THE USE OF MOLECULAR GENETICS TOOLS TO COMPLEMENT A
TRADITIONAL FIELD BASED TURFGRASS BREEDING PROGRAM

by

JOSHUA ANDREW HONIG

A Dissertation submitted to the
Graduate School-New Brunswick
Rutgers, The State University of New Jersey
In partial fulfillment of the requirements
For the degree of
Doctor of Philosophy
Graduate Program in Plant Biology
Written under the direction of
William A. Meyer
And approved by

________________________
________________________
________________________
________________________

New Brunswick, New Jersey
JANUARY, 2011
ABSTRACT OF THE DISSERTATION

The Use of Molecular Genetics Tools to Complement a Traditional Field Based Turfgrass Breeding Program

By JOSHUA ANDREW HONIG

Dissertation Director:
William A. Meyer

Despite being important agricultural crops, turfgrasses lack significant molecular genetic resources. The objectives of the current studies were to develop molecular marker resources for genotyping Kentucky bluegrass (*Poa pratensis* L.) and for the creation of a genetic linkage map in creeping bentgrass (*Agrostis stolonifera* L.) Kentucky bluegrass is an important facultative apomictic temperate perennial grass species utilized for both forage and cultivated turf. Through apomixis, this species is able to propagate diverse and odd ploidy levels, resulting in many genetically distinct phenotypes. A wide range of diverse cultivars and accessions of Kentucky bluegrass have been previously characterized based on common turf performance or morphological characteristics, as well as by random amplified polymorphic DNA (RAPD) markers. In the current studies, 265 Kentucky bluegrass cultivars, experimental selections, collections, and hybrids were genotyped using microsatellite (SSR) markers. Results based on SSR markers were compared with the original Kentucky bluegrass classification system based on pedigree, common turf performance and morphological characteristics.
All cultivars, experimental selections, collections, and hybrids were uniquely identified with the current set of SSR markers. Genetic relationships of individuals as assessed by SSR markers closely matched known pedigrees. Furthermore, genetic relationships based on SSR markers more accurately reflected pedigree than genetic relationships based on morphological characteristics. The current set of SSR markers can be used to rapidly genotype and assign new cultivars/accessions to Kentucky bluegrass classification types and assess genetic relatedness among individuals, and should be considered for use in a Kentucky bluegrass Plant Variety Protection program. Creeping bentgrass is the most widely utilized cool-season turf species for intensively managed sports playing surfaces such as bowling greens and golf course putting greens, tees, and fairways. In the current study, we have constructed the first PCR marker-based genetic linkage map of creeping bentgrass, using SSR, AFLP, CISP, and ILP markers. The latter two marker types are important because they can be used to assess syntenous relationships between orphan crops lacking significant molecular genetic resources, such as creeping bentgrass, and model cereal crops such as rice.
Acknowledgements and Dedication

I would like to thank my advisor, Dr. William A. Meyer, for his wonderful help and guidance during my studies. Simply, he is the best advisor that a student could have. I would also like to thank my entire committee for their time, effort, and instruction throughout my academic career. Much of what I have learned, I have learned from them. My sincerest appreciation goes to the Rutgers Center for Turfgrass Science for financial support throughout my graduate studies. This dissertation is dedicated to my beautiful daughter, Sadie.
TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION ................................................................. ii

ACKNOWLEDGEMENTS AND DEDICATION .................................................. iv

TABLE OF CONTENTS ..................................................................................... v

LIST OF TABLES .............................................................................................. xi

LIST OF FIGURES ............................................................................................ xii

INTRODUCTION ............................................................................................... 1

CHAPTER 1: Review of the literature .............................................................. 4

  Introduction ................................................................................................. 4

  Genetic markers ........................................................................................ 8

  Types of DNA-based markers in plant breeding ......................................... 9

    Restriction fragment length polymorphisms (RFLP) ......................... 13

    Random amplified polymorphic DNA (RAPD) ............................... 14

    Amplified fragment length polymorphisms (AFLP) ....................... 17

    Sequence Related Amplified Polymorphisms (SRAPs) ............... 18

    Target Region Amplification Polymorphism (TRAPs) ................. 19

    Simple sequence repeats (SSR) ........................................................... 20

    Expressed sequence tagged simple sequence repeats (EST-SSRs) .... 35

    Single Nucleotide Polymorphisms (SNPs) ......................................... 39

‘Anchored’ primers for comparative genetics/mapping – EPIC-PCR (Exon Primed Intron Crossing), ILP markers (Intron Length Polymorphisms), COS markers (Conserved Ortholog Set), CATS
Morphological data collection………………………………….292
Microsatellite development and genotyping…………………...293
Morphological Data analysis…………………………………...294
Microsatellite Data analysis…………………………………….295
Results………………………………………………………………….296
Discussion………………………………………………………………305
Literature cited………………………………………………………….308

CHAPTER 5: PCR marker- based genetic linkage map in allotetraploid creeping bentgrass (Agrostis stolonifera L.)……………………………………..312
Introduction…………………………………………………………….312
Materials and methods…………………………………………………316
  Plant material and mapping population………………………….316
  SSR development and analysis……………………………………317
  CISP and ILP analysis………………………………………….319
  AFLP analysis………………………………………………….320
  SSR, CISP, and ILP allele scoring………………………………320
Linkage Analysis and Map Construction…………………………322
Examination of Chromosomal Pairing Behavior………………323
Combined population from reciprocal cross…………………...324
Identification of homologous and homoeologous groups in pooled population…………………………………………………………324
Results………………………………………………………………….325
  AFLP Markers………………………………………………….325
LIST OF TABLES

1. C-series creeping bentgrasses.................................................................139
2. Examples of improved creeping bentgrass varieties developed in the 1970’s through 1990’s.................................................................144
3. Primer sequences and characteristics of 88 Poa pratensis (L.) microsatellite markers obtained from the cultivar ‘Cabernet’, tested on 265 Poa pratensis cultivars, experimental selections, collections, and hybrids.....................223
4. Kentucky bluegrass classification types, brief description of types, and example cultivars..................................................................................238
5. Kentucky bluegrass cultivars, experimental selections, collections, and hybrids genotyped by 80 microsatellite markers in the current study....................243
6. Revised Kentucky bluegrass classification system based on a combination of pedigree, UPGMA cluster analysis, model-based clustering analysis, and principle coordinate analysis (PCO) of microsatellite data.........................................250
7. Analysis of molecular variance (AMOVA) of 265 Kentucky bluegrass individuals from 12 different populations (classification types) based on 80 microsatellite markers.................................................................262
8. Pair-wise $\Phi_{ST}$ values calculated by AMOVA illustrating differences between populations (classification types) of Kentucky bluegrass.........................264
9. Analysis of Variance (ANOVA) of cultivar and year effects on various morphological measurements of Kentucky bluegrass cultivars and experimental selections measured on spaced plants in 2004 and 2005 at the Rutgers University Plant Biology and Pathology Research and Extension Farm in Adelphia, NJ...297
LIST OF FIGURES

