Harvey, 2007). *Apera interrupta* has similar morphological characters to *Agrostis* species (Allred, 2007; Harvey, 2007). Both genera were included in this study to determine how genetically similar they are to *Agrostis* species.

Based on the flow cytometry data (chapter 2), group III accessions *acan-2*, *acan-*[*6* through *8*], *acan-10*, *acan-11*, *asp-4*, *atri-5*, and *avin-1* are predicted to be diploid and *agig-77* and *atri-3* are predicted to be tetraploid. *A. vinealis* is an autotetraploid representing a genome doubling of *A. canina* (Jones, 1956a), consistent with *A. vinealis* (*atri-3*) genetic markers clustering with *A. canina* accessions. Group IV accessions *acap-*7 and *agig-1* are predicted to be diploid and may represent diploid progenitors of *A. capillaris* and *A. castellana* species.

Flow cytometry DNA content measures suggest that *acla-5* is a diploid. Clustering of MITE-display markers place *acla-5* in group V with the cultivated *A. stolonifera* accessions. *A. stolonifera* is naturally found in areas subject to periodic flooding and it is well adapted to damp soils (Harvey, 2007; Beard, 1973). *A. clavata* is also well adapted to wet soils (Harvey, 2007) and *acla-5* was collected from a wet meadow in Mongolia in 1998. Harvey (2007) reports that *A. clavata* is a hexaploid (2n=6x=42). Five *A. clavata* accessions collected from Mongolia were included in the

flow cytometry study (chapter 2) and *acla-2* and *acla-5* are predicted to be diploids and *acla-1*, *acla-3*, and *acla-4* are predicted to be hexaploids or higher ploidy; the diploid accessions were included in the present genetic diversity study. It is not clear if *A. clavata* ranges in ploidy from 2x to 6x, if the diploid accessions represent a distinct ecotype of *A. clavata*, or if the diploid accessions are misidentified. Accession *acla-2* is genotypically distinct from *acla-5* (GS=0.327) and clusters in group VII. The group V *A. clavata* accession, *acla-5*, is genotypically similar to *A. stolonifera* and it is adapted to similar environments, it is therefore hypothesized that the genome of this accession or a close ancestor contributed to one of the sub-genomes of *A. stolonifera*.

Group VI is made up of the tetraploids *amun-7* and *aneb-1* and the diploid *apal-1* and these accessions are sister to the cultivated *A. canina*, *A. capillaris*, and *A. stolonifera* accessions of groups III, IV, and V. All of the Group VII accessions were collected from Mongolia and all are predicted to be diploid based on flow cytometry data (mean 2C DNA=4.93 pg, chapter 2). A common ancestor of these groups may have contributed to the evolution of turf-type *Agrostis* species.

Clustering of MITE-display markers has implications on the evolution of various *Agrostis* species. For example, *A. canina* (2n=14) with the genome designation A_1A_1 is often thought of as contributing to one of the sub-genomes of *A. capillaris* (2n=28; $A_1A_1A_2A_2$), while the other sub-genome (A_3A_3) is of unknown origin. The origins of the *A. stolonifera* sub-genomes (A_2A_2 and A_3A_3) are unknown. Based on clustering of MITE-display markers and the flow cytometry data presented in chapter 2, *A. capillaris* (*acap*-7), *A. gigantea* (*agig-1*), *A. canina* accessions, or group VII accessions have likely

contributed to the sub-genomes of the tetraploid *A. capillaris* and the hexaploid *A. castellana* species. Similarly, *A. clavata* (*acla-5*), *A. canina* accessions, or group VII accessions may have contributed to the evolution of *A. stolonifera* accessions. Examining chromosome pairing behavior of hybrids formed between these species or identifying genome specific molecular markers might resolve the evolution of cultivated *Agrostis*. The candidate progenitor species identified in this study could be used to improve cultivated material.

5. EVOLUTION OF NPGS *AGROSTIS* ACCESSIONS BASED ON *trnL-trnF* AND *atpI-atpH* INTERGENIC SPACERS

There are between 150 and 200 species of *Agrostis* (Harvey, 2007) and the evolutionary relationships between species are not clearly understood. Deducing these relationships is further complicated by a number of factors: The major turf-type species are highly outcrossing and exhibit high rates of heterozygosity. Certain species may have several homoeologous loci resulting from polyploidization events. Homologous genes often have high levels of sequence identity and may not be informative. Intermediate morphologies exist between taxa making taxonomic classification difficult. Interspecific hybridization occurs and these reticulation events are difficult to represent in bifurcating phylograms.

Jones (1956abc) described the chromosome pairing behavior during metaphase I of meiosis between *Agrostis stolonifera*, *A. canina*, *A. capillaris*, *A. vinealis*, and *A. gigantea*. *A. vinealis* is an autotetraploidization of *A. canina* (Jones, 1956a) and these species have the genome designations $A_1A_1A_1A_1$ and A_1A_1 respectively (Brilman, 2001). The tetraploids *A. capillaris* ($A_1A_1A_2A_2$) and *A. stolonifera* ($A_2A_2A_3A_3$) share one sub-genome in common (A_2A_2) and the hexaploid *A. gigantea* ($A_1A_1A_2A_2A_3A_3$) has two sub-genomes in common with both *A. capillaris* and *A. stolonifera* (Jones, 1956bc). The

experiments conducted by Jones are among the most comprehensive describing the relationships of turf-type *Agrostis* species.

Evidence of *Agrostis* species relationships based on current molecular methodologies is lacking. Molecular markers have been used in several studies to explore *Agrostis* genetic diversity (Warnke et al., 1997; Zhao et al., 2007; Karlsen and Steiner, 2007). For example, Vergara and Bughrara (2003) examined 400 polymorphic AFLP makers in a set of 40 *Agrostis* accessions representing 14 species. They reported that *A. transcaspica* is a diploid and based on clustering of AFLP markers this species may have contributed the A_3A_3 genome to the evolution of *A. stolonifera* and *A. gigantea*. Species relationships can be partially inferred from these *Agrostis* diversity studies, but there have been no reported studies designed to explore these relationships.

Species relationships can be inferred from nuclear and plastid DNA sequences. Reichman et al. (2006) sequenced the chloroplast *matK* gene and nuclear ITS (internal transcribed spacer) region from 19 *Agrostis* and three *Polypogon* accessions. Phylogenetic analysis of the DNA sequences were used to determine if transgenic *A. stolonifera* from a controlled area was capable of pollinating wild populations of related species growing at a distance. The individual phylogenies were used to distinguish *Agrostis* accessions by species. Certain *Agrostis* species have been included in phylogenies of the Aveneae and Poaea tribes based on DNA sequence data, placing *Agrostis* species sister to *Calamagrostis* and *Polypogon* species within the Aveneae tribe (Quintanar et al., 2007; Soreng, et al., 2007; Döring et al., 2007).

Inference of *Agrostis* evolution, a highly outcrossing polyploid, based on ITS or other high copy repetitive sequence elements might not provide an accurate depiction of species evolution (Feliner and Rosselló, 2007) and single or low-copy nuclear genes might be more informative (Feliner and Rosselló, 2007; Sang, 2002). One drawback of using low copy genes is that they can be difficult to identify in non-model organisms such as *Agrostis*. Chloroplast sequences would be more informative for deducing *Agrostis* relationships since chloroplasts are maternally inherited in flowering plants and have simple genomes (Warnke, 2003).

AFLP genetic markers are sometimes used in phylogenetic studies when species relationships are difficult to resolve (Després et al., 2003; Xu and Ban, 2004). Simmons et al. (2007) discuss several concerns of using AFLP markers for phylogenetic studies including the unknown origin of markers, band co-migration, the independence of bands, and the treatment of absent bands as homologous even though absences may have arisen for different reasons. MITE-display is a modified AFLP procedure that can effectively generate molecular markers in *Agrostis* (chapters 3 and 4). Markers derived from MITE-display have been used to infer species relationships and for diversity studies (Park et al., 2003; Kavar et al., 2007). MITE-display markers are similar to AFLP markers and the concerns of using AFLP for phylogenetic studies would also be valid for MITE-display markers. The phylogeny of *Agrostis* is complex and random markers in combination with plastid markers may provide a framework for the evolution of the species.

In the present study, the evolutionary relationships of *Agrostis* will be explored based on chloroplast DNA sequence regions. Phylogenies of both cultivated and non-

cultivated NPGS *Agrostis* accessions will be compared. Chloroplast sequence based phylogenetic trees will be compared with the UPGMA cluster analysis presented in chapter 4 to identify incongruence and infer species relationships. This study will provide valuable information to *Agrostis* breeders including species relationships and sources of novel germplasm.

MATERIALS AND METHODS

Plant Material and DNA Extraction

Accessions of *Agrostis* were chosen that represent cultivated and non-cultivated accessions, various ploidy levels as predicted by flow cytometry, different species, and different geographic regions (table 8). Total DNA was extracted by either the DNeasy 96 Plant kit (Qiagen, Valencia, CA) or by a standard CTAB extraction method (Doyle and Doyle, 1987).

Computational Primer Screen

Published chloroplast specific primer sequences (Small et al., 1998; Shaw et al., 2005; Shaw et al., 2007) were formatted into a Primer Catalog (Lasergene, DNASTAR Inc., Madison, WI). The *Agrostis stolonifera* L. cv. Penn-A4 chloroplast genome (EF115543) was obtained from GenBank. The software program PrimerSelect, part of the Lasergene v 6 suite (DNASTAR Inc.), was used to search the Primer Catalog against the chloroplast genome to identify conserved primer pairs.

DNA sequencing and Phylogenetic Analysis

All sequencing reactions of *atpI-atpH* and *trnL-trnF* intergenic spacer regions

Table 8. Accessions used in phylogenetic analyses

Study ID	Accession	Taxon	Cultivar	Origin	2C DNA	Ploidy
aint-1	PI203444	*Apera interrupta*		Turkey	nd	nd
aint-2	PI204389	*Apera interrupta*		Turkey	nd	nd
pchi-1	PI372227	*Polypogon chilensis*		Uruguay	nd	nd
pvir-1	PI204395	*Polypogon viridis*		Turkey	nd	nd
pvir-2	PI223241	*Polypogon viridis*		Afghanistan	nd	nd
acan-10	PI189141	*Agrostis canina*	NOVOBENT	Netherlands	3.64	2x
acan-11	PI290707	*Agrostis canina*		United Kingdom	3.66	2x
acap-7	PI237717	*Agrostis capillaris*	ODENWALDER	Germany	5.74	2x
acap-25	PI578528	*Agrostis capillaris*	EXETER	United States	9.91	*4x
acap-26	PI600936	*Agrostis capillaris*	DUCHESS	United Kingdom	9.70	*4x
acas-11	PI240132	*Agrostis castellana*		Portugal	9.22	4x
acas-3	PI318928	*Agrostis castellana*		Spain	10.42	*4x
acas-8	PI469217	*Agrostis castellana*	HIGHLAND	United States	13.32	6x
acla-2	PI632584	*Agrostis clavata*		Mongolia	4.93	2x
acla-5	W621240	*Agrostis clavata*		Mongolia	4.54	2x
agig-77	PI251099	*Agrostis gigantea*		Yugoslavia	9.06	4x
alim-1	W623620	*Agrostis limprichtii*		China	9.64	*4x
alya-1	PI636652	*Agrostis lyallii*		New Zealand	6.86	*4x
amon-3	PI632549	*Agrostis mongolica*		Mongolia	5.20	2x
amon-5	W619706	*Agrostis mongolica*		Mongolia	5.33	2x
amun-7	PI230236	*Agrostis munroana*		Iran	9.04	4x
apal-1	PI238226	*Agrostis pallida*		Spain	6.09	*2x
asp-4	PI502279	*Agrostis sp.*		Russian	5.18	2x
asto-36	PI578529	*Agrostis stolonifera*	SEASIDE	United States	8.52	4x
asto-40	PI632687	*Agrostis stolonifera*	PENNCROSS	United States	6.02	2x
asto-19	PI204390	*Agrostis stolonifera*		Turkey	8.94	4x
asto-23	PI251945	*Agrostis stolonifera*		Austria	9.04	4x
asto-29	PI302902	*Agrostis stolonifera*		Spain	8.57	4x
asto-30	PI318934	*Agrostis stolonifera*		Spain	8.73	4x
asto-34	PI531251	*Agrostis stolonifera*	KESZTHELYI4	Hungary	7.53	*4x
atri-1	PI598462	*Agrostis trinii*		Russian	7.90	4x
atri-2	PI632559	*Agrostis trinii*		Mongolia	4.19	2x
atri-3	PI636572	*Agrostis trinii*		Mongolia	9.46	4x
atri-4	W619572	*Agrostis trinii*		Mongolia	4.98	2x
avin-1	PI440110	*Agrostis vinealis*		Russian	5.25	2x
agig-1	RUS-07-01-002	*Agrostis gigantea*		Russia	5.12	2x
agig-5	RUS-07-12-041	*Agrostis gigantea*		Russia	8.01	4x
agig-12	RUS-07-32-131	*Agrostis gigantea*		Russia	8.92	4x
asto-1	KGZ-06-08-051	*Agrostis stolonifera*		Kyrgyzstan	8.19	4x
agig-47	PRC-06-IN/118-402	*Agrostis gigantea*		China	9.84	*4x
asto-3	NT-1998-26	*Agrostis stolonifera*	Penn G-1	United States	8.93	4x
asto-4	NT-1998-27	*Agrostis stolonifera*	Penn A-4	United States	9.11	4x
asto-5	NT-1998-29	*Agrostis stolonifera*	Penn A-2	United States	8.69	4x
asto-6	NT-2003-2	*Agrostis stolonifera*	Penn A-1	United States	8.76	4x
asto-7	NT-2003-6	*Agrostis stolonifera*	Alpha	United States	8.76	4x
asto-9	NT-2003-10	*Agrostis stolonifera*	Declaration	United States	9.41	4x
asto-12	NT-2003-15	*Agrostis stolonifera*	Kingpin	United States	7.93	4x
asto-13	NT-2003-21	*Agrostis stolonifera*	007	United States	8.56	4x
acan-2	NT-1998-11	*Agrostis canina*	Vesper (Pick MVB)	United States	4.97	2x
acan-6	NT-2003-16	*Agrostis canina*	Villa(IS-AC 1)	United States	5.48	2x
acan-7	NT-2003-18	*Agrostis canina*	Venus(EFD)	United States	5.15	2x
acan-8	NT-2003-19	*Agrostis canina*	Vesper	United States	5.61	2x

Ploidy levels were predicted based on flow cytometry data. The * indicates those accessions with 2C DNA content values outside of the values defined in chapter 2 for the diploids (2x) and tetraploids (4x). Cultivated accessions are indicated by their cultivar name.

were performed using the BigDye3.1 sequencing kit (Applied Biosystems, Foster City, CA), with the primer pairs A and K specified in table 9, and on a GeneAmp2700 ThermalCycler (Applied Biosystems). Sequences were aligned with ClustalW2.0.9 (Thompson et al., 1994) and manually adjusted to further optimize alignments. Insertions and deletions in the alignment were encoded by completely omitting the gapped regions, replacing gap characters by a fifth base, treating gaps as missing data, or by the Barriel method (Barriel, 1994). The Barriel method encodes gapped regions with multistate characters to account for informative gaps in the data matrix. *Apera* accessions *aint-1* and *aint-2* were designated to the outgroup in all phylogenetic analyses.

Maximum parsimony and maximum likelihood analyses were performed using the software program PAUP v. 4.0b10 (Swofford, 2002) and Bayesian analysis was done using MrBayes v. 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Maximum parsimony analysis was performed using a heuristic search with random addition of sequences. The most likely model of evolution was determined using ModelTest3.7 (Posada and Crandall, 1998) and implemented in both maximum likelihood and Bayesian phylogenetic analyses. Bayesian analysis was run for 1,000,000 MCMC generations and sampled after every 100. The average standard deviation of split frequencies between the independent analyses for each of the chloroplast DNA regions was below 0.01, suggesting high confidence of convergence. Bootstrap analysis was run on both the maximum parsimony (10,000 replicates) and maximum likelihood (1,000 replicates) analyses. Splitstree4 was used to build consensus networks from multiple gene trees to identify discordances between phylogenies with default threshold and edge

Table 9. Conserved chloroplast primers

I	Name	Position	Sequence 5'-3'	Tm	Ref
A	*trnLtrnF*:E	47467	GGTTCAAGTCCCTCTATCCC	49.1	THI
	trnLtrnF:F	47888	ATTTGAACTGGTGACACGAG	47.1	THI
B	*trnTtrnL*:A	46123	CATTACAAATGCGATGCTCT	48	THI
	trnTtrnL:B	47003	TCTACCGATTTCGCCATATC	49.4	THI
C	rpS16R	5218	AACATCWATTGCAASGATTCGATA	55.3	THII
	rpS16F	6105	AAACGATGTGGTARAAAGCAAC	50.5	THII
D	3'trnG(UUC)	13397	GTAGCGGGAATCGAACCCGCATC	64.2	THII
	5'trnG2G	14137	GCGGGTATAGTTTAGTGGTAAAA	49.7	THII
E	trnC(GCA)R	18642	CACCCRGATTYGAACTGGGG	54.1	THII
	rpoB	19952	CKACAAAAYCCYTCRAATTG	46.7	THII
F	psbMR	17235	ATGGAAGTAAATATTCTYGCATTTATTGCT	49.7	THII
	trnD(GUC)R	16464	GGGATTGTAGYTCAATTGGT	45.5	THII
	ycf6R	17690	GCCCAAGCRAGACTTACTATATCCAT	57.3	THII
	trnC(GCA)F	18659	CCAGTTCRAATCYGGGTG	45.2	THII
	ycf6F	17715	ATGGATATAGTAAGTCTYGCTTGGGC	57.3	THII
G	trnE(UUC)	15972	AGGACATCTCTCTTTCAAGGAG	48.6	THII
	trnT(GGU)	15369	CTACCACTGAGTTAAAAGGG	43.4	THII
	trnD(GUC)F	16483	ACCAATTGAACTACAATCCC	45.5	THII
	trnY(GUA)	16164	CCGAGCTGGATTTGAACCA	53.3	THII
H	rpS4R2	45605	CTGTNAGWCCRTAATGAAAACG	49.2	THII
	trnTUGUR	46147	AGGTTAGAGCATCGCATTTG	49.7	THII
I	trnL5(UAA)F(TabC)	46991	CGAAATCGGTAGACGCTACG	52.7	THII
	3trnL(UAA)R(TabD)	47487	GGGGATAGAGGGACTTGAAC	49.1	THII
J	psbB	70842	TCCAAAAAANKKGGAGATCCAAC	54.5	THII
	psbH	71468	TCAAYRGTYTGTGTAGCCAT	50.4	THII
K	atpI	31735	TATTTACAAGYGGTATTCAAGCT	50.2	THIII
	atpH	32393	CCAAYCCAGCAGCAATAAC	48	THIII
L	ndhC	50106	TATTATTAGAAATGYCCARAAAATATCATATTC	53.8	THIII
	trnV(UAC)x2	51106	GTCTACGGTTCGARTCCGTA	49.3	THIII
M	psbE	62673	TATCGAATACTGGTAATAATATCAGC	46.6	THIII
	petL	64048	AGTAGAAAACCGAAATAACTAGTTA	45.9	THIII
N	rpL32-F	105950	CAGTTCCAAAAAAACGTACTTC	48.2	THIII
	trnL(UAG)	106833	CTGCTTCCTAAGAGCAGCGT	51.7	THIII
O	rpS16x1	6084	GTTGCTTTYTACCACATCGTTT	50.5	THIII
	trnQ(UUG)	6978	GCGTGGCCAAGYGGTAAGGC	59.1	THIII
P	TabE	47467	GGTTCAAGTCCCTCTATCCC	49.1	THIII
	ndhJ	48565	ATGCCYGAAAGTTGGATAGG	48.5	THIII
Q	trnT(GGU)-R	15350	CCCTTTTAACTCAGTGGTAG	43.4	THIII
	psbD	9263	CTCCGTARCCAGTCATCCATA	49.2	THIII

Position – Primer location on the *Agrostis stolonifera* Penn A-4 chloroplast genome. Ref – References THI (Small et al., 1998), THII (Shaw et al., 2005), and THIII (Shaw et al., 2007).

weight parameters. The splits graph was drawn as a rectangular cladogram using the cluster network splits transformation and choosing *aint-1* as an outgroup

RESULTS

The software program DNASTAR Lasergene v 6 was used to identify conserved chloroplast specific primer pairs reported by Small et al. (1998) and Shaw et al. (2005 and 2007) in an *Agrostis stolonifera* cv. Penn-A4 chloroplast genome (GenBank Accession EF115543). Seventeen conserved primer pairs were identified and their relative bp position on the chloroplast genome was determined (table 9). A subset of conserved primer pairs were screened on a diverse panel of *Agrostis* DNA and successfully amplified products sequenced. The primer pairs amplifying the *trnL-trnF* and *adhI-adhH* intergenic spacer regions of the chloroplast genome amplified well in all accessions tried and were subsequently sequenced from the 52 accessions composed of one *Apera* accession, three *Polypogon* accessions, and 48 *Agrostis* accessions.