1. Phylogeny of Kentucky bluegrass cultivars, experimental selections, collections, and hybrids. Diagram is a UPGMA cladogram based on the average Jaccard distance of SSR marker alleles between individuals, color coded according to known pedigrees………………………………………………………………..254

2. Estimated population (classification type) structure for Kentucky bluegrass cultivars, experimental selections, collections, and hybrids………………………………………..257

3. Principle coordinate analysis (PCO) of Kentucky bluegrass cultivars, experimental selections, collections, and hybrids based on microsatellite allele data……………………………………………………………………………...260

4. An UPGMA cladogram of 172 Kentucky bluegrass cultivars using 80 SSR primer pairs……………………………………………………………………………………..298

5. An UPGMA cluster analysis of six morphological characteristics of 172 Kentucky bluegrass cultivars and experimental selections in 2004……………………………..300

6. An UPGMA cluster analysis of six morphological characteristics of 172 Kentucky bluegrass cultivars and experimental selections in 2005……………………………..302

7. Creeping bentgrass genetic linkage map…………………………………………………………329
INTRODUCTION

The Rutgers University turfgrass breeding program is the largest cool-season turfgrass breeding program in the world. The program began when the administration of what is now the Rutgers School of Environmental and Biological Sciences offered a turfgrass breeding position to Dr. C. Reed Funk in December 1961 (Funk and Meyer, 2001). The startup funding and first year budget for the position was four hundred dollars and part-time use of a university car for germplasm collection (Funk and Meyer, 2001). From this modest beginning, the program has grown to release well over 400 turfgrass cultivars and collect over 15,000 new germplasm sources from Europe and Asia by 2010 (W. A. Meyer – Personal communication). Additional accomplishments include the release of dozens of landmark cultivars, significant contributions to the discovery and understanding of turfgrass endophytes, creation of the Rutgers Center for Turfgrass Science, and the development of a significant royalty stream from the liscencing of hundreds of cultivars (W. A. Meyer – Personal communication). Interestingly, all of these breeding accomplishments have been achieved solely through classical field breeding techniques, without the use of molecular genetic tools.

Despite being important agricultural crops, turfgrasses lack significant molecular genetic resources. There are four principle reasons why this has occurred: 1) many turfgrasses are genetically complex, being out-crossing polyploids, and thus represent complicated research targets; 2) many traits of interest to turfgrass breeders and scientists are under complex control of a number of genes (quantitative traits); 3) there is a lack of well characterized genetic materials and absence of repeatable phenotyping protocols for many target traits which limits the application of genomics technologies; and 4) the
turfgrass research community has failed to attract major funding for molecular and genomics research from governmental agencies or the private sector (Zhang et al., 2006). Gresshoff et al. (1998) have described turfgrass researchers as being in a “catch-up” mode in relation to molecular genetic technologies. Additionally, Zhang and Mian (2003) summarized the situation by stating that “…many of the fundamental tools required for forage and turf to benefit fully from the revolution in genomics do not exist or are incomplete”.

Although there are valid reasons for the lack of molecular genetic tools in turfgrasses to date, more effort needs to be made to develop molecular genetic tools for turfgrass breeding programs. The use of molecular tools in plant breeding programs falls into three general categories: 1) characterization of germplasm (identification or fingerprinting); 2) genetic dissection of target traits (e. g. identification and characterization of genomic regions involved in the expression of target traits including genetic mapping, linkage analysis, and QTL analysis; and 3) implementation of marker-assisted selection. Two areas of study in the current dissertation address the first and second categories listed above. The first area of study is focused on molecular genetic characterization of Kentucky bluegrass (Poa pratensis L.) cultivars and accessions using microsatellite (SSR) markers (chapters 2, 3, and 4) and addresses the ‘characterization of germplasm’ category, while the second area of study is focused on the development of a PCR-based genetic linkage map in allotetraploid creeping bentgrass (Agrostis stolonifera L.) (chapter 5) and addresses the ‘genetic dissection of target traits’ category. These studies represent significant progress towards developing molecular genetic tools in two
turfgrass species with heretofore limited molecular genetic resources, and also serve to expand the scope of the Rutgers University turfgrass breeding program.
CHAPTER 1

Review of the Literature

Introduction

Plant breeding consists of the creation of genetic variability, selection of elite types from that variability, and synthesis of a stable cultivar from those elite selections (Dudley, 2002). It is interesting to note that plant breeding precedes our understanding of genetics, dating back to the time when primitive peoples first saved seed to plant in succeeding years. These early examples of plant domestication are older than recorded history (Mayo, 1980). In fact, it has been argued that plant (and animal) domestication has led to the development of modern civilization as we know it today (Diamond, 1999). Man has depended upon plants for food, clothing, fuel, drugs, construction materials (Allard, 1960), and even ornamentals due to their aesthetically pleasing nature or for functional recreational surfaces. Considering the many uses and importance of plants, “…it is not surprising that men have long been concerned with developing types better suited to satisfying their needs” (Allard, 1960). Plant breeding began when prehistoric man learned to look for and select plants that better satisfied those needs (Poehlman and Sleper, 1995). Although prehistoric people were practicing early forms of plant breeding by selecting and domesticating useful wild plant species, it is only recently (relative to prehistoric domestication) that plant breeding, mainly in conjunction with advances in genetics, has become a systematic scientific discipline (Allard, 1960).

Plant breeding relies on genetic variation and uses selection (natural and artificial) to improve plants for traits and characteristics that are of interest to the grower and end user. Throughout much of our recorded history this improvement has occurred as a result
of practical phenotypic selection (classical plant breeding). This phenotypic selection process began with early humans interest in larger, non-shattering seeds or fleshier and sweeter fruits and continues today with modern applied techniques such as the development and use of hybrid corn. Current applied or classical plant breeding practices have made exceptional strides in producing a continuous range of improved plant varieties. Lee (1995) writes, “On a worldwide basis, plant breeding has been one of the most successful technologies developed in modern agriculture: its methods are opportunistic and adaptable to myriad production schemes, they require relatively inexpensive input, and their products have pervasive social benefits”.

It is interesting to note that while Lee acknowledges that crop yields have dramatically increased since the early 1900’s as a result of enhanced genetic potential through classical plant breeding, this progress has been accomplished without extensive knowledge of the underlying factors at the gene, pathway or genome level. Phenotypic changes are well documented in breeding programs for all crops, but knowledge at the genotypic level and the relationship to basic biology is less well documented (Lee, 1995). Therefore, recent advances in the fields of biotechnology and molecular biology hold significant promise to enhance the tools available to the modern plant breeder to aid in solving problems that cannot be wholly addressed with traditional applied plant breeding methodologies.

One of the first and most fundamental of these newer technologies was DNA markers, which are simply differences in genetic information between individuals (Paterson et al., 1991). When this technology was first applied to plant improvement programs, Paterson et al. (1991) envisioned how this technology would enhance the
advances made in classical plant breeding: “Classical plant and animal breeding make an indispensable contribution to society, and will continue to do so for the foreseeable future. However, we feel that DNA markers can significantly accelerate many breeding endeavors. Further, DNA markers may provide approaches to some objectives which have proven difficult to achieve with classical techniques…”.

Since the early application of newer genetic technologies significant molecular genomic resources have been developed for model plant species (e.g., *Arabidopsis thaliana*) and important food crops such as rice (*Oryza sativa*), wheat (*Triticum* spp.), maize (*Zea mays*), barley (*Hordeum vulgare*), oats (*Avena sativa*), sorghum (*Sorghum* spp.), etc. These resources include large numbers of biochemical and DNA markers (GrainGenes 2.0 database, 2008), detailed genetic/physical maps (Harushima et al., 1998; Wight et al., 2003; Quarrie et al., 2005; Marcel et al., 2007), extensive collections of expressed sequence tags (ESTs) (Sasaki et al., 1994; Seki et al., 2002; Lazo et al., 2004; Kunne et al., 2005), large insert libraries (such as yeast artificial chromosomes (YAC) or bacterial artificial chromosomes (BAC)) for detailed genetic dissection (Tao et al., 1994; Choi et al., 1995), and ultimately complete sequencing of the model dicot *A. thaliana* (The Arabidopsis Genome Initiative, 2000) and the first major crop species, *O. sativa* (Goff et. al, 2002). The advances achieved with the above model species and economically important cereal crops have resulted in a more thorough understanding of the basic biology of the entire plant kingdom, as well as new powerful molecular genetics tools and methods.

Besides the model plant species (*A. thaliana* and *O. sativa*) and the major cereal genomes, the remainder of plant species have been historically underrepresented in the
new wealth of molecular genetics and genomics approaches outlined above. In terms of molecular genetics and genomics research effort, all plants, except the few species for which significant molecular genetic resources are available, can be referred to as minor or ‘orphan crops’ (Naylor et al., 2004; Nelson et al., 2004; Devos, 2005; Xu et al., 2005; Jackson et al., 2006; Varshney et al., 2006). This is particularly true for forage and turfgrasses where the development of molecular and genomic resources has been limited (Zhang et al., 2006). There are several factors that likely contribute to the status of forage and turfgrasses as orphan crops [adapted from Zhang et al., 2006]: First, many grasses have large and complex genomes (polyploids); second, many important traits of interest are controlled by numerous genes (quantitative) as opposed to single genes (qualitative); third, forage and turfgrasses lack the well characterized genetic stocks and genetic resources available in the model species; and finally, the forage and turfgrass research community has failed to attract major funding from governmental agencies or the private sector to conduct comprehensive molecular genetic analysis on any forage or turfgrass species. Despite the above mentioned obstacles there have been some efforts made by various researchers to develop molecular and genomics tools in these grasses [reviewed in Zhang et al., 2006]. Furthermore, advances in molecular genetics and genomics in model crops provide an opportunity for efficient transfer of information and technology to orphan crops (Nelson et al., 2004). As an example, within plant families, and particularly within Poaceae, many taxa exhibit synteny, where similar genes and gene order along stretches of chromosomes remain conserved to varying degrees (Bonierbale et al., 1988; Chao et al., 1989; Tanksley et al., 1992; Ahn and Tanksley, 1993; Bennetzen and Freeling, 1993; Moore et al., 1995; Devos and Gale, 1997; Bennetzen and Freeling,
Therefore, initial investment in primary molecular genetic resources for orphan crops can be supplemented with genomic data and resources from model crops, thereby using syntenic relationships as a foundation to achieve more rapid progress on lesser-studied species (Naylor et al., 2004; Nelson et al., 2004; Paterson et al., 2005; Xu et al., 2005).

Genetic markers

Genetic markers represent genetic differences between individual organisms, populations, species etc. Markers do not necessarily represent a target gene for a trait of interest, but rather act as ‘signposts’ or tags for a broad to narrow candidate region of a chromosome that is associated or linked with the trait under study. Markers occupy a genomic position on a chromosome just as genes do, and can therefore be referred to as ‘loci’ (Collard et al., 2005).

There are three major types of genetic markers [adapted from Collard et al., 2005]: 1) morphological or visual markers representing a particular phenotype such as flower color, seed size or shape, disease resistance/susceptibility etc.; 2) biochemical markers such as proteins and enzymes (e.g. isozymes represent differences in an enzymatic profile between individuals); and 3) DNA or molecular markers which represent differences at the DNA sequence level. As discussed previously, phenotypic markers have a long history of use in plant domestication and plant breeding. Both visible markers and isozymes were used to distinguish individuals and to create genetic
maps through the late 1970’s (Paterson et al., 1991). In 1980, Botstein et al., proposed using differences in DNA itself (utilizing restriction fragment length polymorphisms – RFLPs) to study inheritance patterns in humans. Prior to this time, progress using genetic markers in systematic scientific experiments had been slowed due to the paucity of morphological and enzymatic markers available to researchers. While in principle it is possible to use morphological and enzymatic markers, there are a far greater number of DNA markers that can be readily found and put into practice (Paterson et al., 1991).

**Types of DNA-based markers in plant breeding**

A significant review of the types of DNA-based markers used in plant breeding programs is provided by Gupta et al. (1999). These authors classify the types of markers into three groups primarily based on methodology:

1) Hybridization-based markers such as restriction fragment length polymorphisms (RFLP).

2) PCR-based markers such as random amplified polymorphic DNAs (RAPDs), microsatellites or simple sequence repeats (SSRs), sequence tagged sites (STS), amplified fragment length polymorphisms (AFLPs), inter-simple sequence repeats (ISSRs), and cleaved amplified polymorphic sequences (CAPs).

3) DNA chip and comparative sequencing-based markers such as single nucleotide polymorphisms (SNPs).
The above review and classification was almost all-inclusive at the time of publication, however, there have been advances in technology and methods that have added to the above list. As an example, additional PCR based techniques such as sequence related amplified polymorphisms (SRAPs) (Li and Quiros, 2001) and conserved intron spanning primers (CISPs) (Feltus et al., 2006) have since been developed. Additionally, PCR techniques and DNA chip technology are constantly changing to reflect recent technological breakthroughs or allow for inclusion of a greater number of testable oligonucleotides and higher throughput, while newer technical adaptations of heteroduplex analysis, single stranded conformational polymorphism (SSCP), and advanced sequencing platforms can be used for SNP identification.