The alignment of a 298 bp conserved region of the *trnL-trnF* intergenic spacer region (Appendix C) was used in the phylogenetic analyses. This region contains two gaps, 6 variable parsimony uninformative characters, 260 constant characters, and 32 parsimony informative characters (10.7% PIC). Thirteen unique haplotypes were identified in the set of 52 accessions (table 10). The 13 *trnL-trnF* haplotypes preserved the eight groups described in chapter 4 based on 1,309 MITE-display markers with the exception of groups III, IV, V, and VI. Similarly the alignment of the *atpI-atpH* intergenic spacer region (Appendix D) had 451 total characters, 12 gaps, 10 variable but

Table 10. Unique *trnL-trnF* intergenic spacer haplotypes

Representative Group ID	Accessions with identical haplotypes	MITE Display Grouping
aint-1	*aint-1*, *aint-2*	I
pchi-1	*pchi-1*	II
pvir-1	*pvir-1*, *pvir-2*	II
acan-10	*acan-10*, *acan-11*, *agig-1**, *acan-2*, *acan-6*, *acan-7*, *acan-8*	III, IV*
acap-7	*acap-7*, *acap-25*, *acap-26*, *acas-11*, *acas-3*, *acas-8*, *apal-1***, *asto-29**, *asto-30**, *atri-1*, *agig-5*, *agig-12*	IV, V*, VI**
acla-2	*acla-2*, *amon-3*, *amon-5*, *atri-2*, *atri-4*	VII
acla-5	*acla-5**, *agig-77*, *asp-4*, *atri-3*, *avin-1*, *asto-1**, *agig-47**	III, V*
alim-1	*alim-1*	VIII
alya-1	*alya-1*	VIII
amun-7	*amun-7*	VI
asto-36	*asto-36*, *asto-40*, *asto-19*, *asto-34*, *asto-3*, *asto-4*, *asto-5*, *asto-6*, *asto-7*, *asto-9*, *asto-13*	V
asto-23	*asto-23*	V
asto-12	*asto-12*	V

Table 11. Unique *atpI-atpH* intergenic spacer haplotypes

Representative Group ID	Accessions with identical haplotypes	MITE Display Grouping
aint-1	*aint-1*, *aint-2*	I
pchi-1	*pchi-1*	II
pvir-1	*pvir-1*, *pvir-2*	II
acan-10	*acan-10*, *acan-11*, *acla-5***, *agig-77*, *asp-4*, *asto-40***, *asto-23***, *asto-34***, *atri-3*, *avin-1*, *agig-1**, *asto-1***, *agig-47***, *asto-4***, *asto-5***, *asto-6***, *asto-7***, *asto-12***, *asto-13***, *acan-2*, *acan-6*, *acan-7*, *acan-8*	III, IV*, V**
acap-7	*acap-7*, *acap-25*, *acap-26*, *acas-11*, *acas-8*, *apal-1***, *asto-29**, *asto-30**, *atri-1*, *agig-5*, *agig-12*	IV, V*, VI**
acas-3	*acas-3*	IV
acla-2	*acla-2*, *amon-3*, *amon-5*, *atri-2*, *atri-4*	VII
alim-1	*alim-1*	VIII
alya-1	*alya-1*	VIII
amun-7	*amun-7*	VI
asto-36	*asto-36*, *asto-19*, *asto-3*, *asto-9*	V

uninformative characters, 393 constant characters, and 48 parsimony informative characters (10.6% PIC). There were 11 distinct haplotypes (table 11) of the *atpI-atpH* intergenic spacer region, which also closely preserved the groupings described by the clustering of MITE-display markers with the exception of groups III, IV, V, and VI.

A phylogenetic analysis of the *atpI-atpH* and *trnL-trnF* intergenic spacer regions was performed separately on the cultivated *Agrostis* accessions from table 8 along with the *Polypogon* and *Apera* accessions or on the set of 52 accessions listed in table 8. Of the four gap-encoding methods, complete omission of the gapped regions reduced the number of *atpI-atpH* haplotypes from 11 to 9 and *trnL-trnF* haplotypes from 13 to 11, and this method was not analyzed further due to the reduction in data. The other three methods, new state, missing, and Barriel method, resulted in similar phylogenies so the Barriel method was chosen because it effectively represents informative gaps. The predicted model of evolution for both the *trnL-trnF* and *atpI-atpH* regions was a general time reversible (GTR) model with equal base frequencies.

The single best tree of the *trnL-trnF* region from the maximum parsimony analysis of the set of 52 accessions had length 45, consistency index of 0.9778, and homoplasy index of 0.0222 (figure 9). The phylogenies of the *atpI-atpH* region are presented in figure 10 and the relationships between haplotypes are similar to those observed in the analyses of the *trnL-trnF* region. The best *atpI-atpH* phylogenetic tree found by maximum parsimony had a length of 79, consistency index of 0.9367, and homoplasy index of 0.0633. All three phylogenetic analysis methods, maximum parsimony, maximum likelihood, and Bayesian inference, gave topologically similar trees

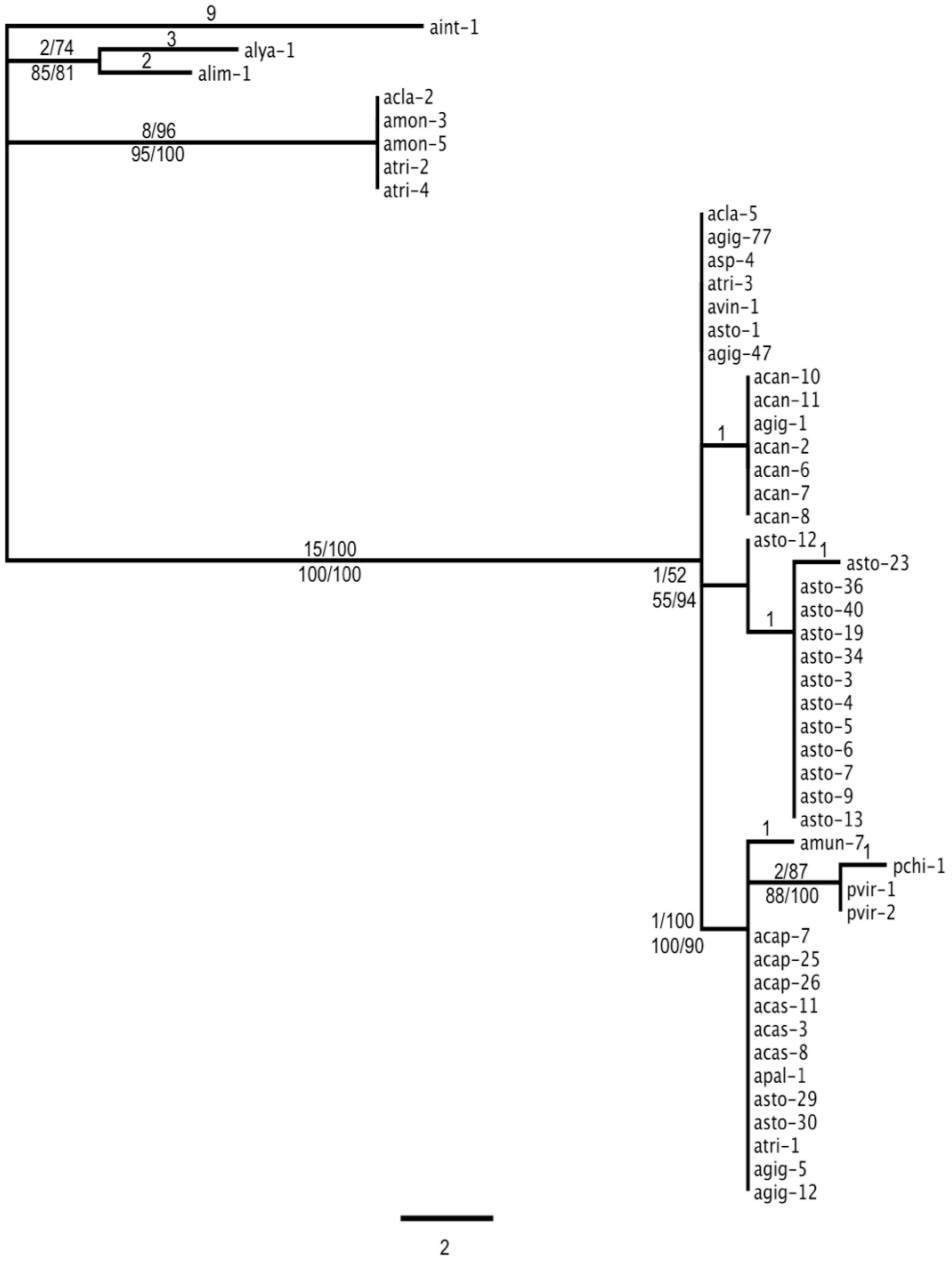


Figure 9. Phylogenetic analysis of *trnL-trnF* intergenic spacer. The most parsimonious tree from the maximum parimony analysis. The number of changes is shown above the branches along with the bootstrap support values from the parsimony analysis (changes/bs). The maximum likelihood bootstrap support values and the clade support values from the Bayesian phylogenetic analysis are indicated below the branches as (bs/bpp).

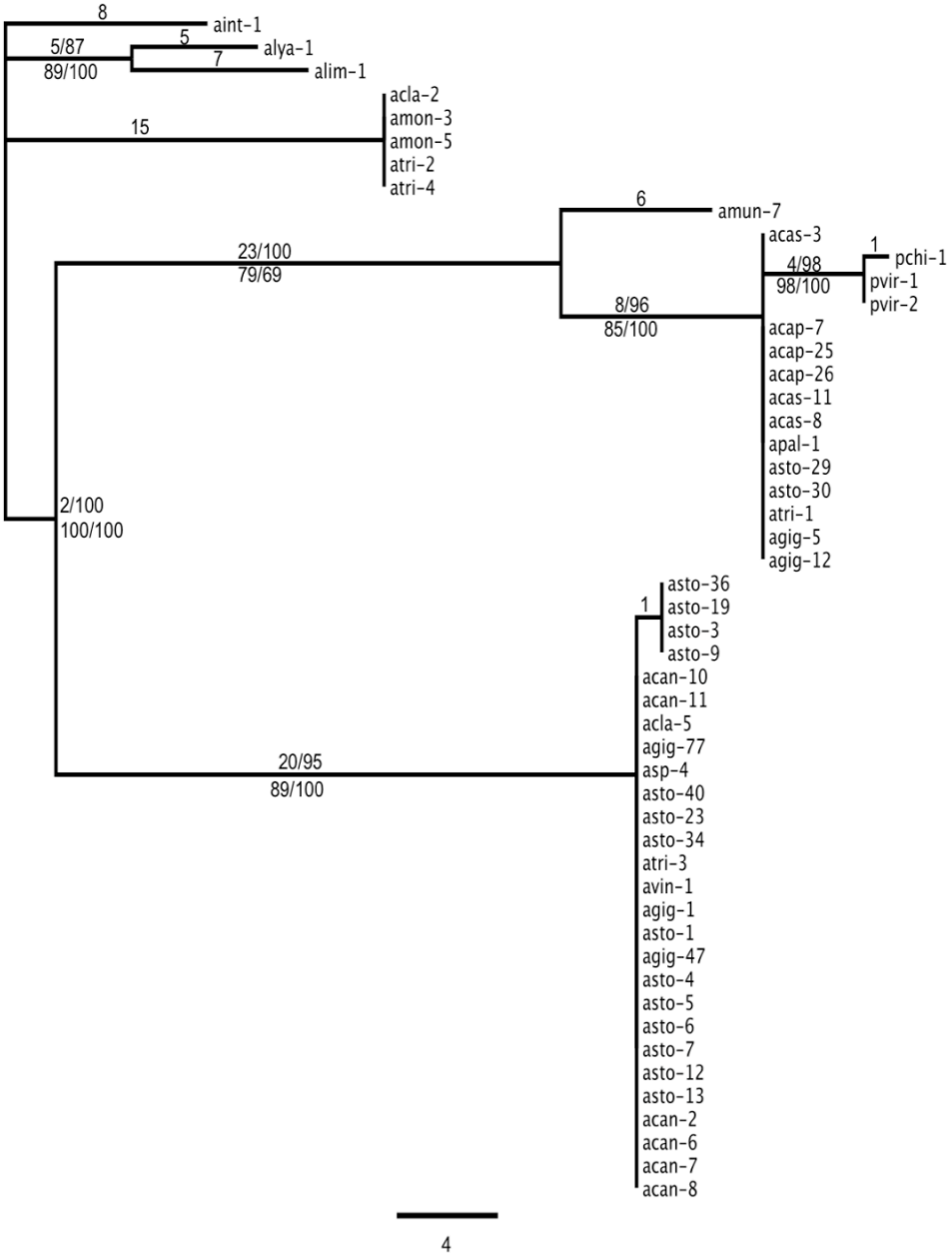


Figure 10. Phylogenetic analysis of *atpI-atpH* intergenic spacer. The most parsimonious tree from the maximum parimony analysis. The number of changes is shown above the branches along with the bootstrap support values from the parsimony analysis (changes/bs). The maximum likelihood bootstrap support values and the clade support values from the Bayesian phylogenetic analysis are indicated below the branches as (bs/bpp).

for both intergenic spacer regions.

A maximum parsimony analysis of the *trnL-trnF* region from cultivated *Agrostis* accessions (table 8) had a single best tree of length 30, consistency index of 1.0, and homoplasy index of 0.0 (figure 11). The maximum parsimony phylogeny of the *atpI-atpH* region had a length of 42, consistency index of 1.0, and homoplasy index of 0.0 (figure 12). Maximum parsimony, maximum likelihood, and Bayesian inference, gave topologically similar trees for both intergenic spacer regions. Both the *trnL-trnF* and *atpI-atpH* intergenic spacer regions divided the accessions into 5 distinct clades. The *Polypogon* accessions formed a single clade that is closely related to the clade representing the *A. capillaris* and *A. castellana* accessions. The remaining *A. stolonifera* and *A. canina* accessions formed two more groups. The *atpI-atpH* intergenic spacer region was not successful at distinguishing the *A. canina* accessions from the *A. stolonifera* accessions. The *A. stolonifera* accessions *asto-3, asto-9,* and *asto-36* formed a distinct subgroup within the *A. stolonifera* and *A. canina* accessions. The *trnL-trnF* spacer region was able to distinguish the *A. canina* accessions from the *A. stolonifera* accessions forming two distinct clades. Phylogenetic analysis of both of the intergenic spacer regions closely supported the groups found by clustering of MITE-display markers (chapter 4).

A Splitstree4 consensus network of the chloroplast encoded intergenic spacer regions showed one disagreement; the placement of *amun-7* was inconsistent between the individual sequence phylogenies (figure 13). One difference between the individual chloroplast DNA sequence analyses of the placement of *amun-7* is in its relationship to

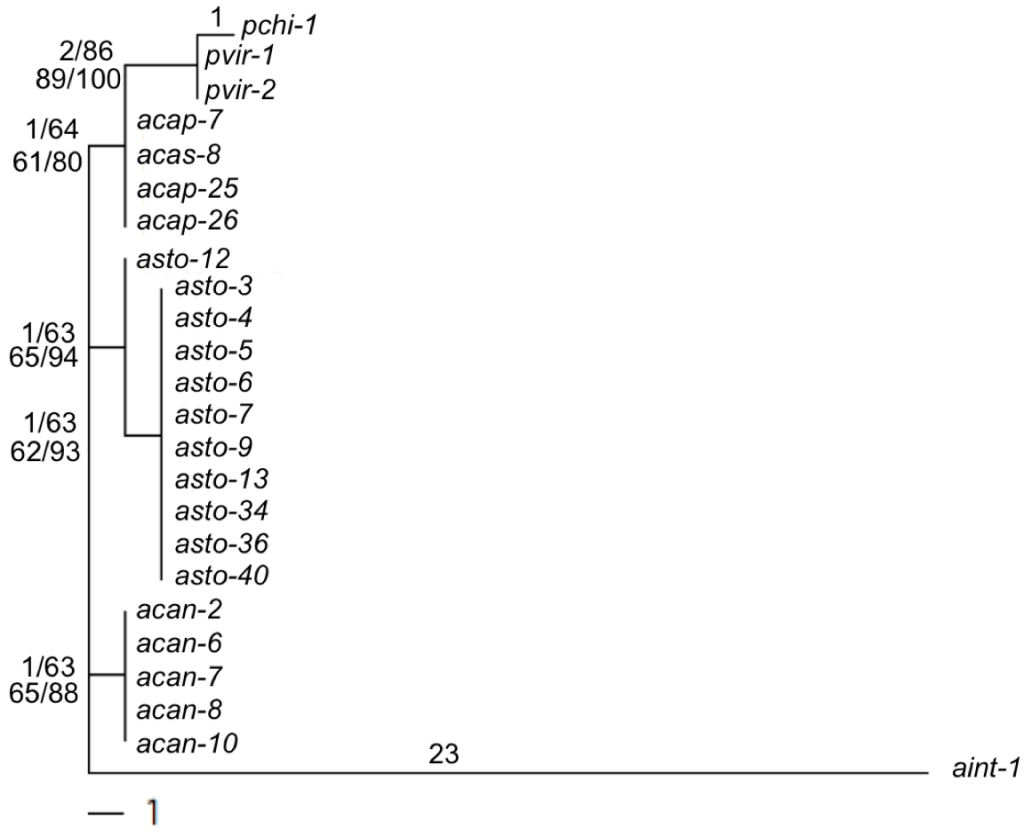


Figure 11. Phylogenetic analysis of *trnL-trnF* intergenic spacer from cultivated *Agrostis*. The most parsimonious tree from the maximum parimony analysis. The number of changes is shown above the branches along with the bootstrap support values from the parsimony analysis (changes/bs). The maximum likelihood bootstrap support values and the clade support values from the Bayesian phylogenetic analysis are indicated below the branches as (bs/bpp).

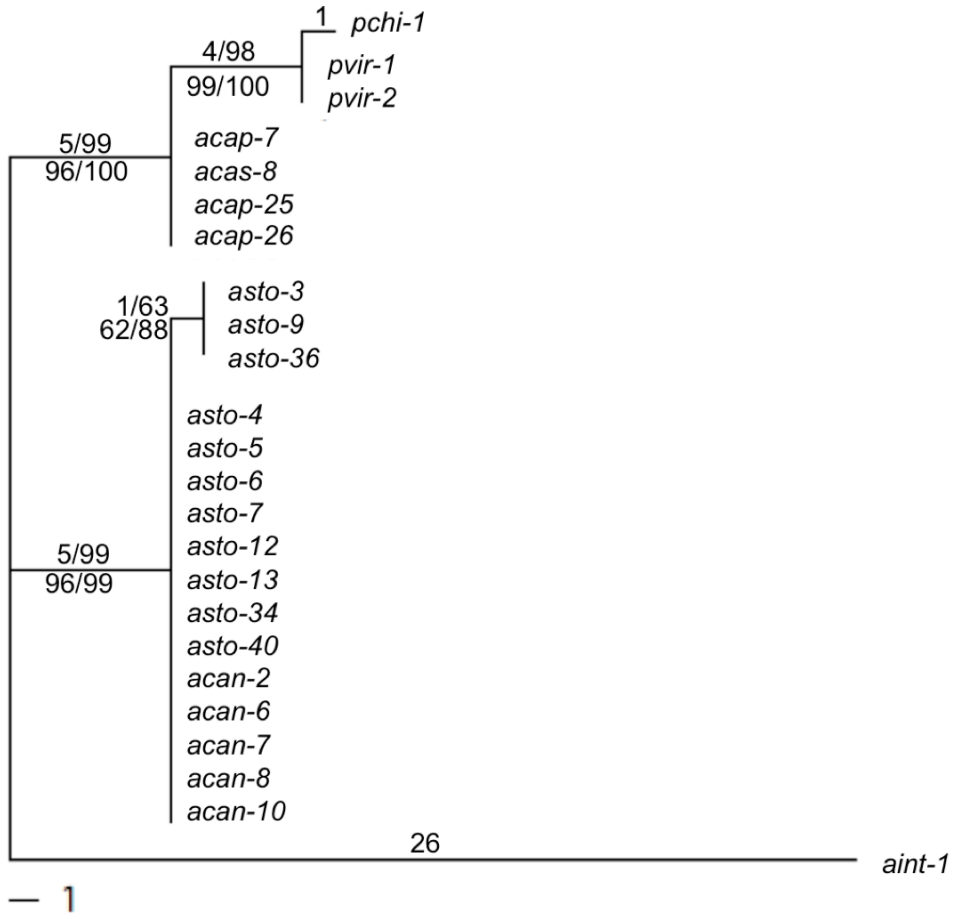
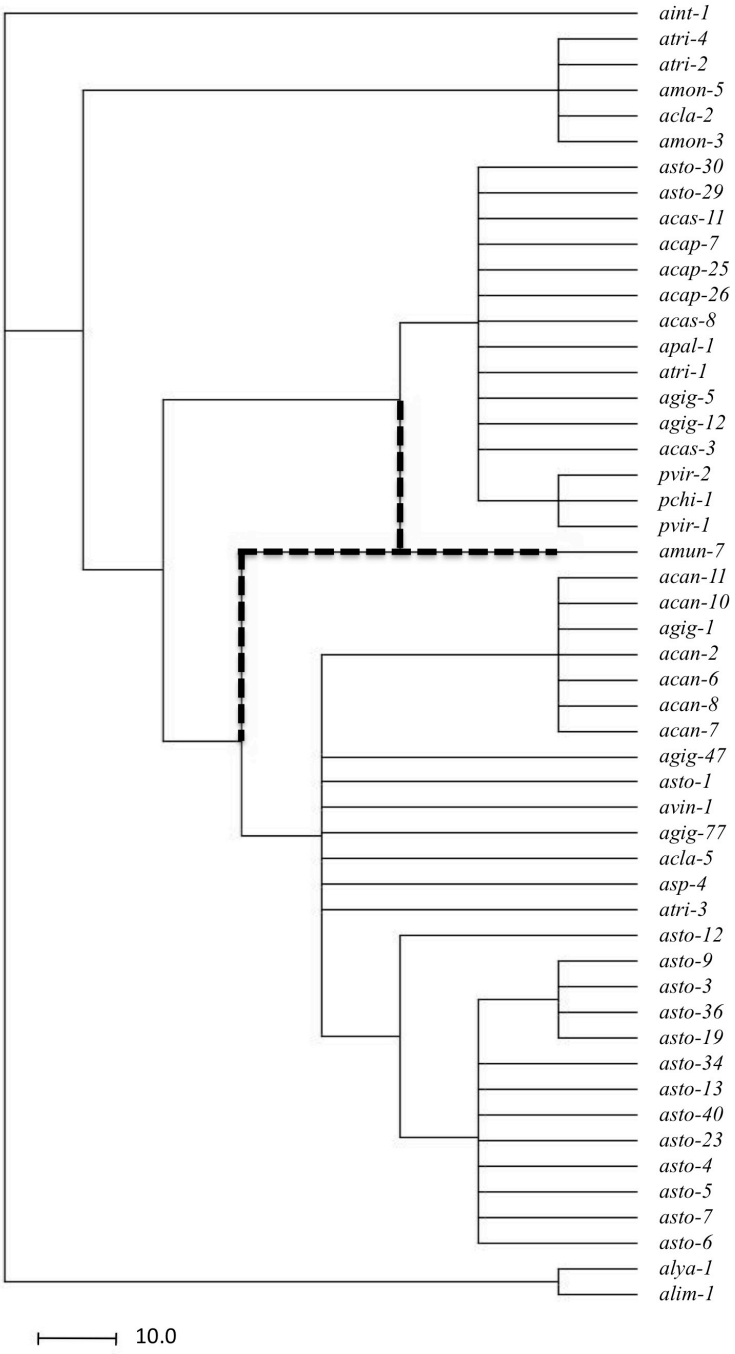


Figure 12. Phylogenetic analysis of *atpI-atpH* intergenic spacer from cultivated *Agrostis*. The most parsimonious tree from the maximum parimony analysis. The number of changes is shown above the branches along with the bootstrap support values from the parsimony analysis (changes/bs). The maximum likelihood bootstrap support values and the clade support values from the Bayesian phylogenetic analysis are indicated below the branches as (bs/bpp).

Figure 13. Consensus network of the *atpI-atpH* and *trnL-trnF* intergenic spacer regions. Splits are indicated with the bold dashed line.



the *A.* stolonifera and *A. canina* accessions. The analysis of the *trnL-trnF* region supports a closer relationship of *amun-7* to the *A. stolonifera* and *A. canina* accessions compared to the same relationships from the *atpI-atpH* phylogeny.

A consensus network of *atpI-atpH*, *trnL-trnF*, and MITE-display phylogenies was generated by SplitsTree4 to test the agreement between the individual chloroplast DNA sequence phylogenies with the MITE-display cluster analysis from chapter 4 (figure 14). This consensus network is complex and shows several splits between the chloroplast sequences and the MITE-display marker derived phylogenies. The MITE display cluster analysis groupings are well preserved. For example, all of the members of group VII coalesce in the consensus network graph. Groups I, II, III, and VI are also preserved. The MITE display cluster data suggest that *agig-1* is a member of group IV, while the *trnL-trnF* and *atpI-atpH* phylogenies place *agig-1* with the *A. canina* accessions. Groups V and VI have a high number of splits compared to the other accession groups. Several splits are also observed between the diploid accessions of groups III and VI and the tetraploid cultivated *Agrostis* accessions of groups IV and V.

DISCUSSION

Polyploidization events, origins of turf-type species, sources of novel germplasm, and targets for developing interspecific hybrids may be inferred from the evolution of *Agrostis* species. Chloroplast DNA sequence regions have been used to resolve phylogenies of several grass species (Fortune et al., 2008; Ge et al., 2002; Gillespie et al., 2006; Guo and Ge, 2005). There were only a few distinct haplotypes of the *trnL-trnF*

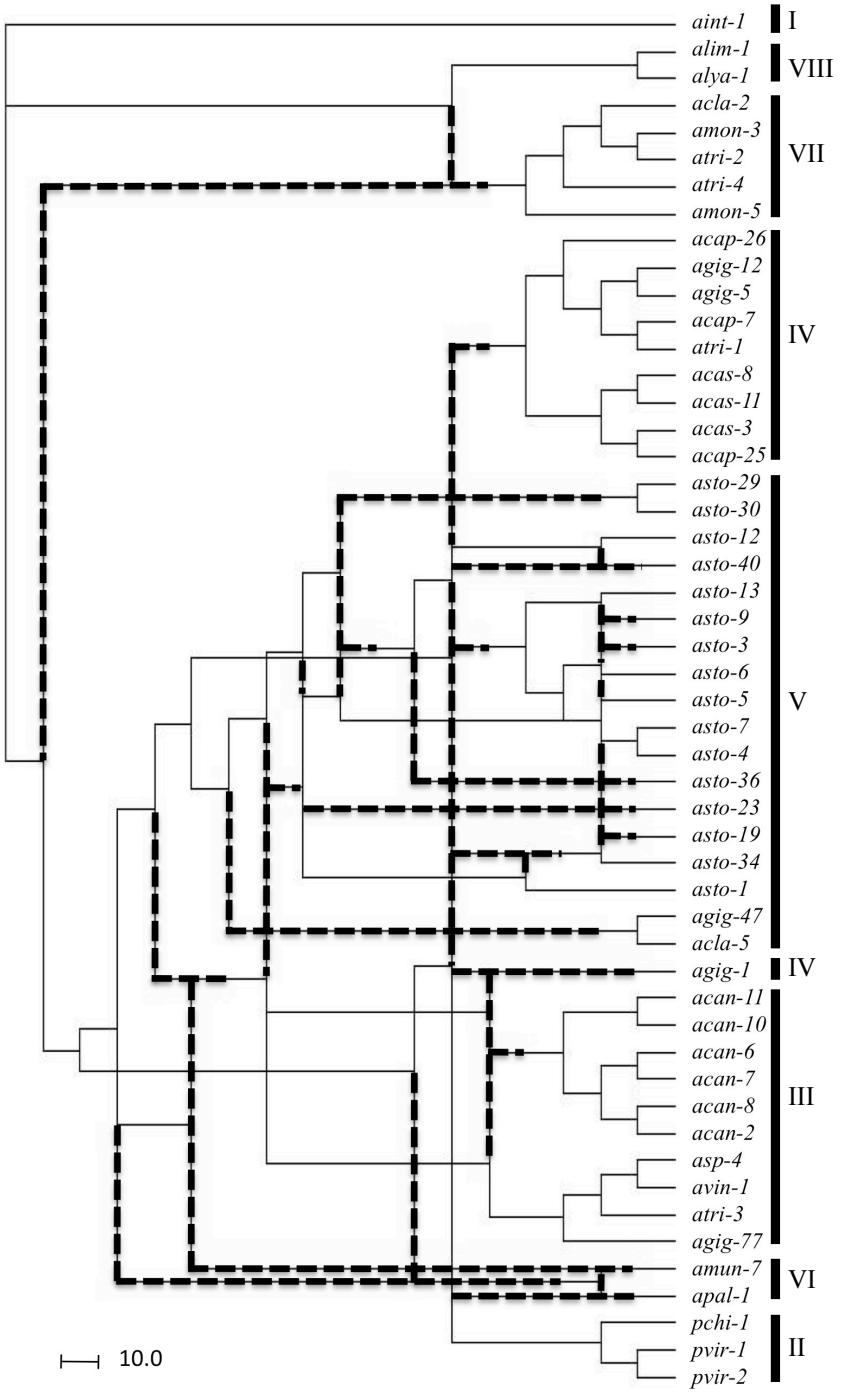


Figure 14. Consensus network of the intergenic spacer regions and MITE-display clusters. Splits are shown with the bold dashed line and the MITE-display groups are labeled and shown with the bold solid bar.

and *atpI-atpH* intergenic spacer regions within this set of *Agrostis* and therefore these chloroplast regions might not provide a detailed view of evolutionary processes in *Agrostis*.

The lack of cohesion of species in the individual chloroplast sequence based phylogenies calls into question the ability to resolve species relationships using these chloroplast encoded DNA sequences. For example *asto-1, asto-29, and asto-30* do not fall into or near the same clade as the rest of the *A. stolonifera* accessions. Similarly the four *A. trinii* accession and the five *A. gigantea* accessions are distributed throughout the phylogenetic tree and do not appear to be closely related. There are several possible reasons for this inability to distinguish species based on chloroplast sequence data, including incorrectly identifying and naming of the National Plant Germplasm System *Agrostis* accessions, multiple origins of a single species contributing to different maternal chloroplast ancestral genomes, and the possibility of paternally inherited chloroplast genomes in *Agrostis*.

The individual phylogenies of the *atpI-atpH* and *trnL-trnF* intergenic spacer regions support the groupings described by the MITE-display cluster analysis of chapter 4. The MITE-display group III is composed primarily of *Agrostis canina* accessions, group IV has both *A. capillaris* and *A. castellana* accessions, group V is made up of the *A. stolonifera* accessions, and group VI is composed of *A. munroana* (*amun-7*) and *A. pallida* (*apal-1*). Interspecific hybridization is known to occur between the species in groups III, IV, and V, which might make it difficult to clearly resolve the relationships between the species in these groups. *A. pallida* is predicted to be diploid and has

identical intergenic spacer haplotypes with the *A. capillaris* and *A. castellana* accessions (group IV). The *apal-1* accession did not cluster with group IV based on MITE-display marker clustering. It is therefore hypothesized that a common ancestor of group IV and *apal-1* is the maternal progenitor of the group IV polyploids.

The phylogenetic analysis of the *atpI-atpH* and *trnL-trnF* intergenic spacer regions of cultivated *Agrostis* accessions included four *Agrostis* species along with the *Apera* and *Polypogon* accessions. It is interesting that the *Polypogon* accessions are closely related to the *A. capillaris* accessions. Since *Polypogon* species are known to form intergeneric hybrids with *Agrostis* species, these species may have evolved from a common ancestor. Furthermore it is believed that the genome of the diploid *A. canina* contributed to the evolution of *A. capillaris* but not *A. stolonifera.* The data presented by these chloroplast phylogenies suggest that *A. canina* is more closely related to *A. stolonifera* than *A. capillaris.* The DNA sequence based phylogenies presented here do not support the findings of Jones (1956 bc). Further experiments examining other chloroplast genome regions or nuclear DNA may help to further resolve species relationships in *Agrostis.*

Interspecific hybridization is known to occur between species in groups III, IV, and V and these reticulation events are one possible cause of incongruence when comparing the individual gene trees to the MITE display groups. There are several disagreements in the consensus network of group V *A. stolonifera* accessions (figure 14). The relatively high number of disagreements might be explained by the interbreeding that has taken place within this group since the majority of these accessions represent

cultivated germplasm. Groups I, II, III, IV, VI, VII, and VIII do not exhibit many within-group splits. These groups may provide novel genes that could be used to further improve cultivated *A. stolonifera*.

The consensus network presented here was derived from chloroplast sequence data along with MITE-display marker data. Phylogenies constructed from a larger and more diverse collection of *Agrostis* or constructed from other chloroplast regions or low-copy nuclear genes may provide a more detailed picture of *Agrostis* evolution.

6. CONCLUSIONS

Agrostis stolonifera is a highly outcrossing allotetraploid species and it is the primary species used in high quality turf stands such as golf course greens, tees, and fairways. There are between 150 and 200 species of *Agrostis* and evolutionary relationships between species and related genera are not clearly understood. Resolving these relationships is important to turfgrass breeders because they can incorporate related material directly into breeding programs, recreate evolutionary events that led to the high quality turf-type *Agrostis*, or introduce novel stress tolerance genes through wide cross breeding. *Agrostis* species represent a polyploid series, which complicates phylogenetic inference so it is important to determine the ploidy level of available germplasm.

Flow cytometry proved to be an efficient tool for measuring DNA content and predicting ploidy. The 2C DNA content from known cultivated diploid germplasm (mean 5.29 pg) was used to identify candidate diploids from the NPGS germplasm collection. Accessions predicted to be diploid are *A. canina* (*acan-10* and *acan-11*), *A. capillaris* (*acap-7*), *A. clavata* (*acla-2* and *acla-5*), *A. gigantea* (*agig-1*), *A. mongolica* (*amon-3* and *amon-5*), species not specified accession (*asp-4*), *A. stolonifera* (*asto-40*), *A. trinii* (*atri-2*, *atri-4*, and *atri-5*), and *A. vinealis* (*avin-1*). Similarly several NPGS accessions were predicted to be tetraploid (table 2). The flow cytometry data suggests that some of the germplasm has been misclassified but also provides new diploid and

tetraploid germplasm to examine for breeding potential. The majority of these accessions were included in subsequent analyses to determine their relationships to NTEP cultivated germplasm.

Molecular markers anchored to MITE sequences are useful to study diversity and for the development of markers that could be used in marker assisted selection since, unlike AFLP markers, they are often found associated with genes. The data presented in chapter 3 is the first report of MITEs in *Agrostis* species and the first report of their use as a molecular marker system in a turfgrass species. A total of 16,122 *Agrostis* DNA sequences were searched for sequence containing MITE-like characteristics. The software program FindMITE identified 495 candidate MITEs. The software program primer3 was very successful at designing both MIP (MITE insertional polymorphism) and MITE-display PCR primers. Of the 157 designed MIP primer pairs, 146 amplified a product and 79 were found to exhibit banding profiles consistent with MITE transposition activity. Since some upfront sequence information is needed to design effective MITE-based molecular markers, this molecular marker technique is not as easy to implement as AFLP, but the potential for developing gene-based markers outweighs this complication.

In chapter 4, 1,309 MITE-display markers were used in a UPGMA cluster analysis to assay the genetic diversity within a set of 75 *Agrostis*, 2 *Apera*, and 4 *Polypogon* accessions. The clustering of MITE-display markers showed that the *A. stolonifera* cultivated NTEP accessions are highly similar which may in-turn suggest a narrowing of this gene pool. MITE-display marker clustering and ploidy predictions suggest that *A. clavata* (*acla-5*) is a likely progenitor of *A. stolonifera* and *A. capillaris*

(*acap-7*) and/or *A. gigantea* (*agig-1*) are potential diploid progenitors of *A. capillaris* and *A. castellana* species. These data also confirm the close relationship of *Polypogon* species to *Agrostis*. It is interesting that MITE-display diversity group VII is composed of all diploid germplasm originating from Mongolia. This region may be the center of origin for *Agrostis* or at least a source of novel diploid germplasm.

The DNA sequencing of *atpI-atpH* and *trnL-trnF* intergenic spacer regions revealed only a few unique haplotypes. Sequencing more chloroplast DNA from larger chloroplast genome regions and from a more diverse set of *Agrostis* germplasm may provide a more detailed picture of chloroplast evolution in the genus. Several reticulation events within the set of 52 accessions examined were observed in the consensus network of the combined chloroplast intergenic spacer regions with the MITE-display marker data. The number of reticulation events was very high within the set of *A. stolonifera* accessions, which might suggest the formation of several hybrids within that group. The consensus network presented in figure 14 demonstrates the complexity of species relationships in *Agrostis*. It would be interesting to develop a network phylogeny from low copy nuclear genes combined with chloroplast sequence data to determine if that provides a more resolved phylogeny of *Agrostis*.