Another consideration, mostly from a practical standpoint, is the importance of categorizing the quality and utility of the data obtained from various PCR based techniques. As an example, PCR procedures can be broadly defined by three major approaches: 1) arbitrary-primed PCR; 2) sequence-tagged-site (STS) PCR; and 3) a combination of the two previous approaches (Peakall, 1997). Sequence-tagged-site PCR uses two different specific primers, complimentary to opposite strands of conserved DNA, to amplify the intervening sequence (Peakall, 1997). A highly desirable feature of STS-PCR is that it produces a co-dominant single locus genetic marker [in the case of a diploid organism] (Peakall, 1997). Consequently, the DNA profiles are generally much easier to score than multi-locus profiles, with at most 2 alleles per sample and heterozygotes that are readily apparent (Peakall, 1997). Another advantage of STS-PCR is that a positive result is always achieved (except in the case of null alleles) from a successful reaction (Peakall, 1997). A disadvantage of STS-PCR is the need for a priori
sequence information from the organism under investigation to generate the site specific primers. This drawback can require significant intial outlays of money, labor, and time.

The second approach to PCR (arbitrary-primed PCR) overcomes the need for sequence knowledge, but brings its own limitations (Peakall, 1997). In 1990, Williams et al. (1990) and Welsh and McClelland (1990) showed that by using a single short primer of arbitrary (but known) sequence, under low stringency PCR conditions, multi-locus DNA profiles could be produced with polymorphisms revealed as band ‘presence’ or ‘absence’ (Peakall, 1997). Numerous acronyms for various iterations of this technique have been proposed, but the term ‘RAPDs’ has become widely used as a generic term for arbitrary-primed PCR (AP-PCR) (Peakall, 1997). The advantage of AP-PCR methods is that DNA sequence knowledge is not required, making the approach theoretically applicable to any organism (Peakall, 1997). Despite this significant advantage over STS-PCR, AP-PCR suffers from numerous drawbacks including the following outlined by Peakall (1997):

1) AP-PCR methods produce multi-locus profiles (multiple bands) that present a challenge to score objectively and consistently, because band intensities invariably differ within and among profiles, and alignment across multiple runs can be difficult.

2) Polymorphisms are only detected as band presence or absence (bi-allelic), where the heterozygote cannot be detected.

3) Run-to-run reproducibility of DNA profiles is questionable, because PCR amplification is performed under low stringency conditions allowing for the
fact that some of the products result from weak interactions between primer and DNA template that are extremely sensitive to reaction conditions. This can result in poor reproducibility for some primers and bands.

4) While some problems of AP-PCR reproducibility can be solved within a laboratory by ‘good practice’, laboratory to laboratory variation is likely to remain a problem.

The third approach to PCR combines the features of STS-PCR and AP-PCR. The common technique that utilizes this approach is AFLP (Vos et al., 1995). This technique involves multiple amplification rounds, where one amplification step amplifies restriction fragments along with ligated sequence specific adaptors from a total digest of genomic DNA (STS-PCR step utilizing ‘known’ restriction sites + ligated adaptors), and subsequent rounds ‘selectively’ amplify subsets of fragments based on utilizing the restriction sites and ligated adaptors as primers along with arbitrary extension bases (AP-PCR step). Compared with AP-PCR alone, AFLPs are performed under high stringency, and are therefore theoretically less sensitive to reaction conditions (Peakall, 1997).

The DNA-based markers briefly introduced above have been used in molecular genetic studies of hundreds of plant species. No single marker system is universally applicable to all problems or investigations, as each marker system has both strengths and weaknesses. Therefore, the choice of marker system is often a compromise that depends on the research question, the molecular genetic resources already available in the study organism, the genetic resolution needed, financial considerations, available research
equipment, and level of technical expertise needed. The more recent technological advances are usually associated with model crops, while orphan species are usually first studied using the older marker systems such as RFLPs, RAPDs, and AFLPs. With this in mind the following section will describe the technical aspects and some advantages and disadvantages of the more common marker systems used in molecular plant breeding programs in more detail, or review additional marker systems that have been developed since the review of Gupta et al. (1999).

**Restriction fragment length polymorphisms (RFLP):**

Among the DNA-based markers developed to date, RFLPs were developed first. Some of the earliest applications of these markers included mutation mapping in adenovirus (Grodzicker et al., 1974) and genomic mapping in humans (Botstein et al., 1980). The technique uses restriction enzymes (endonucleases) to cleave DNA into variable length pieces. DNA fragments are then separated by electrophoresis, transferred to a membrane (Southern blotting), and hybridized to probe DNA for detection and sizing. The advantages of RFLPs include the following: RFLPs are codominant (can distinguish homozygote versus heterozygote), highly reproducible, easily transferable between laboratories (once probe DNA is developed), and can be conserved across species and genera making them useful for phylogenetic studies (Van Deynze et al., 1998; Gupta et al., 1999; Godwin et al., 2001; Ribaut et al., 2002; Semagn et al., 2006; Zhang et al., 2006). Disadvantages include the requirement for high quantity and high quality DNA and the use of Southern blot hybridization. The hybridization procedures make RFLPs time consuming, laborious, and more technically demanding than PCR
based markers (Gupta et al., 1999; Godwin et al., 2001; Ribaut et al., 2002; Semagn et al., 2006; Zhang et al., 2006). The laborious nature of RFLPs is the greatest drawback for use of these markers in molecular breeding programs.

**Random amplified polymorphic DNA (RAPD):**

RAPDs were one of the first PCR based markers, and the first to amplify DNA fragments without *a priori* knowledge of DNA sequence information (arbitrarily primed) (Williams et al., 1990). The RAPD procedure generally uses 10 bp arbitrary primers and low temperature annealing conditions during PCR amplification. PCR amplification generates numerous DNA fragments and primer sequences can be altered resulting in changes in the DNA fragment banding pattern (Williams et al., 1990). The key advantage of RAPDs is the ability to generate DNA fragments without prior knowledge of the target sequence, making this marker system universally applicable in the study of any organism. Additional advantages include ease of use, low start-up costs, low DNA quantity needed, generation of a large number of DNA fragments, non-radioactive visualization, simple interpretation of data, and a simple primer development protocol with few design constraints (Williams et al., 1990; Gupta et al., 1999; Godwin et al., 2001; Ribaut et al., 2002; Semagn et al., 2006; Zhang et al., 2006).

Although initially attractive as a marker system when first described (simple utility, low cost, and high return), RAPDs have since been found to have significant disadvantages. The greatest drawback is the low stringency PCR conditions (low annealing temperature and to a certain extent the short-length primer), which can increase the likelihood of nonspecific binding and primer mismatch (Williams et al., 1990; Tyler
et al., 1997; Perez et al., 1998; Mueller and Wolfenbarger, 1999; Zhang et al., 2006). The clear problem in this instance is the inability of RAPD assays to distinguish between true and artificial PCR products/polymorphisms (Williams et al., 1990; Ellsworth et al., 1993; Lamboy, 1994a; Lamboy, 1994b; Tyler et al., 1997; Perez et al., 1998; Mueller and Wolfenbarger, 1999). As a result, there is legitimate concern about both intra- and inter-laboratory variation when conducting RAPD assays (Kleinhofs, A., 1992; Devos and Gale, 1992; MacPherson et al., 1993; Meunier and Grimont, 1993; He et al., 1994; McClelland and Welsh, 1994; Tyler et al., 1997). Numerous researchers have noted that PCR reaction preparations and conditions can influence the variability of RAPD results. This includes DNA quality and quantity, specific primer, type of polymerase used, magnesium concentration, type of buffer, temperature profile, and even the type of thermal cycler used (Williams et al., 1990; Arnold et al., 1991; Carlson et al., 1991; Devos and Gale, 1992; Buscher et al., 1993; Ellsworth et al., 1993; MacPherson et al., 1993; Meunier and Grimont, 1993; He et al., 1994; McClelland and Welsh, 1994; Schweder et al., 1995). These results all point to significant intra-laboratory variation if all preparatory and reaction conditions are not rigorously controlled. In a study designed to assess the repeatability of RAPDs within a laboratory, Perez et al. (1998) found that the fraction of repeatable bands among three study organisms was less than 40%. Additionally, a North American inter-laboratory study investigating the reproducibility of RAPDs (Penner et al., 1993) and a European inter-laboratory study investigating the reproducibility of RAPDs, AFLPs, and SSRs (Jones et al., 1997) both reported that RAPD markers were difficult to reproduce among study participants even under controlled conditions.
RAPD markers have additional limitations including problems related to sequence homology and efficacy relative to other marker systems. In regards to homology, it is often assumed that bands of apparently identical molecular weight in different individuals are the same homologous DNA fragment. In other words, it is not always correct to assume that PCR bands of the same size represent the same genetic location (locus). Therefore, genetic inferences drawn from RAPD analysis, without prior or confirming sequence information, may be false (Clark and Lanigan, 1993; van de Zande and Bijlsma, 1995; Karp et al., 1996; Karp et al., 1997; Gupta et al., 1999; Kuras et al., 2004; Semagn et al., 2006). Several researchers have indicated that it is critical to confirm that bands generated during RAPD analysis are in fact homologous (Devos and Gale, 1992; Clark and Lanigan, 1993; Gupta et al., 1999). This step would require additional hybridization, sequencing, or other appropriate confirmation which in effect defeats the original purpose of the RAPD procedure (simple utility, speed, low cost etc.).

Finally, now that additional DNA marker systems have been developed there are likely better alternatives to RAPDs in terms of efficacy. Several researchers have indicated that RAPDs are inappropriate markers for the development of linkage maps and QTL analysis in comparison to other markers (Devos and Gale, 1992; Gupta et al., 1999; Semagn et al., 2006; Zhang et al., 2006). Additionally, once AFLPs were developed (Vos et al., 1995) and SSRs and other marker systems became more commonplace, numerous authors have reported that RAPDs are not as informative or efficient for genetic diversity studies in direct comparison to other marker systems (Mueller and Wolfenbarger, 1999; Yue et al., 2002; Agrama and Tuinstra, 2003; Belaj et al., 2003;
Amplified fragment length polymorphisms (AFLP):

Amplified fragment length polymorphisms are PCR based markers for screening genetic diversity and generating linkage maps. When the technique was invented (Vos et al., 1995) the authors described the method as combining the reliability of the RFLP technique with the power of PCR. In this technique, DNA is “cut” with restriction enzymes, and double-stranded DNA adapters are ligated to the DNA. The sequences of the adapters and the adjacent restriction cut sites serve as templates or primer-binding sites for later PCR reactions. During subsequent ‘selective’ amplification, primers are designed to contain the sequences that are complimentary to those of the restriction recognition sites, the adaptors, and one to three selective extension bases added at their 3’ ends. Only a subset of the restriction fragments, with nucleotides complimentary to the selective extension bases will be amplified under stringent PCR conditions. This selective amplification still generates a large number of bands for the detection of polymorphisms.

The following are advantages of the AFLP technique: Like RAPDs, a key advantage of AFLPs is the ability to generate DNA fragments without prior knowledge of the target sequence (AFLP markers can be generated for DNA of any organism without prior sequence analysis). Additional advantages include general insensitivity to template DNA concentration, high stringency PCR reaction (in contrast to RAPDs), relatively high level of polymorphism detection per reaction, high reproducibility within and between
laboratories, high level of homology (locus specificity – with the exception of polyploids) and high multiplex ratio (simultaneous polymorphism detection using multiple primer combinations) (Vos et al., 1995; Blears et al., 1998; Gupta et al., 1999; Mueller and Wolfenbarger, 1999; Ridout and Donini, 1999; Semagn et al., 2006). Disadvantages of AFLPs include a more technically demanding procedure than RAPDs, dominant marker system (inability to distinguish homozygote from heterozygote), and the need for silver staining, autoradiography or fluorescence based detection. Silver staining and autoradiography generate hazardous waste and require specialized training, while fluorescence based detection systems incur significant initial capital expenditures.

**Sequence Related Amplified Polymorphisms (SRAPs)** (Li and Quiros, 2001):

Sequence Related Amplified Polymorphisms (SRAPs) are a relatively new marker system that combines some of the features of both RAPDs and AFLPs in a PCR based assay. This marker system was originally developed for its’ simplicity, reliability, and moderate throughput, and was shown to have utility in both mapping and gene tagging in plants (Li and Quiros, 2001). Later work has shown that SRAPs have also been useful for elucidating genetic relationships in diverse plant populations and cultivars (Budak et al., 2004b; Ferriol et al., 2004; Vandemark et al., 2006; Esposito et al., 2007; Fu et al., 2008). The method design targets coding sequences and results in the amplification of numerous codominant marker bands. Amplification of markers occurs using 17 or 18 nucleotide forward and reverse primers, each composed of three distinct regions of sequence motifs. The first region comprises 10–11 random nucleotides at the 5’ end of each primer. The second region, immediately following the first 10-11 random
nucleotides, is comprised of the sequence CCGG in the forward primer and AATT in the reverse primer. The third, and final region, is three selective nucleotides at the 3’ end of each primer. Just as in the AFLP procedure, the three selective nucleotides can be varied. Functionally, the 10-11 random nucleotides at the 5’ ends of each primer amplify arbitrary sequences, similar to the full length primers of RAPD markers. The CCGG sequence in the forward primer is included to target GC-rich exonic sequences, while the AATT sequence in the reverse primer is included to target AT-rich intronic sequences and promoter regions. This design feature is intended to “anchor” the forward primer in exonic regions of genes then prime into neighboring introns with the reverse primer. Finally, the three selective nucleotides at the 3’ ends of both primers increase selectivity similar to the 3’ selective extension nucleotides used in the selective amplification step of the AFLP procedure.