APPENDIX A. LIST OF GERMPLASM AND 2C DNA CONTENT

Study ID	Accession	Taxon	2C DNA (pg)	Source
aint-1	PI203444	*Apera interrupta*	nd	NPGS
aint-2	PI204389	*Apera interrupta*	nd	NPGS
pchi-1	PI372227	*Polypogon chilensis*	nd	NPGS
pmon-1	PI317467	*Polypogon monspeliensis*	nd	NPGS
pvir-1	PI204395	*Polypogon viridis*	nd	NPGS
pvir-2	PI223241	*Polypogon viridis*	nd	NPGS
acan-10	PI189141	*Agrostis canina*	3.64±0.18	NPGS
acan-11	PI290707	*Agrostis canina*	3.66±0.13	NPGS
acap-1	PI171470	*Agrostis capillaris*	8.06±0.87	NPGS
acap-2	PI172698	*Agrostis capillaris*	8.39±0.36	NPGS
acap-3	PI204397	*Agrostis capillaris*	7.38±0.54	NPGS
acap-4	PI206626	*Agrostis capillaris*	7.41±0.66	NPGS
acap-5	PI234685	*Agrostis capillaris*	7.56±2.17	NPGS
acap-6	PI235217	*Agrostis capillaris*	7.26±0.55	NPGS
acap-7	PI237717	*Agrostis capillaris*	5.74±1.04	NPGS
acap-8	PI252045	*Agrostis capillaris*	6.77±0.78	NPGS
acap-9	PI283173	*Agrostis capillaris*	8.13±1.21	NPGS
acap-10	PI290708	*Agrostis capillaris*	9.85±0.57	NPGS
acap-11	PI311011	*Agrostis capillaris*	7.26±1.07	NPGS
acap-12	PI325192	*Agrostis capillaris*	6.26±5.51	NPGS
acap-13	PI325194	*Agrostis capillaris*	8.72±1.06	NPGS
acap-14	PI392338	*Agrostis capillaris*	9.22±0.01	NPGS
acap-15	PI420235	*Agrostis capillaris*	9.56±1.27	NPGS
acap-16	PI420236	*Agrostis capillaris*	9.66±0.91	NPGS
acap-17	PI440109	*Agrostis capillaris*	7.07±1.73	NPGS
acap-18	PI478595	*Agrostis capillaris*	9.94±1.35	NPGS
acap-19	PI491264	*Agrostis capillaris*	7.68±1.81	NPGS
acap-20	PI494120	*Agrostis capillaris*	11.02±1.04	NPGS
acap-21	PI494121	*Agrostis capillaris*	9.25±3.26	NPGS
acap-22	PI509437	*Agrostis capillaris*	11.01±0.47	NPGS
acap-23	PI538785	*Agrostis capillaris*	13.94±1.48	NPGS
acap-24	PI578527	*Agrostis capillaris*	9.42±0.53	NPGS
acap-25	PI578528	*Agrostis capillaris*	9.91±0.81	NPGS
acap-26	PI600936	*Agrostis capillaris*	9.7±0.31	NPGS
acap-27	W618294	*Agrostis capillaris*	14.04±0.34	NPGS

Study ID	Accession	Taxon	2C DNA (pg)	Source
acas-9	PI210428	*Agrostis castellana*	12.41±1.56	NPGS
acas-10	PI240131	*Agrostis castellana*	12.56±0.62	NPGS
acas-11	PI240132	*Agrostis castellana*	9.22±0.81	NPGS
acas-12	PI240133	*Agrostis castellana*	13.49±1.02	NPGS
acas-13	PI240134	*Agrostis castellana*	14.64±0.55	NPGS
acas-14	PI240135	*Agrostis castellana*	13.45±1.3	NPGS
acas-15	PI240136	*Agrostis castellana*	13.46±1.23	NPGS
acas-16	PI240137	*Agrostis castellana*	15.7±1.83	NPGS
acas-17	PI240138	*Agrostis castellana*	15±0.66	NPGS
acas-18	PI240139	*Agrostis castellana*	15.3±1.32	NPGS
acas-19	PI240140	*Agrostis castellana*	14.2±0.6	NPGS
acas-20	PI240141	*Agrostis castellana*	10.34±0.58	NPGS
acas-21	PI240142	*Agrostis castellana*	16±0.78	NPGS
acas-22	PI240143	*Agrostis castellana*	13.34±0.5	NPGS
acas-23	PI240144	*Agrostis castellana*	11.96±0.76	NPGS
acas-24	PI240145	*Agrostis castellana*	12.62±1.62	NPGS
acas-25	PI287740	*Agrostis castellana*	15.67±1.27	NPGS
acas-26	PI287741	*Agrostis castellana*	13.47±1.37	NPGS
acas-27	PI287742	*Agrostis castellana*	13.86±0.6	NPGS
acas-28	PI287743	*Agrostis castellana*	13.56±1.46	NPGS
acas-29	PI287744	*Agrostis castellana*	13.76±0.23	NPGS
acas-30	PI287745	*Agrostis castellana*	15.55±0.8	NPGS
acas-31	PI289644	*Agrostis castellana*	14.1±1.53	NPGS
acas-32	PI302830	*Agrostis castellana*	13.6±0.92	NPGS
acas-33	PI318924	*Agrostis castellana*	15.28±0.94	NPGS
acas-34	PI318925	*Agrostis castellana*	15.99±0.96	NPGS
acas-1	PI318926	*Agrostis castellana*	15.17±0.93	NPGS
acas-2	PI318927	*Agrostis castellana*	14.6±0.94	NPGS
acas-3	PI318928	*Agrostis castellana*	10.42±0.56	NPGS
acas-4	PI318929	*Agrostis castellana*	14.31±0.32	NPGS
acas-5	PI318930	*Agrostis castellana*	12.03±0.44	NPGS
acas-6	PI318931	*Agrostis castellana*	14.17±1.87	NPGS
acas-7	PI318932	*Agrostis castellana*	12.82±0.54	NPGS
acas-8	PI469217	*Agrostis castellana*	13.32±1.48	NPGS
acla-1	PI632483	*Agrostis clavata*	18.46±0.13	NPGS
acla-2	PI632584	*Agrostis clavata*	4.93±0.18	NPGS
acla-3	W619681	*Agrostis clavata*	16.31±0.6	NPGS
acla-4	W619776	*Agrostis clavata*	12.59±1.11	NPGS
acla-5	W621240	*Agrostis clavata*	4.54±0.35	NPGS
aexa-1	W631008	*Agrostis exarata*	18.78±0.75	NPGS

Study ID	Accession	Taxon	2C DNA (pg)	Source
agig-48	PI199244	*Agrostis gigantea*	11.27±1.1	NPGS
agig-49	PI204391	*Agrostis gigantea*	11.57±0.52	NPGS
agig-50	PI204393	*Agrostis gigantea*	13.53±0.52	NPGS
agig-51	PI204394	*Agrostis gigantea*	11.56±0.41	NPGS
agig-52	PI206627	*Agrostis gigantea*	11.83±1.26	NPGS
agig-53	PI206628	*Agrostis gigantea*	12.45±0.49	NPGS
agig-54	PI206630	*Agrostis gigantea*	12.36±0.48	NPGS
agig-55	PI206631	*Agrostis gigantea*	12.15±0.2	NPGS
agig-56	PI206879	*Agrostis gigantea*	12.01±0.57	NPGS
agig-57	PI206880	*Agrostis gigantea*	9.18±1.04	NPGS
agig-58	PI207454	*Agrostis gigantea*	12.82±1.06	NPGS
agig-59	PI207455	*Agrostis gigantea*	12.76±0.53	NPGS
agig-60	PI212281	*Agrostis gigantea*	12.73±0.41	NPGS
agig-61	PI221903	*Agrostis gigantea*	11.6±1.43	NPGS
agig-62	PI221904	*Agrostis gigantea*	12.7±1	NPGS
agig-63	PI221905	*Agrostis gigantea*	12.17±0.9	NPGS
agig-64	PI221907	*Agrostis gigantea*	11.84±1.33	NPGS
agig-65	PI221909	*Agrostis gigantea*	12.47±1.11	NPGS
agig-66	PI221910	*Agrostis gigantea*	12.39±1.34	NPGS
agig-67	PI222030	*Agrostis gigantea*	12.32±0.63	NPGS
agig-68	PI222031	*Agrostis gigantea*	13.09±1.05	NPGS
agig-69	PI223239	*Agrostis gigantea*	11.99±0.95	NPGS
agig-70	PI223240	*Agrostis gigantea*	11.37±0.64	NPGS
agig-71	PI223242	*Agrostis gigantea*	12.12±0.37	NPGS
agig-72	PI230124	*Agrostis gigantea*	12.48±0.46	NPGS
agig-73	PI230125	*Agrostis gigantea*	14.57±1.21	NPGS
agig-74	PI235439	*Agrostis gigantea*	9.68±0.49	NPGS
agig-75	PI251097	*Agrostis gigantea*	7.28±0.09	NPGS
agig-76	PI251098	*Agrostis gigantea*	8.54±0.94	NPGS
agig-77	PI251099	*Agrostis gigantea*	9.06±0.49	NPGS
agig-78	PI251250	*Agrostis gigantea*	7.37±0.6	NPGS
agig-79	PI251524	*Agrostis gigantea*	9.04±0.52	NPGS
agig-80	PI251568	*Agrostis gigantea*	10.44±0.53	NPGS
agig-81	PI251569	*Agrostis gigantea*	13.45±	NPGS
agig-82	PI251570	*Agrostis gigantea*	10.78±0.45	NPGS
agig-83	PI251944	*Agrostis gigantea*	10.34±1.75	NPGS
agig-84	PI255869	*Agrostis gigantea*	7.52±0	NPGS
agig-85	PI267051	*Agrostis gigantea*	9.37±0.4	NPGS
agig-86	PI272111	*Agrostis gigantea*	9.25±1.05	NPGS
agig-87	PI274601	*Agrostis gigantea*	9.52±0.59	NPGS

Study ID	Accession	Taxon	2C DNA (pg)	Source
agig-88	PI278696	*Agrostis gigantea*	10.14±0.55	NPGS
agig-89	PI283171	*Agrostis gigantea*	9.82±0	NPGS
agig-90	PI283172	*Agrostis gigantea*	9.95±0.14	NPGS
agig-91	PI298084	*Agrostis gigantea*	9.85±1.14	NPGS
agig-92	PI298085	*Agrostis gigantea*	9.34±1.63	NPGS
agig-93	PI305494	*Agrostis gigantea*	10.02±1.92	NPGS
agig-94	PI306681	*Agrostis gigantea*	9.93±1.19	NPGS
agig-95	PI311006	*Agrostis gigantea*	11.06±0.59	NPGS
agig-96	PI311007	*Agrostis gigantea*	10.87±1.04	NPGS
agig-97	PI311008	*Agrostis gigantea*	9.79±0.99	NPGS
agig-98	PI311009	*Agrostis gigantea*	9.35±0.76	NPGS
agig-99	PI311010	*Agrostis gigantea*	7.44±0.78	NPGS
agig-100	PI315069	*Agrostis gigantea*	10.15±0.5	NPGS
agig-101	PI371945	*Agrostis gigantea*	10.27±1.14	NPGS
agig-102	PI383584	*Agrostis gigantea*	8.37±1.59	NPGS
agig-103	PI387897	*Agrostis gigantea*	10.35±0.28	NPGS
agig-104	PI387898	*Agrostis gigantea*	11.04±0.65	NPGS
agig-105	PI387899	*Agrostis gigantea*	10.77±1.7	NPGS
agig-106	PI387900	*Agrostis gigantea*	9.76±0.61	NPGS
agig-107	PI387901	*Agrostis gigantea*	10.72±0.27	NPGS
agig-108	PI387902	*Agrostis gigantea*	10.12±0.88	NPGS
agig-109	PI387903	*Agrostis gigantea*	9.46±1.97	NPGS
agig-110	PI387904	*Agrostis gigantea*	9.65±1.17	NPGS
agig-111	PI387905	*Agrostis gigantea*	10.09±0	NPGS
agig-112	PI392336	*Agrostis gigantea*	11.92±1.21	NPGS
agig-113	PI392337	*Agrostis gigantea*	11.37±0.66	NPGS
agig-114	PI406637	*Agrostis gigantea*	11.38±0.98	NPGS
agig-115	PI440108	*Agrostis gigantea*	11.82±0.25	NPGS
agig-116	PI440111	*Agrostis gigantea*	13.15±0.91	NPGS
agig-117	PI440113	*Agrostis gigantea*	13.82±0.66	NPGS
agig-118	PI440115	*Agrostis gigantea*	13.19±1.24	NPGS
agig-119	PI443034	*Agrostis gigantea*	12.22±1.01	NPGS
agig-120	PI443037	*Agrostis gigantea*	13.6±1.5	NPGS
agig-121	PI443038	*Agrostis gigantea*	13.29±0.23	NPGS
agig-122	PI443042	*Agrostis gigantea*	13.42±0.63	NPGS
agig-123	PI443051	*Agrostis gigantea*	8.91±0.66	NPGS
agig-124	PI499394	*Agrostis gigantea*	12.94±0.09	NPGS
agig-125	PI502272	*Agrostis gigantea*	14.78±0.91	NPGS
agig-126	PI502273	*Agrostis gigantea*	13.96±0.91	NPGS
agig-127	PI502274	*Agrostis gigantea*	13.64±0.88	NPGS

Study ID	Accession	Taxon	2C DNA (pg)	Source
agig-128	PI502275	*Agrostis gigantea*	13.22±1.64	NPGS
agig-129	PI502276	*Agrostis gigantea*	13.41±0.99	NPGS
agig-130	PI502277	*Agrostis gigantea*	14.39±0.72	NPGS
agig-131	PI502278	*Agrostis gigantea*	12.42±0.65	NPGS
agig-132	PI527690	*Agrostis gigantea*	15.12±0.84	NPGS
agig-133	PI531250	*Agrostis gigantea*	13.85±0.67	NPGS
agig-134	PI532916	*Agrostis gigantea*	13.98±1.59	NPGS
agig-135	PI538782	*Agrostis gigantea*	12.99±0.51	NPGS
agig-136	PI538783	*Agrostis gigantea*	12.77±1.4	NPGS
agig-137	PI538784	*Agrostis gigantea*	9.35±1.13	NPGS
agig-138	PI590428	*Agrostis gigantea*	11.64±0.98	NPGS
agig-139	PI598457	*Agrostis gigantea*	13.42±0.82	NPGS
agig-140	PI598458	*Agrostis gigantea*	13.28±0.6	NPGS
agig-141	PI619441	*Agrostis gigantea*	13.88±0.55	NPGS
agig-142	PI619442	*Agrostis gigantea*	13.99±0.5	NPGS
agig-143	PI619443	*Agrostis gigantea*	13.49±0.26	NPGS
agig-144	PI619444	*Agrostis gigantea*	13.02±0.74	NPGS
agig-145	PI619538	*Agrostis gigantea*	11.54±0.79	NPGS
agig-146	PI619542	*Agrostis gigantea*	13.45±0.28	NPGS
agig-147	PI619551	*Agrostis gigantea*	12.26±1.79	NPGS
agig-148	PI619558	*Agrostis gigantea*	12.5±1.13	NPGS
agig-149	PI645596	*Agrostis gigantea*	13.07±1.64	NPGS
ahye-1	PI234681	*Agrostis hyemalis*	14.95±0.99	NPGS
ahyg-1	PI477045	*Agrostis hygrometrica*	21.66±0.89	NPGS
aina-2	W623575	*Agrostis inaequiglumis*	19.94±0.24	NPGS
aina-1	W623579	*Agrostis inaequiglumis*	19.5±0.34	NPGS
akin-1	PI462363	*Agrostis keniensis*	13.04±1.09	NPGS
alac-1	PI195917	*Agrostis lachnantha*	12.37±0.76	NPGS
alac-2	PI299461	*Agrostis lachnantha*	12.77±0.92	NPGS
alim-2	PI636682	*Agrostis limprichtii*	10.07±1.63	NPGS
alim-1	W623620	*Agrostis limprichtii*	9.64±1.03	NPGS
alya-1	PI636652	*Agrostis lyallii*	6.86±0.49	NPGS
alya-2	PI639829	*Agrostis lyallii*	19.7±1.29	NPGS
amon-1	PI362190	*Agrostis mongolica*	13.33±0.73	NPGS
amon-2	PI632490	*Agrostis mongolica*	11.71±1.01	NPGS
amon-3	PI632549	*Agrostis mongolica*	5.2±0.24	NPGS
amon-4	W618291	*Agrostis mongolica*	11.97±0.72	NPGS
amon-5	W619706	*Agrostis mongolica*	5.33±0.36	NPGS
amun-6	PI229720	*Agrostis munroana*	9.4±0.31	NPGS
amun-7	PI230236	*Agrostis munroana*	9.04±0.33	NPGS

Study ID	Accession	Taxon	2C DNA (pg)	Source
aneb-1	PI196319	*Agrostis nebulosa*	9.53±1.52	NPGS
apal-1	PI238226	*Agrostis pallida*	6.09±0.9	NPGS
asp-1	PI440112	*Agrostis sp.*	11.76±0.86	NPGS
asp-2	PI478596	*Agrostis sp.*	8.59±0.1	NPGS
asp-3	PI486302	*Agrostis sp.*	7.94±0.39	NPGS
asp-4	PI502279	*Agrostis sp.*	5.18±0.21	NPGS
asp-5	PI516483	*Agrostis sp.*	11.9±0.68	NPGS
asp-6	PI516484	*Agrostis sp.*	13.46±0.5	NPGS
asp-7	PI516485	*Agrostis sp.*	10.89±1.13	NPGS
asp-8	PI595077	*Agrostis sp.*	7.58±0.45	NPGS
asp-9	PI595078	*Agrostis sp.*	9.06±0.78	NPGS
asp-10	PI618782	*Agrostis sp.*	9.1±0.74	NPGS
asp-11	PI634313	*Agrostis sp.*	13.31±1.96	NPGS
asto-36	PI578529	*Agrostis stolonifera*	8.52±0.62	NPGS
asto-37	PI601363	*Agrostis stolonifera*	6.68±1.44	NPGS
asto-38	PI618723	*Agrostis stolonifera*	11.36±0.62	NPGS
asto-39	PI618724	*Agrostis stolonifera*	11.21±0	NPGS
asto-40	PI632687	*Agrostis stolonifera*	6.02±1.83	NPGS
asto-19	PI204390	*Agrostis stolonifera*	8.94±0.67	NPGS
asto-20	PI221906	*Agrostis stolonifera*	13.21±0.84	NPGS
asto-21	PI235440	*Agrostis stolonifera*	9.43±0.47	NPGS
asto-22	PI235541	*Agrostis stolonifera*	8.75±1.29	NPGS
asto-23	PI251945	*Agrostis stolonifera*	9.04±0.84	NPGS
asto-18	PI578530	*Agrostis stolonifera*	9.49±0.54	NPGS
asto-24	PI600776	*Agrostis stolonifera*	9.17±0.67	NPGS
asto-25	W66606	*Agrostis stolonifera*	8.86±0.83	NPGS
asto-26	PI222073	*Agrostis stolonifera*	13.78±0.27	NPGS
asto-27	PI230235	*Agrostis stolonifera*	12.36±1.12	NPGS
asto-28	PI269838	*Agrostis stolonifera*	9.54±0.55	NPGS
asto-29	PI302902	*Agrostis stolonifera*	8.57±0.75	NPGS
asto-30	PI318934	*Agrostis stolonifera*	8.73±0.78	NPGS
asto-31	PI439027	*Agrostis stolonifera*	11.64±1.37	NPGS
asto-32	PI494118	*Agrostis stolonifera*	8.33±0.82	NPGS
asto-33	PI494119	*Agrostis stolonifera*	7.32±1.04	NPGS
asto-34	PI531251	*Agrostis stolonifera*	7.53±0.29	NPGS
asto-35	W66602	*Agrostis stolonifera*	8.19±0.27	NPGS
atra-1	PI283174	*Agrostis transcaspica*	10.22±0.65	NPGS
atri-1	PI598462	*Agrostis trinii*	7.9±0.44	NPGS
atri-2	PI632559	*Agrostis trinii*	4.19±0.66	NPGS
atri-3	PI636572	*Agrostis trinii*	9.46±0.52	NPGS