Additional unique features of the SRAP marker system include the use of two annealing temperatures during PCR amplification. An initial annealing temperature of 35°C is used for the first 5 PCR cycles. This low initial annealing temperature is used to ensure partial match binding of both primers to the target DNA (Li and Quiros, 2001). The annealing temperature is then raised to 50°C for the remaining 35 cycles, which ensures that the products amplified in the first 5 cycles are efficiently and consistently amplified during the remaining PCR cycles (Li and Quiros, 2001).

**Target Region Amplification Polymorphism (TRAPs)** (Hu and Vick, 2003):

Target Region Amplification Polymorphism (TRAP) is a fairly new PCR marker based on similar principles to SRAP markers, but utilizes sequence-based information as
part of the primer design. One primer, the ‘fixed’ primer, of about 18 nucleotides is
designed from an EST or gene of interest from previously generated sequence data. The
fixed primer is then paired with an arbitrary primer that is designed with the same
principles as SRAP primer design: 10-11 arbitrary nucleotides at the 5’ end, followed by
either a GC- or AT-rich core motif, and ending with 3 or 4 ‘selective’ nucleotides at the
3’ end. The choice of a GC- or AT-rich core motif will promote annealing with an exon
or intron, respectively. The cycling condition of low temperature during early rounds of
PCR amplification and higher temperature during subsequent rounds are utilized for the
same reasons previously described for SRAP markers.

**Simple sequence repeats (SSR):**

The discovery of highly repetitive ‘satellite’ DNA sequences was a fortuitous
accident that occurred during RFLP mapping of the human genome (Wyman and White,
1980). Wyman and White identified eight alleles at a single locus within Mormon
pedigrees, and believed that they had identified a series of allelic restriction fragments
created by a transposable element. Later, Jeffreys et al. (1985) further investigated the
repetitive sequences accidentally found by White and Wyman and a few others (Higgs et
al., 1981; Bell et al., 1982; Goodbourn et al., 1983), and renamed the sequences
‘minisatellites’. The isolated minisatellites shared a core motif of 10-15 bp, were highly
repetitive, varied in length from individual to individual, and had multiple alleles at some
loci. Jeffreys and colleagues were the first to recognize the utility of these markers and
suggested that minisatellites would be useful for linkage mapping, forensics, and
detecting inbreeding. However, a few years later it was found that ‘minisatellites’ were
generally clustered in telomeres along chromosomes (Royle et al., 1988; Armour et al., 1989). This made minisatellites less useful for genetic mapping and other studies due to nonrandom distribution within the genome.

At the time that research was being conducted to characterize minisatellites, Hamada et al. (1982) were the first to describe novel repeat elements composed of either poly(dT-dG), poly(dC-dA), or poly(dG-dC). Hamada and colleagues found these repeat elements in a number of diverse eukaryotic genomes (human, calf, mouse, chicken, Xenopus, salmon, Drosophila, and yeast), and, although not proven, postulated that the repeats ought to have important biological significance (e.g. gene regulation, recombination hot spots, sites of mutagenesis) due to abundant occurrence and evolutionary conservation. Soon after, Litt and Luty (1989) estimated that the human genome contained approximately 50,000 repeat elements with the sequence (dT-dG), and showed that, like minisatellites, the new repeat elements were highly polymorphic and inherited in a codominant fashion. Litt and Luty proposed that the new repeat elements should be called ‘microsatellites’ (smaller repeat motif than minisatellites), and noted that since the elements were abundant and interspersed throughout the genome they could serve as useful genetic markers in regions where minisatellites were not present.

Since this early work numerous researchers have noted the ubiquitous nature of microsatellites across highly divergent taxa, and have more recently indicated that microsatellites are present in every organism studied to date. (Tautz and Renz, 1984; Epplen et al., 1993; Morgante and Olivieri, 1993; Gupta et al., 1996; Powell et al., 1996; Goldstein and Pollock, 1997; Gupta et al., 1999; Toth et al., 2000; Li et al., 2002; Morgante et al., 2002; Varshney et al., 2002; Ellegren, 2004; Li et al., 2004; Liu and
Since the original descriptions of microsatellites, these genetic elements have attracted great attention for a wide range of fundamental and applied investigations due to high polymorphism rate, ease of visualization, and stable codominant inheritance. This has led numerous researchers to conclude that microsatellites are the most popular or have the greatest utility of any marker system.

The terminology used to describe microsatellites has been somewhat confusing. The term “variable numbers of tandem repeats (VNTR)” has been used to describe both minisatellites and microsatellites, referring to the repetitive nature of a core sequence motif (e.g. Litt and Luty, 1989). Microsatellites have also been referred to as simple sequence repeats (SSRs) by Jacob et al. (1991) or short tandem repeats (STRs) by Edwards et al. (1991). Additional terms that have been used, albeit infrequently, include sequence tagged microsatellite site (STMS) (Beckman and Soller, 1990) and simple sequence length polymorphism (SSLP) (Tautz, 1989). Currently, the terms simple sequence repeats (SSRs) and microsatellites are preferred and used interchangeably.

The currently accepted definition of a microsatellite or SSR is a tandemly repeated DNA sequence with a core repeat motif length of 1 to 6 bp (Litt and Luty, 1989; Tautz, 1989; Liu et al., 1995; Gupta et al., 1996; Goldstein and Pollock, 1997; Toth et al.,
2000; Li et al., 2002; Liu and Cordes, 2004; Chistiakov et al., 2006; Selkoe and Toonen, 2006). They are usually referred to in the notation (GA)_n, where the nucleotides in parenthesis represent the core motif or unit that is repeating and n is the number of times the motif is repeated in a row. Microsatellites with one, two, three or four nucleotides making up the core repeat motif are referred to as mononucleotide, dinucleotide, trinucleotide, and tetranucleotide repeats, respectively. Microsatellites with five or six nucleotides making up the core repeat motif are rare and will not be discussed here.

Microsatellites are usually broken into three categories: perfect (or pure), imperfect (or interrupted), and compound (Weber, 1990a; Jarne and Lagoda, 1996). Perfect microsatellites have no interruptions within the run of nucleotide repeats; imperfect microsatellites have one or more interruptions within the repeat region; and compound microsatellites are a run of perfect or imperfect repeats adjacent to another simple sequence repeat.