Study ID	Accession	Taxon	2C DNA (pg)	Source
atri-4	W619572	*Agrostis trinii*	4.98±0.24	NPGS
atri-5	W621432	*Agrostis trinii*	4.9±0.45	NPGS
avin-1	PI440110	*Agrostis vinealis*	5.25±0.24	NPGS
agig-1	RUS-07-01-002	*Agrostis gigantea*	5.12±0	DJ
agig-2	RUS-07-02-003	*Agrostis gigantea*	6.8±0.7	DJ
agig-3	RUS-07-04-016	*Agrostis gigantea*	7.84±0.19	DJ
agig-4	RUS-07-10-032	*Agrostis gigantea*	7.46±0.78	DJ
agig-5	RUS-07-12-041	*Agrostis gigantea*	8.01±0.89	DJ
agig-6	RUS-07-15-054	*Agrostis gigantea*	12.72±1.34	DJ
agig-7	RUS-07-16-057	*Agrostis gigantea*	7.31±0.9	DJ
agig-8	RUS-07-20-076	*Agrostis gigantea*	7.56±0.64	DJ
agig-9	RUS-07-24-089	*Agrostis gigantea*	7.04±0.58	DJ
agig-10	RUS-07-27-103	*Agrostis gigantea*	7.95±0.33	DJ
agig-11	RUS-07-29-116	*Agrostis gigantea*	8.48±0.88	DJ
agig-12	RUS-07-32-131	*Agrostis gigantea*	8.92±0.53	DJ
agig-13	RUS-07-36-142	*Agrostis gigantea*	8.53±0.58	DJ
agig-14	RUS-07-39-156	*Agrostis gigantea*	12.68±0.42	DJ
agig-15	RUS-07-40-158	*Agrostis gigantea*	12.1±0.65	DJ
agig-16	RUS-07-43-166	*Agrostis gigantea*	7.7±1.02	DJ
agig-17	KGZ-06-02-009	*Agrostis gigantea*	12.32±1.04	NPGS
agig-18	KGZ-06-04-019	*Agrostis gigantea*	8.49±0.21	NPGS
agig-19	KGZ-06-07-044	*Agrostis gigantea*	12.45±1.2	NPGS
asto-1	KGZ-06-08-051	*Agrostis stolonifera*	8.19±0.5	NPGS
agig-20	KGZ-06-09-057	*Agrostis gigantea*	10.49±1.35	NPGS
agig-21	KGZ-06-13-069	*Agrostis gigantea*	12.61±1.12	NPGS
agig-22	KGZ-06-15-085	*Agrostis gigantea*	12.7±0.93	NPGS
agig-23	KGZ-06-17-095	*Agrostis gigantea*	12.9±0.67	NPGS
agig-24	KGZ-06-19-101	*Agrostis gigantea*	11.5±1.54	NPGS
agig-25	KGZ-06-20-105	*Agrostis gigantea*	12.01±0.65	NPGS
agig-26	KGZ-06-21-109	*Agrostis gigantea*	12.61±0.61	NPGS
agig-27	KGZ-06-24-115	*Agrostis gigantea*	12.77±0.65	NPGS
agig-28	KGZ-06-25-124	*Agrostis gigantea*	12.44±0.91	NPGS
agig-29	KGZ-06-27-133	*Agrostis gigantea*	13.58±0.37	NPGS
agig-30	KGZ-06-28-136	*Agrostis gigantea*	12.52±0.47	NPGS
agig-31	KGZ-06-29-140	*Agrostis gigantea*	12.11±0.89	NPGS
agig-32	KGZ-06-31-157	*Agrostis gigantea*	12.79±1.39	NPGS
agig-33	KGZ-06-32-160	*Agrostis gigantea*	12.34±0.16	NPGS
agig-34	KGZ-06-33-167	*Agrostis gigantea*	12.39±0.25	NPGS
agig-35	KGZ-06-34-174	*Agrostis gigantea*	12.25±0	NPGS
agig-36	KGZ-06-35-180	*Agrostis gigantea*	11.33±0.35	NPGS

Study ID	Accession	Taxon	2C DNA (pg)	Source
agig-37	KGZ-06-42-212	*Agrostis gigantea*	11.91±0.63	NPGS
agig-38	KGZ-06-44-219	*Agrostis gigantea*	11.35±0.66	NPGS
agig-39	KGZ-06-52-255	*Agrostis gigantea*	12.31±1.03	NPGS
agig-40	PRC-06-IN/060-131	*Agrostis gigantea*	9.93±1.22	NPGS
agig-41	PRC-06-IN/080-238	*Agrostis gigantea*	11.84±1.84	NPGS
agig-42	PRC-06-IN/084-257	*Agrostis gigantea*	10.65±1.17	NPGS
agig-43	PRC-06-IN/086-271	*Agrostis gigantea*	9.99±1.65	NPGS
agig-44	PRC-06-IN/097-309	*Agrostis gigantea*	11.06±1.2	NPGS
agig-45	PRC-06-IN/101-329	*Agrostis gigantea*	10.57±0.29	NPGS
agig-46	PRC-06-IN/102-336	*Agrostis gigantea*	11.7±0.52	NPGS
agig-47	PRC-06-IN/118-402	*Agrostis gigantea*	9.84±1.89	NPGS
asto-2	NTEP-1998-25	*Agrostis stolonifera*	7.94±0.31	NTEP
asto-3	NTEP-1998-26	*Agrostis stolonifera*	8.93±1.15	NTEP
asto-4	NTEP-1998-27	*Agrostis stolonifera*	9.11±0.75	NTEP
asto-5	NTEP-1998-29	*Agrostis stolonifera*	8.69±1.21	NTEP
asto-6	NTEP-2003-2	*Agrostis stolonifera*	8.76±0.97	NTEP
asto-7	NTEP-2003-6	*Agrostis stolonifera*	8.76±1.11	NTEP
asto-8	NTEP-2003-7	*Agrostis stolonifera*	8.81±0.61	NTEP
asto-9	NTEP-2003-10	*Agrostis stolonifera*	9.41±1.02	NTEP
asto-10	NTEP-2003-11	*Agrostis stolonifera*	8.83±0.59	NTEP
asto-11	NTEP-2003-14	*Agrostis stolonifera*	8.62±0.43	NTEP
asto-12	NTEP-2003-15	*Agrostis stolonifera*	7.93±0.51	NTEP
asto-13	NTEP-2003-21	*Agrostis stolonifera*	8.56±0.85	NTEP
asto-14	NTEP-2003-23	*Agrostis stolonifera*	8.51±0.3	NTEP
asto-15	NTEP-2003-24	*Agrostis stolonifera*	8.01±0.39	NTEP
asto-16	NTEP-2003-25	*Agrostis stolonifera*	8.74±0.51	NTEP
asto-17	NTEP-2003-26	*Agrostis stolonifera*	9.46±0.73	NTEP
acan-1	NTEP-1998-2	*Agrostis canina*	5.32±0.21	NTEP
acan-2	NTEP-1998-11	*Agrostis canina*	4.97±0.2	NTEP
acan-3	NTEP-1998-21	*Agrostis canina*	5.18±0.03	NTEP
acan-4	NTEP-2003-8	*Agrostis canina*	4.92±0.37	NTEP
acan-5	NTEP-2003-12	*Agrostis canina*	4.92±0.8	NTEP
acan-6	NTEP-2003-16	*Agrostis canina*	5.48±0.03	NTEP
acan-7	NTEP-2003-18	*Agrostis canina*	5.15±0.1	NTEP
acan-8	NTEP-2003-19	*Agrostis canina*	5.61±0.15	NTEP
acan-9	NTEP-2003-22	*Agrostis canina*	6.1±0.29	NTEP

Source: NPGS – National Plant Germplasm System, DJ – Dr. Doug Johnson, NTEP – National Turfgrass Evaluation Program

APPENDIX B. MITE-INSERTIONAL POLYMORPHISM PRIMER PAIRS

Clone ID	Primer1	Primer2	PCR	MT
CO-EST03T_C06	TCGAGTTCCTTGTTGGCAG	TCGGATCTGAACAACCCAA	472	348
CO-EST04T_D04	CATGGTTTCAATACGGGGA	GGTGAATCAAATCCAGGGG	199	142
CO-EST04T_F08	CAGGAGTTCAACATTATCGTCG	ATCAAGCTCATGGCTGTCG	378	227
CO-EST05T_D12	CAACCATGAAGTTGTCGGG	CAGAGCATCCTCCCCATTT	400	70
CO-EST06T_B08	GTATCCAGCCAGTCAGCCA	TTTTCCATGTGGGTGTCGT	478	244
CO-EST09T_H06	CAAACAAACTGACACCGCC	TGGTGGAATGGGTGGTATG	441	69
CO-EST10T_D02	TGTATGTGTGAAGTTTGAGATTAGGA	TGCACCTGGAACTGGAGAG	224	98
COL336T_G07	AAGGAATTGCTACCGTGGG	TCCATTCCTCGAATCAAACA	749	473
COL338T_A11	CAAATTTGACACCACTTTATTATGC	AACCATATGCGACGTCAGG	831	728
COL338T_C08	CCTCTTCAGCATCATGGCA	CCCTCCTATCGCTGGTTGT	389	339
COL338T_H12	GAGATTGGCCACAAGTTCG	TGAAATCATGCATTTTGCCT	803	663
COL339T_B06	TGTGTTTCGTTCTCGGAGC	TATTGGTCAGGCTTCGCTG	322	71
COL339T_D06	TCTGCAGATGCGTGGTACA	GCAGATAACTCGGCCTTCC	397	68
COL340T_C03	GGTGTTTCCGTTGATGCAC	CGGGGCTTCAAGTTTCATT	482	258
COL341T_A11	CCTGAGGCGTTTCTGTTCA	GGGAAATATACATGGCCCG	438	233
COL341T_E07	GAAGATGCCAAGGAGGAGC	TCAAAGCAATCGCACACAA	462	64
COL342T_D04	GATACCAATCACCGGGACC	AAGAGAGAGGATCCCGCAA	316	73
COL342T_G10	TCTTGTGCACAGCAACAGC	GAACAGCGGCAAAGTCAAA	816	721
COL344T_F01	CATCACCAGATTGGATCGC	AAAATTCGGATGGCTGAGG	765	499
COL344T_G08	GCCACCTCATATGCCTCCT	TTGCCTACACGGAGCAGAC	699	198
COL345T_E06	TCTGGGTTGGTTCCTGACA	AGGGATCTGATGGTCCGAG	719	133
COL347T_F07	TGGAAAAGGGCTCTCAACA	TGTCGGGGAAACAACTTCA	641	363
COL349T_D02	GGTTCACATGCACCAAAGC	ATAAAACACCCCGAGTCGC	741	409
COL350T_A12	CCAATAGCCCAAGGTGCTT	CAAACTCAGTCGCCTGCTC	240	141
COL352T_A09	TGTCATTCGGCAAAACCAT	GACAACGAAAGCCTGCTGA	633	286
COL352T_E09	CCCCCTCTCAATTTCTCCA	GGGGGTTCAACATGGAAAG	656	300
COL353T_A02	CTGCGCCCAAGTACTACGA	TCTGGAGCTTGTTCATGGC	474	346
COL353T_G04	TGTCGAAGGGACCTGTCTG	CGTGCTGGGCATTGTTTAT	327	128
COL354T_B08	TGTAAACTGCTTCTGGCCG	CACAGCAAAAAGAACCAACA	771	62
COL354T_G07	CTCGGACGACGCAAACTTA	CCTCGACGGTCTTTTGGTT	544	257
COL355T_E08	AACGTGGCCTAAAAGTGCC	TCATCGATCTGCTTCTGGC	760	176
COL355T_E11	GTCCGGCCAGTTTTTGAAT	ACTGCCGATTACGGAAAAA	527	397
COL356T_F01	TCAGCGGCACCATCTCTAC	CCCGCTAAGACGAACACAA	617	369
COL357T_D08	AGTCATCCCAGGAGATCCG	TGGTGGCATCTCTGCTTGT	767	74
COL358T_E07	TTACTCAAAGCGCCACCAG	CAAGGGGCATCATCAACAG	669	265

Clone ID	Primer1	Primer2	PCR	MT
COL359T_A07	GCTGTCCAGGTCGTGCTTA	CATACATCCAGAGTCCAGAGCA	370	186
COL363T_A06	CACCCTCTTTTTGGAAGCC	CCTCCAAAAGGGACCTCAA	635	337
COL364T_D10	CCATGGCTGGTACTGCAAC	ACGAACCCAGCTGCTCTTT	419	273
COL365T_F11	TCTTACAGCGTCCAGTGGC	TCCTGCCTGCTTAAACGTG	561	397
COL366T_F11	CCCCTTTTACCGCCAAAT	TGCCCCAGGGAGGTTATAC	1011	866
COL367T_E07	GCGTCTTCTTGTTGGCGTA	CCCGGAGAGATACCGTTGT	658	342
COL368T_E07	GGCTGCACACACAACTTCC	TTCTGCTTCGGTGAATTGC	537	178
COL374T_H07	GGAGGAGCTCCCTCTGGC	GAGAGTTCTCCCCAGCGAC	777	394
COL381T_C07	TTTGTGAAGGGGGAGAACC	AAGCACAAACACACACGCA	878	62
COL382T_D04	ACCACCCATGGAACCTGAT	CTACGGAACGGAGGAGAGG	645	208
COL383T_C01	AAGAAGGGCGAGATGCTGT	GGGCAAGATTGCGAAAAGT	553	325
COL383T_D08	GAACCGTGGAAAACCGACT	ACGCCACCGTTACAAACAG	825	668
COL383T_D11	TCGACCAACGCAAGTTAGC	GGTATGTTTTGCCGCCTTT	498	73
COL383T_E03	TTGCCACTGCCTTCTTTTG	ATCCCCTGACCAATCTTGC	694	331
COL383T_G12	ATGCACGATGAACCCAAGA	CAGCAGCAGATTCTTCACCA	441	240
COL384T_D01	CGGCAACTACTTCAGCGAC	TGCGAGCAACACACCTACA	506	269
COL384T_D12	ATTCTTTGTGTCCCTGCCC	GAATTCGGCACGAGGATTT	569	86
COL384T_E06	AGCAGCTGCACCATACCAA	ACAGGATGTCATGCTCGGA	930	761
COL384T_G09	TGTTTGGATATCACGACCCA	GTGAAGCTCTCCCTTTGGC	469	76
COL385T_E08	CTCGACGTGGAGTCTGCAT	AGCGCATCCTAGACTCGGT	329	160
COL385T_H11	CTGCTTGCATACCTTGCGT	AAGCTTCCGGGGTCTTCTT	608	229
COL386T_E07	GTTGAGCACTGCTGACCGT	TTTTCTACTCCGATGGGGC	763	200
COL387T_D05	GCCATATGTTCCGACAACG	TATCTTTCCGGGGCTTTTG	457	180
COL387T_F09	GACAGCCTGCACCACATCT	GCGCTTGAAGTGCAGAGAG	250	77
COL389T_D08	TACGCGAAGCTTTGGAATG	CATCCAGCTCAATTCGAACA	403	151
COL392T_D03	TCACCGAGCACAACAACAA	CGCATCCGAAACTCAAAAA	479	333
COL393T_B04	GAATTTAACCCCTCGTGCC	AGGAACTGCCTCCCATACG	630	534
COL393T_G01	GTTCCCAGGAGGAAGACGA	GTAAGGGGTGAAGCTGGGA	645	139
COL395T_D06	CACCATCCTCCATCCACC	AAACCGTCTTGCTCCATCC	662	178
COL398T_B02	GCGTCCTTGGGAGTAGCTG	CTGTGGTATGGGTTGTGGC	331	120
COL398T_D10	AAAATAGGGCCTCGTGCC	GAAAACTAGCACGCACCGA	760	492
COL399T_G12	TCGACCAACGCAAGTTAGC	AGCTACTCCCTCCGATCCA	231	73
COL400T_G08	ACCACCACTGGGGATCATT	GGACCCATTTCTCTTTGCG	775	491
COL403T_B01	TGCTCCCATTGAAGAGTGC	ATACAGCGTCCAGCAGCAG	793	460
COL403T_D12	GTACGCCCCACCAGAAGAC	GACCCCTACCTCTCCGTCA	240	66
COL403T_G05	CAGCCGTTGAACTTCTGGA	TTGGATGTGTGCTGATTGC	700	165
COL403T_H08	AGAAAGAACCGCCCATACG	TACTGAATTCGGCACGAGG	543	423
COL404T_E06	TCTCCCTCCTCTTCGTCGT	AACAGCCCTTTGCAGTCCT	451	85
COL404T_F01	GGCCTGTACGGCAGCAAG	CGGACATAGGCTTCAGGGT	531	445
COL404T_H03	CCTCTGTCTCACTTTACTCTGCG	TGGCAAGCCTGCAATTTAG	791	676

Clone ID	Primer1	Primer2	PCR	MT
COL405T_G01	ACAGACATGGCTTGCACCT	GCGCGCCAATTAAACATTA	629	169
COL406T_B06	AAGTTGAGCCACTACCCGC	CATAGCTGCCCTCTCACCA	700	265
COL406T_F10	GCTCCAGACTACGTTGCCA	GAGTCAAGGGCACCAACTCT	559	482
COL407T_H11	TTCTCCCGAATCTCTCGCT	AAAGGCGATTTTGCGTGTT	504	226
COL411T_H05	GGAAGCATCATCCGTTCCT	GGTCGTTGACCAATACGGA	865	186
COL413T_A03	CTTCTCGGCAGCGACGAT	ATGTCTTTGCAACCTCCCC	561	227
COL416T_G02	ACGATCGATGCTGAACCAA	ACCGATCAGCAGGTGTTTG	695	405
COL422T_A10	GCCCGAGGAATTTTAAATGA	TGTCTATTTAATGAAGGCAATTATTTT	635	580
COL426T_E01	ACTACTCCCGCGCCATACT	TGCGCTCATTCTTCAAACC	714	320
COL428T_E02	GTACAACATCACCGACGGC	GCGCGATGGGTACAGTCT	667	109
COL428T_F02	TGATTTCAGAACCCGTTGC	CATGGCAGGATTATGCGAC	619	478
COL430T_H05	ATGCCGTGGAGAGAAATCC	CGTTAACATTCCCCCATCC	549	73
COL431T_B07	ACGTCACCTTCAAGGACCC	CAGCAAAGCAGCCATTGTT	475	165
COL432T_A05	TGACCAGCCATCAGGAGAG	GTAGCAATTTGGAGCGCAA	641	352
COL432T_D12	GCTGTCCAGGTCGTGCTTA	CATCCCGAGTCCAGAGCAT	366	186
COL432T_G07	TTCTCTCCCCCTTGTGTCC	ATAGGGCCCTGCTGCTCT	856	447
Colonial-1_G07	TCGGTTTGTCGTGTAGGGA	TTTTACTGCGAGTTTGGCG	432	191
Colonial2T_H05	GACCCACTACGGCCTCATT	ATCGAGGGGCTAAAGCTGA	578	188
Colonial4T_C05	AGGGTACGAACCCGTCCTA	ACGGAGTTAAGGGCTGTGC	301	132
Colonial6T_C10	CGCCAAAGTCCAAAGACAA	TTAGCGTGCCGGTCAGTAG	714	136
Colonial6T_G04	CCAGCAACACCATTTGACC	TTTGGCATAGCTACGGTGC	406	136
Colonial6T_H07	AGGCCACTTTCTCTGCCAT	CGGCTCCATCTCATCTGAA	469	147
Colonial7T_A06	ATTCACCTCATGCAACGGA	TTTGGCCGCACATTTACTC	590	146
Colonial7T_F06	TGACAAAATGACAAAATGGTTCA	CCGCGTGAATACTCATTGG	664	419
Colonial7T_H07	CCACAATGCTCCTGGTTTG	TGGAGAGTGAAATTTTCCGA	564	183
Colonial8T_H09	TCAGGGTGAACCGTAATCG	TTTGAAAACCGTGCACACA	366	71
Colonial9T_A03	CCGTAATGACCACAGCACC	ACATACGGGACCACGTCCT	775	359
Colonial9T_G02	TTAACCAAGGGCATGCAAA	ACCACGGCTTAATTTTGGC	709	390
ColRandom2T_D10	GCAGTCCCTCTCCCTTGAG	AAGGAGGCGGTGTTCAAGA	216	123
ColRandom2T_E10	ATCACACCTCCACCAGCAA	CGAAAACCGAGGAGAGGAG	386	165
CR-EST10T_C01	GCACGGTGTCGGATGTAAA	CTGGATGTTATTGGCACCG	295	62
Creeping-1_E10	TGTTAGGCAAAATTTGAGCA	AGCAGCAGAGTTCCAGACTTT	465	361
Creeping10T_B08	AAATCCTGCCGCCGTATAG	CAATGAGCCCCATAGGCTT	597	268
Creeping10T_E10	GAAACGTACCCACCTGCCT	TCCGCACCACATTGACTTT	691	347
Creeping4T_C04	AAGAAATGGCCAACCTTGC	CTTCTGCGTGTGTGTGCTG	368	228
Creeping4T_E02	GAACTTGACGTCCGAGCAA	GGCAATTTCAATAGTGCCTCA	405	71
Creeping5T_H02	GCTTATTGGTCCTGGGCAT	CAATCCCATCCCATCTTCC	321	63
Creeping5T_H10	CATGATCGAGGACGATTGG	AAGGTGGTGCGCCACTAAT	735	80
Creeping9T_B09	CATCTCGCCCCTCGATATT	TCGTTGAGGGTGTTTTTGG	475	204
Creeping9T_E07	CCGCCATGCTGTATGAATC	TAGCCCCAAAGTCGAGAGG	353	82