The utility of microsatellites as molecular markers is a function of the mutational mechanism(s) that creates variation (polymorphism) at the SSR loci. In general, the mutation rates of microsatellites range from \(10^{-6} - 10^{-2}\) events per locus per generation (Levinson and Gutman, 1987; Dallas 1992; Dietrich et al., 1992; Weber and Wong 1993; Schug et al., 1997; Wierdl et al., 1997; Schlotterer et al., 1998; Schug et al. 1998; Kovalchuk et al. 2000; Vazquez et al., 2000; Udupa and Baum 2001; Li et al., 2002; Thuillet et al., 2002; Vigouroux et al., 2002), which is considerably higher than the rate of point mutations at other loci. This high mutation rate leads to ‘hypervariability’ in the number of alleles present at microsatellite loci (Ellegren, 2004). The majority of polymorphisms at these alleles are due to changes in the repeat number of the
microsatellite core motif, i.e. the changes involve the addition or subtraction of a certain number of repeat units [see below for exceptions] (Valdes et al., 1993; Di Rienzo et al., 1994; Kimmel and Chakraborty, 1996; Nauta and Weissing, 1996; Feldman et al., 1997; Eisen, 1999; Li et al., 2002; Ellegren, 2004). Early research on the mutation process for microsatellites was based on the stepwise mutation model (SMM) (Ohta and Kimura, 1973) whereby the length of the repeat is altered by the addition or removal of one repeat unit at a fixed rate (Shriver et al., 1993; Valdes et al., 1993; Kimmel and Chakraborty, 1996; Kimmel et al., 1996). Subsequent research indicated that a simple SMM did not fit observational data. Deviations from a simple SMM include multistep changes (Weber and Wong, 1993; Amos et al., 1996; Primmer et al., 1996; Fitzsimmons, 1998; Primmer et al., 1998; Brinkmann et al., 1998; Jones et al., 1999; Shimoda et al., 1999; Ellegren, 2000a; Gardner et al., 2000; Kayser et al., 2000; Xu et al., 2000; Holtkemper et al., 2001; Harr et al., 2002; Huang et al., 2002; Beck et al., 2003; Myhre Dupuy et al., 2004), an upper limit ceiling on repeat expansions (Garza et al., 1995; Nauta and Weissing, 1996; Feldman et al., 1997; Dermitzakis et al., 1998; Samadi et al. 1998; Pollock et al., 1998; Zhivotovsky, 1999; Ellegren, 2000a; Harr and Schlotterer, 2000; Stefanini and Feldman, 2000; Calabrese and Durret, 2003), bias such that mutation rate is dependent upon SSR length where long loci mutate more often than short loci (Weber, 1990a; Zhang et al., 1994; Primmer et al., 1996; Brinkman et al., 1998; Primmer et al., 1998; Schlotterer et al., 1998; Crozier et al., 1999; Ellegren, 2000a; Harr and Schlotterer, 2000; Kayser et al., 2000; Xu et al., 2000; Bacon et al., 2001; Brohede et al., 2002; Huang et al., 2002; Webster et al., 2002; Beck et al., 2003; Lai and Sun, 2003; Leopoldino and Pena, 2003; Sibly et al., 2003) and expansions that are stopped by interrupted point mutations,
breaking up one long perfect repeat array into two shorter repeats, which has the affect of reducing mutation rate (Bell and Jurka, 1998; Kruglyak et al., 1998; Palsboll et al., 1999; Taylor et al., 1999; Kruglyak et al., 2000). The interrupted point mutations could represent the initial ‘death’ of a microsatellite (Ellegren, 2000b), however, Harr et al. (2000) have shown that interruptions in a microsatellite array can be removed and, therefore, may only represent a transition state during the course of microsatellite evolution. These deviations from the standard SMM have resulted in new models of microsatellite mutation that have real implications for population genetics and phylogenetic studies.

There have been two main mechanistic models for microsatellite mutation: replication slippage and recombination-associated mutation. Recombination associated mutation, where SSR length is altered by ‘unequal crossing over’ or gene conversion events (Harding et al., 1992; Jakupciak and Wells, 2000; Richard and Paques, 2000), has lost favor as a general microsatellite mutation mechanism relative to the replication slippage mechanism (Ellegren, 2004). It could also be that microsatellite recombination associated mutation is a more appropriate model to explain uncontrolled triplet repeat expansions associated with human disease (Jakupciak and Wells, 2000; Richard and Paques, 2000). The microsatellite replication slippage mechanistic model (originally described for microsatellites by Levinson et al., 1985; Levinson and Gutman, 1987; first in vitro evidence for microsatellites by Schlotterer and Tautz, 1992) involves transient dissociation of the replicating DNA strands containing the SSR array followed by misaligned reassociation. During DNA replication, DNA polymerase pauses and then dissociates from the DNA allowing the terminal portion of the newly synthesized strand
to separate from the template and anneal to a different repeat unit (Hile and Eckert, 2004). In vitro experiments have shown that DNA polymerase is the only enzyme needed for replication slippage to occur (Schlotterer and Tautz, 1992). When the DNA strands realign out of register, continued replication results in the insertion or deletion of repeat units relative to the original template strand. There are two realignment possibilities: 1) misalignment resulting in a loop in the nascent strand which produces a longer SSR repeat; and 2) misalignment resulting in a loop in the template strand which produces a shorter SSR repeat (Ellegren, 2004). Loops created during misalignment can form ‘hairpin-like’ secondary structures that have been shown to escape DNA repair (Moore et al., 1999; Panigrahi et al., 2005; Pearson et al., 2005). It should be noted that most of the misalignment mutations are corrected by proofreading or the cellular mismatch repair (MMR) system (Strand et al., 1993), while only a few of the changes become permanent mutations producing the variability (polymorphism) of alleles present at SSR loci.

As noted above, most microsatellite mutations/polymorphisms are assumed to be changes in the length of the repeat array itself. From a mechanistic point of view (strand slippage) this is likely how many of the mutations occur, however, there are known exceptions. Tautz et al. (1986) were one of the earliest to show that flanking regions of microsatellites were composed of less than perfect, but still repetitive DNA that had higher levels of genetic variation than single-copy DNA. Brohede and Ellegren (1999) found higher rates of nucleotide substitution in the ends and 5-10 bp of flanking sequences of dinucleotide repeats, while Fernandez et al. (2008) found a SNP or insertion/deletion polymorphism (INDEL) in all of the 91 flanking regions of
microsatellites of grapes (*Vitis* spp.) they studied. With these results in mind, it may not be safe to assume that all length polymorphisms in microsatellites are the result of changes in the length of the repeat array itself, however, sequencing of all microsatellite loci to confirm the nature of the polymorphism would substantially increase time and costs in any study that uses SSRs as molecular markers.