Clone ID	Primer1	Primer2	PCR	MT
CRP227T_G12	CTGGGCATTGGGTACATTG	AGAATCGCTGCCACATCAG	373	242
CRP228T_B12	GCCTTGCGGTCAGGTAAAT	CATCACCGATCATCCGAAG	662	166
CRP229T_E09	CGTCATTAGTCTCGGGGGT	ATTAAGCCCGCTTAGGCTC	811	201
CRP230T_F08	CGAGGCGACAAGAAGCTAGT	GCAATGAAAACGGTTTATTAGTTCA	671	539
CRP231T_E07	CAACCAGAACCCAGCAATG	CCCTCCTCCTGCTTACTCG	714	367
CRP231T_F07	CTCAACAGCAGATACGGGC	AATGCAACTGCTCACCAGC	413	220
CRP238T_C07	TGCTGGTGGCTTACACCAT	AGCAGACTCCCGACTTGGT	436	208
CRP238T_E03	TTGGAAAGTCTCAAGGGGG	CATTACGCCACCGGATTTT	196	141
CRP239T_G02	GCGTCATCGAACTGTTGCT	GTCGTAGAAGGCCGAGACC	492	286
CRP240T_E09	CTCACACGGTGGATTGGAC	TAGTTGCAATTGGCTGCTG	738	679
CRP242T_A01	TCTCCGGAGGAATGTGTGA	GTTGGTCACCGCATTTTTG	233	141
CRP243T_G08	TGACAGCAGCCATGTGTGT	TGCACATCAAAGCATCAGG	369	278
CRP244T_F04	ATGGTGAGAGTTTTGGGGC	TGGGGTCGATTTGAATGAA	547	124
CRP246T_F01	AAATTTTCCGTACCCAGCG	TTTGGTGCAAAACGGACAT	273	78
CRP248T_B08	CCAAGATGCCACCAATGTT	CTTCCCAAGACGTTCCGAT	628	533
CRP250T_C02	ACGAGGCAGCTTTGGAGTT	TCACAATGGCAAGCACAGA	616	335
CRP252T_G05	ATGCACGCTTGTTTTGAGC	GACATGGTGCACGAATCTG	839	470
CRP257T_G06	AAAGACCACCGTCCTTTGG	TGTCGCCCTTGACTTCATC	815	295
CRP263T_C09	CGAGGGCTTTGTGGAAGAT	ACTTGCTCAACCACAACCG	646	444
CRP264T_C12	GAGGCATCAACCGTGAGAA	TTCGTCGCATCTACGGAAC	489	396
CRP269T_C04	CGCAAGGCGATCATGTTTA	CCGGATCTAAATGCAGGAAA	387	180
CRP277T_B11	AGCAGCATGTCAATTTCGC	GCGTGGAGGAAAGCTGGT	476	272
CRP280T_A03	GAGGTGACATGCACCAACA	AGATTTCCCGTGGGTTGAG	965	915
CRP280T_C03	TGAAGGCCAGCTTGACAAC	TATGTCACAACCCAACCCC	538	297
CRP281T_D09	CCTTCCCGAAAACGGTAGA	GATCCGTTGGGACACACAG	477	203
CRP282T_D01	GCACGAGGGGAATTACCAA	TGCCACCTTTTTCAGCAAG	290	123
CRP285T_G09	AGTGTTAAGGCTGGGGTGG	AAGGGGGTTTTGTTTTCCA	533	278
CRP286T_H09	GCGAAATTCTTGCATCCCT	CGATAGCAGATCATGCGAATA	460	277
CRP289T_E09	ATCAGCAAACTGTGAGCCG	CACTCAAGCTCGGTTGGAA	759	422
CRP292T_G01	ATGAGGCCTTGGAAAGTGG	CATAGCGGGCTTAGCAACA	431	261
CRP298T_C04	TCAACCGGTTCTACATCTTGG	ATTGTCCCCATCACGGTTT	768	299
CRP304T_A04	GGGACTACAACGGCAACCT	TAAGGATCTACCGCCGCTT	641	165
CRP304T_F02	GGAACTCCGACCCAAGTGT	GTATGCCCCTCGAACTCGT	322	139
CRP314T_B05	GCTCAAGGGGAAGGTGGTA	AAACAAAAGTCTCGGAAGGAA	877	651
CRP314T_C11	CAGCTGAAGAGCAGCAAGG	AAACAATAAGGCGGGATCG	466	334
CRP315T_B09	ACGTCATTTTTCTTCGGGC	GAATCTGTCCCCTCCCAGA	621	97
CRP315T_C10	CGGTCTTCCTCATCCCCTA	AAAAGATCGGCCACAGACC	911	177
CRP318T_A09	AGGTCTTCGAGATGCCCAC	CCCCCACCAATAACTACCG	391	74
CRP318T_F09	AGGGACACCCACAGTGACA	TGCTACCAAAAATGCAGCC	365	238
CRP319T_C01	GAGGAGAGGAGATGGCAGG	TCACACACGAGTGCGAAAA	484	165

Clone ID	Primer1	Primer2	PCR	MT
CRP320T_G05	AGGTCTTCGAGATGCCCAC	AACGTCCAAGGGGTTCTTG	301	74
CRP321T_H01	CTATCTGAGGGGTTGGCGT	TTCTTTCCTGAAACAAGGACAAA	746	520
CRP328T_C12	GGTATTTCCAAACCCACCG	TTACATCTGCAGGGCAAGC	511	66

PCR – Expected PCR product size; MT – Size of predicted MITE

```
aint-1      CTTTTTCATTCTACTCTTTCAAAAAGGAGTGCGAAGAAAACTCAATGGAT
aint-2      CTTTTTCATTCTACTCTTTCAAAAAGGAGTGCGAAGAAAACTCAATGGAT
pchi-1      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
pvir-1      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
pvir-2      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acan-10     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acan-11     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acap-7      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acap-25     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acap-26     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acas-11     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acas-3      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acas-8      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acla-2      CTTTCTCATTCTACTCTTTCACAAAAGAGTGCGAAGAGAACTCAATGGAT
acla-5      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
agig-77     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
alim-1      CTTTTTCATTCTACTCTTTCACAAAGGAGTACGAAGAGAACTCAATGGAT
alya-1      CTTTTTCATTCTACTCTTTCACAAAGGAGTGCGAAGAGAACTCAATGGAT
amon-3      CTTTCTCATTCTACTCTTTCACAAAAGAGTGCGAAGAGAACTCAATGGAT
amon-5      CTTTCTCATTCTACTCTTTCACAAAAGAGTGCGAAGAGAACTCAATGGAT
amun-7      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
apal-1      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asp-4       CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-36     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-40     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-19     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-23     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-29     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-30     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-34     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
atri-1      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
atri-2      CTTTCTCATTCTACTCTTTCACAAAAGAGTGCGAAGAGAACTCAATGGAT
atri-3      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
atri-4      CTTTCTCATTCTACTCTTTCACAAAAGAGTGCGAAGAGAACTCAATGGAT
avin-1      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
agig-1      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
agig-5      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
agig-12     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-1      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
agig-47     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
```

```
asto-3      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-4      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-5      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-6      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-7      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-9      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-12     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
asto-13     CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acan-2      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acan-6      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acan-7      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT
acan-8      CTTTCTCACTCTACTCTTTCCCAAAGGAGTGCGAAGAGAACTAAATGGAT

aint-1      CTTATCCTAGAATATATTTATTTTTTATTAGAGTATCGGGAAGGAATCCC
aint-2      CTTATCCTAGAATATATTTATTTTTTATTAGAGTATCGGGAAGGAATCCC
pchi-1      CTTATCCTATAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
pvir-1      CTTATCCTATAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
pvir-2      CTTATCCTATAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acan-10     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acan-11     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acap-7      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acap-25     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acap-26     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acas-11     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acas-3      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acas-8      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acla-2      CTTATCCTAGAATATATTTCTTT--------------------------
acla-5      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
agig-77     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
alim-1      CTTATCCTAGAATATATTTCTTTTTTATTAGAGTATCGGGAAGGAATCCC
alya-1      CTTATCCTAGAATATATTTCTTTTTTATTAGAGTATCGGGAAGGAATCCC
amon-3      CTTATCCTAGAATATATTTCTTT--------------------------
amon-5      CTTATCCTAGAATATATTTCTTT--------------------------
amun-7      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
apal-1      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
asp-4       CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
asto-36     CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-40     CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-19     CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-23     CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-29     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
asto-30     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
asto-34     CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
atri-1      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
atri-2      CTTATCCTAGAATATATTTCTTT--------------------------
```

```
atri-3      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
atri-4      CTTATCCTAGAATATATTTCTTT-------------------------
avin-1      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
agig-1      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
agig-5      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
agig-12     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
asto-1      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
agig-47     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
asto-3      CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-4      CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-5      CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-6      CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-7      CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-9      CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
asto-12     CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
asto-13     CTTATCCTAGAATAGATTTCTTTTTTAGTCGAGTATCGGGAAGGAATCCC
acan-2      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acan-6      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acan-7      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC
acan-8      CTTATCCTAGAATAGATTTCTTTTTTATTCGAGTATCGGGAAGGAATCCC

aint-1      GGGTATTCACTCTATTTTT-CAGTATTATTAAGTAAACCATGTACAATGC
aint-2      GGGTATTCACTCTATTTTT-CAGTATTATTAAGTAAACCATGTACAATGC
pchi-1      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
pvir-1      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
pvir-2      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acan-10     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acan-11     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acap-7      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acap-25     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acap-26     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acas-11     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acas-3      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acas-8      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acla-2      ----------------------AGTATTAAGTAAACCATGTACAATGC
acla-5      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
agig-77     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
alim-1      GGTTATTCAATCTATTTTTTCAGTATTATTAAGTAAACCATGTACAATGC
alya-1      GGTTATTCAATCTCTTTTTTCAGTATTATTAAGTAAACCATGTACAATGC
amon-3      ----------------------AGTATTAAGTAAACCATGTACAATGC
amon-5      ----------------------AGTATTAAGTAAACCATGTACAATGC
amun-7      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
apal-1      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asp-4       GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-36     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
```

```
asto-40     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-19     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-23     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-29     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-30     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-34     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
atri-1      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
atri-2      ----------------------AGTATTAAGTAAACCATGTACAATGC
atri-3      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
atri-4      ----------------------AGTATTAAGTAAACCATGTACAATGC
avin-1      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
agig-1      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
agig-5      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
agig-12     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-1      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
agig-47     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-3      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-4      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-5      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-6      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-7      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-9      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-12     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
asto-13     GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acan-2      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acan-6      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acan-7      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC
acan-8      GGTTATTCACTCTATTTTT-AAGTATTATTAAGTAAGCCATATACAATGC

aint-1      ATAGGACTACCCACCCA-TTTTCCAA-TTTAGAATTTGAAATATTTTATT
aint-2      ATAGGACTACCCACCCA-TTTTCCAA-TTTAGAATTTGAAATATTTTATT
pchi-1      GTAGGACTACCCTCCC--ATTTTTAAATTTCGAATTTAAAATACTTTATT
pvir-1      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTAAAATACTTTATT
pvir-2      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTAAAATACTTTATT
acan-10     GTAGGACTACCCTCCC--ATTTTAAAATTTAGAATTTGAAATACTTTATT
acan-11     GTAGGACTACCCTCCC--ATTTTAAAATTTAGAATTTGAAATACTTTATT
acap-7      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
acap-25     GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
acap-26     GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
acas-11     GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
acas-3      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
acas-8      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
acla-2      ATAGGACTACCCCCCCCCACTTTCAAATTTAGAATTTGAAATACTTTATT
acla-5      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACTTTATT
agig-77     GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACTTTATT
```

```
alim-1      ATAGGACTACCTCCCCG-TTTTTCAAATTTAGAATTTGAAATACTTTATT
alya-1      ATAGGACTACTCCCCCG-TTTTTCAAATTTAGAATTAGAAATACTTTATT
amon-3      ATAGGACTACCCCCCCCCACTTTCAAATTTAGAATTTGAAATACTTTATT
amon-5      ATAGGACTACCCCCCCCCACTTTCAAATTTAGAATTTGAAATACTTTATT
amun-7      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
apal-1      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
asp-4       GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACTTTATT
asto-36     GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-40     GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-19     GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-23     GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACCTTATT
asto-29     GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
asto-30     GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
asto-34     GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
atri-1      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
atri-2      ATAGGACTACCCCCCCCCACTTTCAAATTTAGAATTTGAAATACTTTATT
atri-3      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACTTTATT
atri-4      ATAGGACTACCCCCCCCCACTTTCAAATTTAGAATTTGAAATACTTTATT
avin-1      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACTTTATT
agig-1      GTAGGACTACCCTCCC--ATTTTAAAATTTAGAATTTGAAATACTTTATT
agig-5      GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
agig-12     GTAGGACTACCCTCCC--ATTTTCAAATTTCGAATTTGAAATACTTTATT
asto-1      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACTTTATT
agig-47     GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACTTTATT
asto-3      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-4      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-5      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-6      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-7      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-9      GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-12     GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
asto-13     GTAGGACTACCCTCCC--ATTTTCAAATTTAGAATTTGAAATACCTTATT
acan-2      GTAGGACTACCCTCCC--ATTTTAAAATTTAGAATTTGAAATACTTTATT
acan-6      GTAGGACTACCCTCCC--ATTTTAAAATTTAGAATTTGAAATACTTTATT
acan-7      GTAGGACTACCCTCCC--ATTTTAAAATTTAGAATTTGAAATACTTTATT
acan-8      GTAGGACTACCCTCCC--ATTTTAAAATTTAGAATTTGAAATACTTTATT

aint-1      TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTTTACTA
aint-2      TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTTTACTA
pchi-1      TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
pvir-1      TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
pvir-2      TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acan-10     TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acan-11     TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acap-7      TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
```

acap-25 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acap-26 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acas-11 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acas-3 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acas-8 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acla-2 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTCTACTA
acla-5 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
agig-77 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
alim-1 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTCTACTA
alya-1 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTCTACTA
amon-3 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTCTACTA
amon-5 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTCTACTA
amun-7 TAATTTATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
apal-1 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asp-4 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-36 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-40 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-19 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-23 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-29 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-30 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-34 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
atri-1 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
atri-2 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTCTACTA
atri-3 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
atri-4 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATACTCTACTA
avin-1 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
agig-1 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
agig-5 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
agig-12 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-1 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
agig-47 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-3 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-4 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-5 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-6 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-7 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-9 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-12 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
asto-13 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acan-2 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acan-6 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acan-7 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA
acan-8 TAATTGATTTTTTAGTCCCTTTAATTGACATAGATACAAATCCTCTACTA

aint-1 GGATTATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
aint-2 GGATTATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
pchi-1 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
pvir-1 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
pvir-2 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acan-10 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acan-11 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acap-7 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acap-25 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acap-26 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acas-11 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acas-3 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acas-8 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acla-2 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acla-5 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
agig-77 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
alim-1 GGATTATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
alya-1 GGATTATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
amon-3 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
amon-5 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
amun-7 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
apal-1 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asp-4 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-36 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-40 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-19 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-23 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-29 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-30 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-34 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
atri-1 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
atri-2 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
atri-3 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
atri-4 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
avin-1 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
agig-1 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
agig-5 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
agig-12 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-1 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
agig-47 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-3 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-4 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-5 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-6 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-7 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG

asto-9 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-12 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
asto-13 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acan-2 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acan-6 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acan-7 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG
acan-8 GGATGATGCACAAGAAAAGGTCAGGATAGCTCAGTTGGTAGAGCAGAG

```
aint-1      CAAGAAATTCTTACTTCATAAGCTCTATTGGAAAGAAGTATACTTCTAAG
aint-2      CAAGAAATTCTTACTTCATAAGCTCTATTGGAAAGAAGTATACTTCTAAG
pchi-1      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
pvir-1      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
pvir-2      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
acan-10     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
acan-11     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
acap-7      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
acap-25     CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
acap-26     CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
acas-11     CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
acas-3      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-ACC------
acas-8      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
acla-2      CAATAAATTCTTACTTCATAAGCTCTATTGAGGAGAAGTATACTTCTAAG
acla-5      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
agig-77     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
alim-1      CAAGAAATTCTTATTTCATAAGCTCGATTGGGAAGAAGTATACTTCTAAG
alya-1      CAAGAAATTCTTACTTCATAAGCTCTCTTGGGAAGAAGTATACTTCTAAG
amon-3      CAATAAATTCTTACTTCATAAGCTCTATTGAGGAGAAGTATACTTCTAAG
amon-5      CAATAAATTCTTACTTCATAAGCTCTATTGAGGAGAAGTATACTTCTAAG
amun-7      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
apal-1      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
asp-4       CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-36     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-40     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-19     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-23     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-29     CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
asto-30     CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
asto-34     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
atri-1      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
atri-2      CAATAAATTCTTACTTCATAAGCTCTATTGAGGAGAAGTATACTTCTAAG
atri-3      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
atri-4      CAATAAATTCTTACTTCATAAGCTCTATTGAGGAGAAGTATACTTCTAAG
avin-1      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
agig-1      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
agig-5      CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
agig-12     CAAGAAATTCTTACTTC-TAAGATCTATTGGGGAGAAGTA-AC-------
asto-1      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
agig-47     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
```

```
asto-3      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-4      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-5      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-6      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-7      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-9      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-12     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
asto-13     CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
acan-2      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
acan-6      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
acan-7      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------
acan-8      CAAGAAATTCTTACTTC-TAAGCTCTATTGGGGAGAAATA-AC-------

aint-1      CTCTATCTTTATTGTGGAGAAGTAACTAAAAAAGTACAAATTGAAATTAT
aint-2      CTCTATCTTTATTGTGGAGAAGTAACTAAAAAAGTACAAATTGAAATTAT
pchi-1      -------------------------TAAAAAAGTACAAATTGAAATGAT
pvir-1      -------------------------TAAAAAAGTACAAATTGAAATGAT
pvir-2      -------------------------TAAAAAAGTACAAATTGAAATGAT
acan-10     -------------------------TAAAAAAGTACAAATTGAAATGAT
acan-11     -------------------------TAAAAAAGTACAAATTGAAATGAT
acap-7      -------------------------TAAAAAAGTACAAATTGAAATGAT
acap-25     -------------------------TAAAAAAGTACAAATTGAAATGAT
acap-26     -------------------------TAAAAAAGTACAAATTGAAATGAT
acas-11     -------------------------TAAAAAAGTACAAATTGAAATGAT
acas-3      -------------------------TAAAAAAGTACAAATTGAAATGAT
acas-8      -------------------------TAAAAAAGTACAAATTGAAATGAT
acla-2      CTCTATCTCTATTGTGGAGAAGTAACTAAAAAAGTACAAATGAAAATGAT
acla-5      -------------------------TAAAAAAGTACAAATTGAAATGAT
agig-77     -------------------------TAAAAAAGTACAAATTGAAATGAT
alim-1      CTCTATCTCTATTGTGGAGAAGTCACTAAAAAAGTAAAAATGGAAATGAT
alya-1      CTCTATCTCTATTGTGGAGAAGTCACTAAAAA-GTAAAAATTGAAATGAT
amon-3      CTCTATCTCTATTGTGGAGAAGTAACTAAAAAAGTACAAATGAAAATGAT
amon-5      CTCTATCTCTATTGTGGAGAAGTAACTAAAAAAGTACAAATGAAAATGAT
amun-7      -------------------------TAAAAAAGTACAAATTGAAATGAT
apal-1      -------------------------TAAAAAAGTACAAATTGAAATGAT
asp-4       -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-36     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-40     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-19     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-23     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-29     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-30     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-34     -------------------------TAAAAAAGTACAAATTGAAATGAT
atri-1      -------------------------TAAAAAAGTACAAATTGAAATGAT
atri-2      CTCTATCTCTATTGTGGAGAAGTAACTAAAAAAGTACAAATGAAAATGAT
```

```
atri-3      -------------------------TAAAAAAGTACAAATTGAAATGAT
atri-4      CTCTATCTCTATTGTGGAGAAGTAACTAAAAAAGTACAAATGAAAATGAT
avin-1      -------------------------TAAAAAAGTACAAATTGAAATGAT
agig-1      -------------------------TAAAAAAGTACAAATTGAAATGAT
agig-5      -------------------------TAAAAAAGTACAAATTGAAATGAT
agig-12     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-1      -------------------------TAAAAAAGTACAAATTGAAATGAT
agig-47     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-3      -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-4      -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-5      -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-6      -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-7      -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-9      -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-12     -------------------------TAAAAAAGTACAAATTGAAATGAT
asto-13     -------------------------TAAAAAAGTACAAATTGAAATGAT
acan-2      -------------------------TAAAAAAGTACAAATTGAAATGAT
acan-6      -------------------------TAAAAAAGTACAAATTGAAATGAT
acan-7      -------------------------TAAAAAAGTACAAATTGAAATGAT
acan-8      -------------------------TAAAAAAGTACAAATTGAAATGAT

aint-1      AATGTGAATTGTCTGAACTACATAAAAGAGAATTCCATATATCGGATTAG
aint-2      AATGTGAATTGTCTGAACTACATAAAAGAGAATTCCATATATCGGATTAG
pchi-1      AATGTGAATTGTCCGAACTA------AGGGAATTCCATATCTCGGATTAG
pvir-1      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
pvir-2      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acan-10     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acan-11     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acap-7      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acap-25     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acap-26     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acas-11     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acas-3      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acas-8      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acla-2      AATGTGAATTATCTGAACTACATAGAAGATAATTCTATATATCGGATTAG
acla-5      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
agig-77     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
alim-1      AATGTGAATTGTATGAACTACATAGAAGATAATTCCATATATCGGATTAG
alya-1      AATGTGAATTGTCTGAACTACATAGAAGAGAATTCCATATATCGGATTAG
amon-3      AATGTGAATTATCTGAACTACATAGAAGATAATTCTATATATCGGATTAG
amon-5      AATGTGAATTATCTGAACTACATAGAAGATAATTCTATATATCGGATTAG
amun-7      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
apal-1      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asp-4       AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-36     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
```

```
asto-40     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-19     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-23     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-29     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-30     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-34     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
atri-1      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
atri-2      AATGTGAATTATCTGAACTACATAGAAGATAATTCTATATATCGGATTAG
atri-3      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
atri-4      AATGTGAATTATCTGAACTACATAGAAGATAATTCTATATATCGGATTAG
avin-1      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
agig-1      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
agig-5      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
agig-12     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-1      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
agig-47     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-3      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-4      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-5      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-6      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-7      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-9      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-12     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
asto-13     AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acan-2      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acan-6      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acan-7      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG
acan-8      AATGTGAATTGTCCGAACTACATATAAGGGAATTCCATATCTCGGATTAG

aint-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
aint-2      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
pchi-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
pvir-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
pvir-2      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
acan-10     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
acan-11     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
acap-7      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
acap-25     ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
acap-26     ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
acas-11     ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
acas-3      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
acas-8      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
acla-2      ATAATGAATCTAACCTAGGAA-----AATACCTTACA-TACATTTGTTTC
acla-5      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
agig-77     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
```

```
alim-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
alya-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
amon-3      ATAATGAATCTAACCTAGGAA-----AATACCTTACA-TACATTTGTTTC
amon-5      ATAATGAATCTAACCTAGGAA-----AATACCTTACA-TACATTTGTTTC
amun-7      AGAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
apal-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
asp-4       ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-36     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-40     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-19     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-23     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-29     ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
asto-30     ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
asto-34     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
atri-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
atri-2      ATAATGAATCTAACCTAGGAA-----AATACCTTACA-TACATTTGTTTC
atri-3      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
atri-4      ATAATGAATCTAACCTAGGAA-----AATACCTTACA-TACATTTGTTTC
avin-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
agig-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
agig-5      ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
agig-12     ATAATGAATCTAACCTAGGAATATATAATACCCTATA-TACATTTGTTTC
asto-1      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
agig-47     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-3      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-4      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-5      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-6      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-7      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-9      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-12     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
asto-13     ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
acan-2      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
acan-6      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
acan-7      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC
acan-8      ATAATGAATCTAACCTAGGAATATATAATACCCTATAATACATTTGTTTC

aint-1      TT------CTATTTTGTTTGCATATTTTCTCATTTTCTATTGAATCGGAT
aint-2      TT------CTATTTTGTTTGCATATTTTCTCATTTTCTATTGAATCGGAT
pchi-1      TT------CTATTTTGTTTCCGTATTTTATCATTTTCTATTGAATCGGAT
pvir-1      TT------CTATTTTGTTTCCGTATTTTATCATTTTCTATTGAATCGGAT
pvir-2      TT------CTATTTTGTTTCCGTATTTTATCATTTTCTATTGAATCGGAT
acan-10     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acan-11     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acap-7      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
```

```
acap-25     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acap-26     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acas-11     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acas-3      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acas-8      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acla-2      TTCTACTTCTATTTTGTTTGCATATTTTCTCATTTTTTATTGAATCGGAT
acla-5      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
agig-77     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
alim-1      TT------CTATTTTGTTTGCATATTTTATCATTTTCTATTGAATCGGAT
alya-1      TT------CTATTTTGTTTGCATAGTTTATCATTTTCTATTGAATCGGAT
amon-3      TTCTACTTCTATTTTGTTTGCATATTTTCTCATTTTTTATTGAATCGGAT
amon-5      TTCTACTTCTATTTTGTTTGCATATTTTCTCATTTTTTATTGAATCGGAT
amun-7      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
apal-1      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asp-4       TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-36     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCAGAT
asto-40     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-19     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCAGAT
asto-23     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-29     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-30     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-34     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
atri-1      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
atri-2      TTCTACTTCTATTTTGTTTGCATATTTTCTCATTTTTTATTGAATCGGAT
atri-3      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
atri-4      TTCTACTTCTATTTTGTTTGCATATTTTCTCATTTTTTATTGAATCGGAT
avin-1      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
agig-1      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
agig-5      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
agig-12     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-1      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
agig-47     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-3      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCAGAT
asto-4      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-5      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-6      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-7      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-9      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCAGAT
asto-12     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
asto-13     TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acan-2      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acan-6      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acan-7      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
acan-8      TT------CTATTTTGTTTGCGTATTTTCTCATTTTCTATTGAATCGGAT
```

```
aint-1      TCTAAAATCATTGCTTAAAGCGAAAACCCGCACAAAAGATGACTC-----
aint-2      TCTAAAATCATTGCTTAAAGCGAAAACCCGCACAAAAGATGACTC-----
pchi-1      CTTCAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
pvir-1      CTTCAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
pvir-2      CTTCAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
acan-10     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
acan-11     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
acap-7      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
acap-25     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
acap-26     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
acas-11     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
acas-3      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
acas-8      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
acla-2      TCTAAAATCATTGCTTAAAGCGGAAACCCGCACAAAAGATGACTC-----
acla-5      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
agig-77     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
alim-1      TCGAAAATCATTGCTTAAAGCGGAAACCCGCACAAAAGATGACTC-----
alya-1      TCTAAAATCATTGCTTAAAGCGGAAACCCGCACAAAAGATGACTC-----
amon-3      TCTAAAATCATTGCTTAAAGCGGAAACCCGCACAAAAGATGACTC-----
amon-5      TCTAAAATCATTGCTTAAAGCGGAAACCCGCACAAAAGATGACTC-----
amun-7      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
apal-1      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
asp-4       CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-36     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-40     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-19     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-23     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-29     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
asto-30     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
asto-34     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
atri-1      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
atri-2      TCTAAAATCATTGCTTAAAGCGGAAACCCGCACAAAAGATGACTC-----
atri-3      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
atri-4      TCTAAAATCATTGCTTAAAGCGGAAACCCGCACAAAAGATGACTC-----
avin-1      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
agig-1      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
agig-5      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
agig-12     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCCACTT
asto-1      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
agig-47     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-3      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-4      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-5      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-6      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-7      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
```

```
asto-9      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-12     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
asto-13     CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
acan-2      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
acan-6      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
acan-7      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC
acan-8      CCTAAAATCATTGCTTAACACAGAAAACCGCACAAAAGATGACTCTACTC

aint-1      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
aint-2      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
pchi-1      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
pvir-1      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
pvir-2      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acan-10     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acan-11     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acap-7      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acap-25     CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acap-26     CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acas-11     CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acas-3      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acas-8      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acla-2      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
acla-5      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
agig-77     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
alim-1      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
alya-1      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
amon-3      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
amon-5      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
amun-7      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
apal-1      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asp-4       CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-36     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-40     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-19     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-23     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-29     CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-30     CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-34     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
atri-1      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
atri-2      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
atri-3      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
atri-4      CACTTA-TAGACATTAAGGATATATATAGTATAGATCTAGCCTGACTCCA
avin-1      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
agig-1      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
agig-5      CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
```

```
agig-12     CACTTAATAGGCATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-1      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
agig-47     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-3      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-4      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-5      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-6      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-7      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-9      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-12     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
asto-13     CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acan-2      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acan-6      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acan-7      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA
acan-8      CACTTA-TAGACATTAAGAATATATA--GTATAGATCTAGCCTGGCTCCA

aint-1      CCCTCCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTATATC
aint-2      CCCTCCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTATATC
pchi-1      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
pvir-1      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
pvir-2      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acan-10     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acan-11     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acap-7      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acap-25     CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acap-26     CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acas-11     CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acas-3      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acas-8      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acla-2      CCCTGCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTAT-TC
acla-5      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
agig-77     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
alim-1      CCCTCCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTAT-TC
alya-1      CCCTTCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTAT-TC
amon-3      CCCTGCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTAT-TC
amon-5      CCCTGCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTAT-TC
amun-7      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
apal-1      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asp-4       CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-36     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-40     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-19     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-23     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-29     CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-30     CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
```

```
asto-34     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
atri-1      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
atri-2      CCCTGCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTAT-TC
atri-3      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
atri-4      CCCTGCTTACTGCATATATACTTTGACAATTCCATAATATAGTCTAT-TC
avin-1      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
agig-1      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
agig-5      CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
agig-12     CCCTCCTTACTGCATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-1      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
agig-47     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-3      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-4      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-5      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-6      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-7      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-9      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-12     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
asto-13     CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acan-2      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acan-6      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acan-7      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT
acan-8      CCCTCCTTACTACATCTATACTTTGACAATTCCATAATATAGTCTAT-TT

aint-1      TCT----------CTCCTACAACTCTGGGTTGTATATTCATATGCATTTC
aint-2      TCT----------CTCCTACAACTCTGGGTTGTATATTCATATGCATTTC
pchi-1      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
pvir-1      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
pvir-2      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
acan-10     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
acan-11     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
acap-7      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
acap-25     TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
acap-26     TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
acas-11     TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
acas-3      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
acas-8      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
acla-2      TATTCTCTTCTCTCTCCTACAACTCTAGGTTGTATATTCATATGCATTTC
acla-5      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
agig-77     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
alim-1      TGT----------CTCCTATAACTCTAAGTTGTATATTCATATGCATTTC
alya-1      TGT----------TTCCTACAACTCTAAGTTGTATATTCATATGCATTTC
amon-3      TATTCTCTTCTCTCTCCTACAACTCTAGGTTGTATATTCATATGCATTTC
amon-5      TATTCTCTTCTCTCTCCTACAACTCTAGGTTGTATATTCATATGCATTTC
amun-7      TCT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
```

```
apal-1      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
asp-4       TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-36     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-40     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-19     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-23     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-29     TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
asto-30     TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
asto-34     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
atri-1      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
atri-2      TATTCTCTTCTCTCTCCTACAACTCTAGGTTGTATATTCATATGCATTTC
atri-3      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
atri-4      TATTCTCTTCTCTCTCCTACAACTCTAGGTTGTATATTCATATGCATTTC
avin-1      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
agig-1      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
agig-5      TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
agig-12     TCT----------CTCCTACAAATCTAGGTTGTATATTCATACGCATTTC
asto-1      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
agig-47     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-3      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-4      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-5      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-6      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-7      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-9      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-12     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
asto-13     TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
acan-2      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
acan-6      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
acan-7      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC
acan-8      TAT----------CTCCTACAACTCTAGGTTGTATATTCATACGCATTTC

aint-1      T
aint-2      T
pchi-1      T
pvir-1      T
pvir-2      T
acan-10     T
acan-11     T
acap-7      T
acap-25     T
acap-26     T
acas-11     T
acas-3      T
acas-8      T
```

acla-2	T
acla-5	T
agig-77	T
alim-1	T
alya-1	T
amon-3	T
amon-5	T
amun-7	T
apal-1	T
asp-4	T
asto-36	T
asto-40	T
asto-19	T
asto-23	T
asto-29	T
asto-30	T
asto-34	T
atri-1	T
atri-2	T
atri-3	T
atri-4	T
avin-1	T
agig-1	T
agig-5	T
agig-12	T
asto-1	T
agig-47	T
asto-3	T
asto-4	T
asto-5	T
asto-6	T
asto-7	T
asto-9	T
asto-12	T
asto-13	T
acan-2	T
acan-6	T
acan-7	T
acan-8	T

REFERENCES

REFERENCES

1. Ahloowalia, B.S. 1965. A root tip squash technique for screening chromosome number in lolium. Euphytica 14:170-172.

2. Allred, K.W. 2007. Apera Adans, p. 27-, In Barkworth, M.E., Anderton, L.K., Capels, K.M., Long, S., and Piep, M.B. ed. Manual of Grasses for North America. Utah State University Press, Logan, Utah.

3. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-10

4. Amundsen, K.L. and Warnke, S.E. 2005. Poster229:Using Genographer to analyze marker data collected on an Applied Biosystems 3730 DNA analyzer. Plant and Animal Genome XIII Conference, San Diego, CA.

5. Anderson, J.A., Churchill, G.A., Autrique, J.E., Tanksley, S.D., and Sorrells, M.E. 1993. Optimizing parental selection for genetic linkage maps. Genome 36:181-6.

6. Arumuganathan, K., and Earle, E.D. 1991a. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. 9:208–218.

7. Arumuganathan, K., Tallury, S.P., Fraser, M.L., Bruneau, A.H., and Qu, R. 1999. Nuclear DNA content of thirteen turfgrass species by flow cytometry. Crop Science 39:1518-1521.

8. Barkworth, M.E. 2007. Polypogon Desf, In Barkworth, M.E., Anderton, L.K., Capels, K.M., Long, S., and Piep, M.B. ed. Manual of Grasses for North America. Utah State University Press, Logan, Utah.

9. Barriel, V. 1994. Phylogénies moléculaires et insertions-délétions de nucléotides. C. R. Acad. Sci. Paris ser. 3 317:693–701.

10. Bashaw, E.C. and C.R. Funk. 1987. Apomictic grasses, p. 40-82. In W.R. Fehr ed. Principles of cultivar development. Vol. 2 Crop sciences. Macmillan Publishing Co., New York.

11. Beard, J.B. 1973. Creeping Bentgrass, p. 71-78, In Turfgrass: Science and Culture. Prentice-Hall, Inc., Englewood Cliffs, NJ.

12. Benham, J., J-U. Jeung, M. Jasieniuk, V. Kanazin, and T. Blake. 1999. Genographer: A graphical tool for automated AFLP and microsatellite analysis. J. Agric. Genomics 4: article3

13. Belanger, F.C., Bonos, S., and Meyer, W.A. 2004. Dollar spot resistant hybrids between creeping bentgrass and colonial bentgrass. Crop Sci. 44:581-586.

14. Belanger, F.C. Meagher, T.R., Day, P.R., Plumley, K., and Meyer, W. A. 2003a. Interspecific hybridization between Agrostis stolonifera and related Agrostis species under field conditions. Crop Sci 43:240-246

15. Belanger, F.C., Plumley, K.A., Day, P.R., and Meyer, W. A. 2003b. Interspecific hybridization as a potential method for improvements of Agrostis species. Crop Sci 43:2172-2176.

16. Bonos, S.A., Clarke, B.B., and Meyer, W.A. 2006. Breeding for Disease Resistance in the Major Cool-Season Turfgrasses. Annu. Rev. Phytopathol. 44:213-34.

17. Bonos, S.A., Plumley, K.A., and Meyer, W.A. 2002. Ploidy Determination in Agrostis Using Flow Cytometry and Morphological Traits. Crop Sci 42:192-196.

18. Brilman, L.A. 2001. Utilization of interspecific crosses for turfgrass improvement. Int. Turfgrass Soc. Res. J. 9:157-161.

19. Brummer, E.C., Cazcarro, P.M., and Luth, D. 1999. Ploidy determination of alfalfa germplasm accessions using flow cytometry. Crop Sci 39:1202-07

20. Bureau, T.E. and Wessler, S.R. 1992. Tourist: a large family of small inverted repeat elements frequently associated with maize genes. Plant Cell. Oct;4(10):1283-94.

21. Bureau, T.E. and Wessler, S.R. 1994a. Mobile inverted-repeat elements of the Tourist family are associated with the genes of many cereal grasses. Proc Natl Acad Sci U S A. Feb 15;91(4):1411-5.

22. Bureau, T.E. and Wessler, S.R. 1994b. Stowaway: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell. Jun;6(6):907-16.

23. Bureau, T.E., Ronald, P.C., and Wessler, S.R. 1996. A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. Proc Natl Acad Sci U S A. 93(16):8524-9.

24. Casa, A.M., Brouwer, C., Nagel, A., Wang, L., Zhang, O., Kresovich, S., and Wessler, S.R. 2000. Inaugural article: the MITE family heartbreaker (Hbr): molecular markers in maize. PNAS 97(18):10083-9.

25. Casa, A.M., Nagel, A., and Wessler, S.R. 2004. MITE Display, p. 175-188. In Miller, W.J. and Capy, P., ed. Methods in Molecular Biology, v. 260: Mobile Genetic Elements: Protocols and Genomic Applications. Humana Press, Totowa, N.J.

26. Caceres, M.E., Pupilli, F., Piano E., and Arcioni, S. 2000. RFLP markers are an effective tool for the identification of creeping bentgrass (Agrostis stolonifera L.) cultivars. Genet. Res. Crop Evol. 47:455-459.

27. Casler, M.D. and Duncan, R.R. 2003. Introduction, p. 3-26, In Casler, M.D. and Duncan, R.R., ed. Turfgrass biology, genetics, and breeding. John Wiley & Sons, Hoboken, N.J.

28. Castiglioni, P., Ajmone-Marsan, P., Van Wijk, R., and Motto, M. 1999. AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theor. Appl. Genet. 99(3):425-431.

29. Chakraborty, N., Bae, J., Warnke, S., Chang, T. and Jung, G. 2005. Linkage map construction in allotetraploid creeping bentgrass (Agrostis stolonifera L.). Theor. Appl. Genet. 111(4):795-803.

30. Després, L., Gielly, L., Redoutet, B., and Teberlet, P. 2003. Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. Mol Phylogenet Evol. 27(2):185-96.

31. Dice, L.R. 1945. Measures of the amount of ecologic association between species. Ecology 26:297–302

32. Döring, E, Schneider, J, Hilu, K.W., and Roser, M. 2007. Phylogenetic relationships in the Aveneae/Poeae complex (Pooideae, Poaceae). Kew Bulletin 62:407-424.

33. Doyle, J.J. and Doyle, J.L. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. Phytochem Bull. 19:11-15.

34. Doyle, J. J., Doyle, J. L., Rauscher, J. T., and Brown, A. H. D. 2003. Diploid and *polyploidy* reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). New Phytol. 161: 121-132

35. Emshwiller, E. 2002. Ploidy levels among species in the 'Oxalis tuberosa Alliance' as inferred by flow cytometry. Annals of Botany 89:741-53.

36. Feliner, G.N. and Rosselló, J.A. 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Molecular Phylogenetics and Evolution 44:911-919.

37. Feschotte, C., Zhang, X, and Wessler, S.R. 2002. Miniature inverted-repeat transposable elements and their relationship to established DNA transposons, p. 1147-1158, In Craig, N.L., Craigie, R., Gellert, M, and Lambowitz, A.M., ed. Mobile DNA II. ASM Press, Washington, D.C.

38. Fortune, P.M., Pourtau, N., Viron, N., and Ainouche, M.L. 2008. Molecular phylogeny and reticulate origins of the polyploid Bromus species from section Genea (Poaceae). Am. J. Bot. 95:454-64.

39. Fulton, T.M., Van der Hoeven, R., Eannetta, N.T., and Tanksley, S. D. 2002. Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14(7):1457-67.

40. Ge, S., Li, A., Lu, B.-R., Zhang, S.-Z., and Hong, D.-Y. 2002. A phylogeny of the rice tribe Oryzeae (Poaceae) based on matK sequence data. Am. J. Bot. 89:1967-72.

41. Ge, S., Sang, T., Lu, B.-R., and Hong, D.-Y. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc Nat Acad Sci 96(25):14400-14405.

42. Gillespie, L.J., Archambault, A., and Soreng, R.J. 2006. Phylogeny of Poa (Poaceae) based on trnT-trnF sequence data: major clades and basal relationships. Aliso 23:420-34.

43. Girardin, H., Latgé, J.P., Srikantha, T., Morrow, B. and Soll, D.R. 1993. Development of DNA probes for fingerprinting Aspergillus fumigatus. J. Clin. Microbiol. 31, 1547-1554.

44. Goldman, J.J. and Sims, P.L. 2005. Production of an interspecific hybrid between Texas and Argentine bluegrass. Plant Breeding. 124:419-420

45. Grover, C.E., Kim, H, Wing, R.A., Paterson, A.H., and Wendel, J.F. 2008. Incongruent patterns of local and global genome size evolution in cotton. Genome Research 14:1474-82.

46. Guo, Y-L. and Ge, S. 2005. Molecular phylogeny of Oryzeae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. Am. J. Bot. 92(9):1548-58.

47. Hancock, J.F. 1992. Chromasomal sequence alterations, p. 16-27, In Plant evolution and the origin of crop species. Prentice Hall, Englewood Cliffs, NJ.

48. Harvey, M.J. 2007. Agrostis, p. 148-154, In Barkworth, M.E., Anderton, L.K., Capels, K.M., Long, S., and Piep, M.B. ed. Manual of Grasses for North America. Utah State University Press, Logan, Utah.

49. Holligan, D. Zhang, X, Jiang, N, Pritham, E.J., and Wessler, S.R. 2006. The transposable element landscape of the model legume Lotus japonicus. Genetics 174:2215-28.

50. Hollman, A.B., Stier, J. C., Casler, M. D., Jung, G., and Brilman, L. A. 2005. Identification of putative velvet bentgrass clones using RAPD markers. Crop Sci 45:923-930.

51. Huelsenbeck, J. P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754-755.

52. Huson, D.H. and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23(2):254-267.

53. Huson, D.H. and Kloepper, T.H. 2005. Computing recombination networks from binary sequences. Bioinformatics 21(suppl. 2):ii159-ii165.

54. Huson D.H., Kloepper, T., Lockhart, P.J. and Steel, M.A. 2005. Reconstruction of Reticulate Networks from Gene Trees, Proceedings of RECOMB, Springer, Heidelberg, 2005, LNBI 3500, p 233-249

55. Jensen, L.B., Holm, P.B., and Lübberstedt, T. 2007. Cross-species amplification of 105 Lolium perenne SSR loci in 23 species within the Poaceae. Molecular Ecology Notes 7:1155-1161.

56. Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S.R., McCouch, S.R. and Wessler, S.R. 2003. An active DNA transposon family in rice. Nature 421:163–167.

57. Jones, K. 1955. Species differentiation in Agrostis. J. Genet. 54:370-376.

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

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

60. Jones K. 1956c. Species differentiation in Agrostis III. Agrostis gigantea Roth. and its hybrids with A. tenuis Sibth. and A. stolonifera L. J Genet 54:394–399

61. Kalloo, G. 1992. Utilization of wild species, p. 149-167, In Kalloo, G. and Chowdhury J.B. eds. Distant hybridization of crop plants. Springer-Verlag, New York, New York.

62. Kang, B.C., Nahm, S.H., Huh, J.H., Yoo, H.S., Yu, J.W., Lee, M.H., and Kim, B.-D. 2001. An interspecific (Capsicum annuum x C. Chinese) F2 linkage map in pepper using RFLP and AFLP markers. Thero. Appl. Genet. 102(4):531-9.

63. Karlsen, Å.K. and Steiner., J.J. 2007. Scandinavian colonial bentgrass diversity described by RAPD, variable chlorophyll fluorescence, and collecting site ecogeography. Acta Agriculturae Scandinavica, Section B - Plant Soil Science 57:23-34.

64. Kavar, T., Meglic, V., and Rozman, L. 2007. Diversity of Slovenian maize (Zea mays) populations by Hbr (MITE) markers and morphological traits. Genetika. 43(9):1190-7

65. Kellogg, E.A. 1998. Relationships of cereal crops and other grasses. Proceedings of the National Academy of Sciences 95(5):2005-10.

66. Kikuchi, K., Terauchi, K., Wada, M., and Hirano, H.Y. 2003. The plant MITE mPing is mobilized in anther culture. Nature 421:167–170.

67. Kopecky, D., Lukaszewski, AJ., Gibeault, V. 2005. Reduction of ploidy level by androgenesis in intergeneric Lolium-Festuca hybrids for turf grass breeding. Crop Science 45:274 – 281

68. Kuoh, C-S. 2003. Agrostis dimorpholemma (Gramineae), a new record from souther Taiwan and it's morphological variation. Taiwania. 48(2):94-98

69. Linder, C.R., and Rieseberg, L.H. 2004. Reconstructing patterns of reticulate evolution in plants. American Journal of Botany 91(10):1700-08.

70. Lyons, M., Cardle, L., Rostoks, N., Waugh, R. and Flavell, A.J. 2008. Isolation, analysis and marker utility of novel miniature inverted repeat transposable elements from the barley genome. Molecular Genetics and Genomics 280(4):275-85.

71. McBreen, K and Lockhart, PJ. 2006. Reconstructing reticulate evolutionary histories of plants. Trends Plant Sci. 11(8):398-404.

72. McClintock, B. 1950. The origin and behavior of mutable loci in maize. Proc. Nat. Acad. Sci. 36(6):344-55.

73. Meudt, H.M. and Clarke, A.C. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends Plant Sci. 12(3):106–17.

74. Meyers, B.C., Tingey, S.V., Morgante, M. 2001. Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. Genome Res 11: 1660–1676

75. Monteith, J. 1930. Classification of redtop and the common bent grasses. Bull. U.S. Golf Assoc., Green Comm. 10(3):44-54.

76. Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283-92.

77. Olson, M., Hood, L., Cantor, C., and Doststein, D. 1989. A common language for physical mapping of the human genome. Science 245: 1434-1435.

78. Paran, I. and Michelmore, R.W. 1993. Development of reliable PCR based markers to downy mildew resistance genes in Lettuce. Theor. Appl. Genet. 85: 985-993.

79. Park, K-C, Lee, J.K., Kim, N-H, Shin, Y-B, Lee, J-H, and Kim, N-S. 2003. Genetic variation in Oryza species detected by MITE-AFLP. Genes Genet. Syst. 78:235-243.

80. Phillips, T.D. 2007. Breeding Turfgrasses, p. 203-211, In Pessarakli, P. and Pessarakli, M ed. Handbook of Turfgrass Management and Physiology. CRC Press, Boca Raton, Fl.

81. Posada, D and Crandall, K.A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics. 14(9):817-8.

82. Quintanar, A., Castroviejo, S., and Catalán, P. 2007. Phylogeny of the tribe Aveneae (Pooideae, Poaceae) inferred from plastid trnT-F and nuclear ITS sequences. AM. J. Botany 94:1554-1569.

83. Ramsey, J. and Schemske, D.W. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annu. Rev. Ecol. Syst. 29:467-501.

84. Raymond, O., Piola, F., Sanlaville-Boisson, C. 2002. Inference of reticulation in outcrossing allopolyploid taxa: caveats, likelihood and perspectives. Trends in Ecology and Evolution 17(1):3-6

85. Reichman, JR, Watrud, LS, Lee, EH, Burdick, CA, Bollman, MA, Storm, MJ, King, GA, and Mallory-Smith, C. 2006. Establishment of transgenic herbicide-resistant creeping bentgrass (Agrostis stolonifera L.) in nonagronomic habitats. Mol. Ecology 15:4243-4255.

86. Rohlf, F.J. (2000) NTSYS-pc numerical taxonomy and multivariate analysis system, version 2.20e. Exeter Publication, New York

87. Ronquist, F. and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

88. Rotter, D., A.K. Bharti, H.M. Li, C. Luo, S.A. Bonos, S. Bughrara, G. Jung, J. Messing, W.A. Meyer, S. Rudd, S.E. Warnke, and F.C. Belanger. 2007. Analysis of EST sequences suggests recent origin of allotetraploid colonial and creeping bentgrasses. Mol Genet Genomics 278:197-209.

89. Rozen, S. and Skaletsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers, p 365-38, In S. Krawetz, S. Misener, eds. Bioinformatics Methods and Protocols in the series Methods in Molecular Biology. Humana Press, Totowa, NJ

90. Sang, T. 2002 Utility of low-copy nuclear gene sequences in plant phylogenetics. Crit Rev Biochem Mol Bio 37(3):121-47.

91. Sang, T., Crawford, D.J., and Stuessy, T.F. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). American Journal of Botany 84(9):1120-36.

92. Sang T., Crawford, D.J., and Stuessy, T.F. 1995. Documentation of reticulate evolution in peonies (Paeonia) using internal trasnscribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. PNAS 92:6813-17.

93. Sang, T., Donoghue, M.J., and Zhang, D. 1997. Evolution of alcohol dehydrogenase genes in peonies (Paeonia): phylogenetic relationships of putative nonhybrid species. Mol Bio Evol 14:994-1007.

94. Sang, T. and Zhang, D. 1999. Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five Paeonia species based on Adh gene phylogenies. Systematic Botany 24(2):148-63.

95. Sanmiguel, P. and Bennetzen, J.L. 1998. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann. Bot. (Lond.) 82:37–44.

96. Scheef, E.A., Casler, M., and Jung, G. 2003. Development of species-specific SCAR markers in Bentgrass. Crop Sci 43:345-349.

97. Senchina, D.S., Alvarez, I, Cronn, R.C., Liu, B., Rong, J., Noyes, R.D., Paterson, A.H., Wing, R.A., Wilkins, T.A., and Wendel, J.F. 2003. Rate variation among nuclear genes and the age of polyploidy in Gossypium. Mol Biol Evol 20:633-43.

98. Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E., and Small, R.L. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92:142-66.

99. Shaw, J., Edgar, L., Edward, S., and Randall, S. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am. J. Bot. 94:275-88.

100. Shildrick, J.P. 1976. Highland bent: a taxonomic problem. J. Sports Turf Res. Inst. 52:142-150.

101. Simmons, M.P., Zhang, L.-B., Webb, C.T., and Müller, K. 2007. A penalty of using anonymous dominant markers (AFLPs, ISSRs, and RAPDs) for phylogenetic inference. Mol. Phyl. Evol. 42(2):528-42.

102. Small, R.L., Ryburn, J.A., Cronn, R.C., Seelanan, T., and Wendel, J.F. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. Am. J. Bot. 85:1301-15.

103. Sneath, P.H.A. and Sokal, R.R. 1973. Numerical Taxonomy. Freeman, San Francisco, CA.

104. Soreng, R.J. and Davis, J.I. 1998. Phylogenetics and character evolution in the grass family (Poaceae): simultaneous analysis of morphological and chloroplast DNA restriction site character sets. Botanical Review 64:1-85.

105. Soreng, R.J., Davis, J.I., and Doyle, J.J. 1990. A phylogenetic analysis of chloroplast DNA restriction site variation in Poaceae subfam. Pooideae. Pl. Syst. Evol. 172:83-97.

106. Soreng, R.J., Davis, J.I., and Voionmaa, M.A. 2007. A phylogenetic analysis of Poaceae tribe Poeae sensu lato based on morphological characters and sequence data from three plastid-encoded genes: evidence for reticulation, and a new classification for the tribe. Kew Bulletin 62:425-454.

107. Stalker, H.T. 1980. Utilization of wild species for crop improvement. Adv. Agron. 33:111-47.

108. Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

109. Thomas, H., and M.O. Humphreys. 1991. Progress and potential of interspecific hybrids of Lolium and Festuca. J. Agric. Sci. (Cambridge) 117:1–8

110. Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties, and weighted matrix choice. Nucleic Acids Research 22(22):4673-80.

111. Tu, Z. 2001. Eight novel families of miniature inverted repeat transposable elements in the African malaria mosquito, Anopheles gambiae. Proc. Natl. Acad. Sci. USA 98, 1699-1704.

112. Tuna, M., Vogel, K.P., Arumuganthan, K., and Gill, K.S. 2001. DNA content and ploidy determination of bromegrass germplasm accessions by flow cytometry. Crop Science 41:1629-34.

113. Turcotte, K., Srinivasan, S., and Bureau, T. 2001. Survey of transposable elements from rice genomic sequences. Plant J 25:169–179

114. Turgeon, A.J. 1996. Turf Quality, p. 4-12 and Turfgrass Species, 43-84, In Turfgrass Management. Prentice Hall, Upper Saddle River, NJ.

115. Van Dijk, G.E. and Winkelhorst, G.D. 1981. Interspecific crosses as a tool in breeding Poa pratensis L. 1. Poa longifolia Trin × P. pratensis L. Euphytica 31(1):215- 223.

116. Vargas, J.M. 1994. Dollar spot, p. 23-27, In Management of Turfgrass Diseases Second Ed. CRC Press, Boca Raton, FL.

117. Vergara, G.V. and Bughrara, S.S. 2003. AFLP analyses of genetic diversity in bentgrass. Crop Sci 43:2162-2171.

118. Vergara, G.V. and Bughrara, S.S. 2004. Genetic differentiation of tetraploid creeping bentgrass and hexaploid redtop bentgrass genotypes by AFLP and their use in turfgrass breeding. Crop Sci 44:884-890.

119. Vicient, C.M., Suoniemi, A., Anamthawat-Jónsson, K., Tanskanen, J., Beharav, A., Nevo, E., and Schulman, A.H. 1999. Retrotransposon BARE-1 and its role in genome evolution in the genus hordeum. Plant Cell 11(9):1769-84.

120. Vos, P., Hogers, R., Bleeker, M., Reijans, M., van del Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., and Kuiper, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23(21):4407-4414.

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

122. Wang, Y., Reighard, G.L., Layne, D.R., Abbott, A.G., and Huang, H. 2005. Inheritance of AFLP Markers and their use for genetic diversity analysis in wild and domesticated pawpaw [Asimina triloba (L.) Dunal]. J. Amer. Soc. Hort. Sci. 130(4):561-568.

123. Warnke, S. 2003. Creeping bentgrass (Agrostis stolonifera L.), p. 175-185, In Casler, M.D. and Duncan, R.R., ed. Turfgrass biology, genetics, and breeding. John Wiley & Sons, Hoboken, N.J.

124. Warnke, S. E., Douches, D. S., and Branham, B. E. 1997. Relationships among creeping bentgrass cultivars based on isozyme polymorphisms. Crop Sci. 37(1):203-207.

125. Watson, L. 1990. The grass family. Poaceae, p. 1-31, In Chapman, G.P. ed. Reproductive versatility in the grasses. Cambridge University Press, Cambridge, MA.

126. Wessler, S.R., Bureau, T.E., and White, S.E. 1995. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. Curr Opin Genet Dev. 5:814-21.

127. Wessler, S.R., Nagel, A., and Casa, A. 2001. Miniature inverted repeat transposable elements help to create diversity in maize and rice. Rice Genetics IV. pp. 107- 116. International Rice Research Institute, Philippines.

128. Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531-6535.

129. Xu, D.H. and Ban, T. 2004. Phylogenetic and evolutionary relationships between Elymus humidus and other Elymus species based on sequencing of non-coding regions of cpDNA and AFLP of nuclear DNA. Theor. Appl. Genet. 108(8):1443-8.

130. Xu, J.H. and Messing, J. 2006. Maize haplotypes with a helitron-amplified cytidine deaminase gene copy. BMC Genet 7:52

131. Yang, G. and Hall, T.C. 2003a. MAK, a computational tool kit for automated MITE analysis. Nucleic Acids Research 31(13):3659-65.

132. Yang, G. and Hall, T.C. 2003b. *MDM*-1 and *MDM*-2: Two *Mutator*-derived MITE families in rice. J. Mol. Evol. 56:255-264.

133. Yang, Z-N, Ye, X-R, Molina, J, Roose, M.L. and Mirkov, T.E. 2003. Sequence analysis of a 282-kilobase region surrounding the citrus tristeza virus resistance gene (Ctv) locus in Poncirus trivoliata L. Raf. Plant physiology 131:482-92.

134. Yeater, K.M., Bollero, G.A., Bullock, D.G., and Rayburn, A.L. 2004. Flow cytometric analysis for ploidy level differentiation of 45 hairy vetch accessions. Annals of Applied Biology 145(1):123-7.

135. Zhao, H., Bughrara, S.S., and Oliveira, J.A. 2006. Genetic diversity in colonial bentgrass (Agrostis capillaris L.) revealed by EcoRI-MseI and PstI-MseI AFLP markers. Genome 49:328-35.

136. Zhao, H., Bughrara, S.S.,and Wang, Y. 2007. Cytology and pollen grain fertility in creeping bentgrass interspecific and intergeneric hybrids. Euphytica 156:227-35.

CURRICULUM VITAE

Keenan L. Amundsen graduated from Potterville High School, Potterville, Michigan, in 1991. He received his Bachelor of Science in Mathematics from Alma College, Alma, Michigan in 1998 and his Bachelor of Science in Crop and Soil Sciences with an emphasis in turfgrass management from Michigan State University, East Lansing, Michigan, also in 1998. He completed his Master of Science in Crop and Soil Sciences with an emphasis in Plant Breeding and Genetics at Michigan State University in 2003 under the supervision of Dr. Michael Thomashow. He has been employed with the United States Department of Agriculture as a research technician since 2003.