The first report of microsatellites in plants was provided by Condit and Hubbell (1991) in a study of six different tropical tree species. Immediately following this initial report, a number of researchers indicated that microsatellites were also abundant in other plant species (Akkaya et al., 1992; Langercrantz et al., 1993; Morganti and Oliveri, 1993; Wu and Tanksley, 1993; Wang et al., 1994). Initial reports and categorization of microsatellites estimated that microsatellites were less frequent in plants than in mammals (Morganti and Oliveri, 1993), while early database searches indicated that microsatellites were five times less abundant in plant genomes than in animal genomes (Langercrantz et al., 1993). Early studies also indicated that microsatellite abundance differed between monocots and dicots (Condit and Hubbell, 1991; Wang et al., 1994). Wang et al. (1994) reported an average of one microsatellite every 64.6 kb in monocots, compared with one every 21.1 kb in dicots, suggesting that microsatellites occur three times more frequently in dicots than in monocots. The location and distribution of microsatellites in plants was also studied in these early papers. DNA sequence database searches, of nuclear and organellar DNA, in EMBL and Genbank showed microsatellites were present in 58 and 24 plant species, respectively (Wang et al., 1994). This work indicated that microsatellites were rarely found in organelles but were more abundant in nuclear DNA. Mono-, di-, and tetra- nucleotide repeats were predominantly found in
non-coding regions of the DNA, while 57% of the tri-nucleotide repeats containing GC were located in coding regions. Trinucleotide repeats, in all organisms, are still found to be significantly more abundant than other repeat types in coding sequences. The preponderance of trinucleotide repeats in coding sequences is likely due to negative selection against frameshift mutations in coding regions (Metzgar et al., 2000).

The early work characterizing microsatellites in plants, mentioned above, has led to many lingering false perceptions about the distribution, abundance, location and function of microsatellites. For instance, recent reports of microsatellite frequency in plant genomes (Cardle et al., 2000; Morgante et al., 2002; Varshney et al., 2002) has been much higher (average of 1 SSR every 6-7 kb) than originally reported by earlier studies (Langercrantz et al., 1993; Morganti and Oliveri, 1993; Wang et al., 1994). This is likely due to an increase in genomic and EST sequence data now available for plant genomes, and makes the frequency of microsatellites in plant genomes comparable to the frequencies described in humans and other mammal systems (Schulman et al., 2004).

Another lingering misperception has been that microsatellites are preferentially associated with the repetitive DNA (non-transcribed; heterochromatic) portion of both animal and plant genomes. Early research suggested that microsatellites originated in these non-transcribed regions, and could be preferentially associated with retrotransposable elements (Arcot et al., 1995; Duffy et al., 1996; Nadir et al., 1996; Motohashi et al., 1997; Ramsay et al., 1999; Elsik and Williams, 2001; Wilder and Hollocher, 2001). Again, with the abundance of new sequence data from a variety of organisms it is now possible to conduct more thorough studies of microsatellite locations within genomes. Morgante et al. (2002) conducted a comprehensive analysis of
abundance and distribution of microsatellites among transcribed and non-transcribed regions of Arabidopsis, rice, soybean, maize, and wheat. Contrary to previous assumptions, the frequency of microsatellites was significantly higher in gene-rich, non-repetitive, transcribed regions than in repetitive gene-poor genomic DNA across all species. In Arabidopsis, 3’ UTRs had a higher microsatellite frequency than non-transcribed regions, with significant increases in tri- and tetra-nucleotide repeats. Additionally, 5’ UTRs exhibited an almost threefold higher frequency of microsatellites than any other genomic region, with the increase in frequency due mostly to di- and tri-nucleotide repeats. Similar findings related to increased SSR density in the non-repetitive, transcribed regions of genomes have been reported for rice, where genomic SSRs, that are not even part of EST-derived SSRs, were preferentially associated with genic regions near promoters or inside introns and away from highly repetitive gene-poor DNA in heterochromatin (La Rota et al., 2005). Additionally, an interesting observation by Fujimori et al. (2003) noted a gradient of microsatellite density along the direction of transcription in rice and Arabidopsis, where the SSR density began gradually increasing upstream of the transcriptional start sites (TSS) of genes, rose dramatically immediately before the TSS, and then decreased gradually downstream of the TSS.

The exact function of microsatellites is not completely understood. Initial studies often ascribed potential function to mini- and microsatellites such as the alternating purine/pyrimidine repeats (such as (GT)n and (GC)n) playing a role in gene regulation, genetic recombination or chromosome packing/condensing by forming a Z-DNA structure (Hamada et al., 1982; Korol et al., 1994; Karlin et al., 1998; Biet et al., 1999). Although a few researchers recognized the utility of repetitive mini- and microsatellite
DNA for generating molecular markers, the study of function associated with ‘non-coding’ DNA has not been a ‘fashionable’ subject relative to in-depth analysis of coding sequences (Shapiro, 2002; Makalowski, 2003). Orgel and Crick (1980) coined the term “junk DNA” to describe the excess, non-coding DNA that did not play an obvious role in protein coding. Unfortunately, the term “junk DNA” has stuck with repetitive DNA such as microsatellites, and likely has impeded progress on ascribing universally recognized functions. After all, who wants to waste their time studying someone else’s ‘junk’? This perception has endured despite the fact that existing literature does describe possible function associated with repetitive DNA sequences (Shapiro, 2002).

In general, it has been known for some time that SSRs commonly occur within transcribed genes (in both exons and intervening introns) and the non-transcribed regulatory regions of genes (King et al., 1997). As an example, 12% of identified SSRs in Japanese pufferfish (Edwards et al., 1998), 10% in primates (Jurka and Pethiyagoda, 1995), 15% in rabbit (van Lith and van Zutphen, 1996), 9.1% in pig (Moran 1993), and 10.6% in chicken (Moran 1993) occur in protein coding regions of genes. Additionally, microsatellites that are found associated with 5’ untranslated regions, promoters, upstream regulatory elements, transcription factors, and transcription start sites or in introns have been found or hypothesized to be involved in the regulation of gene expression and/or transcriptional activity (Hamada et al., 1984a; Hamada et al., 1984b; Solomon et al., 1986; Lue et al., 1989; Winter and Varshavsky, 1989; Hoffman et al., 1990; Punt et al., 1990; Lafyatis et al., 1991; Stallings et al., 1991; Aharoni et al., 1993; Epplen et al., 1993; Chamberlain et al., 1994; Gerber et al., 1994; King, 1994; Kim and Mullet, 1995; Lanz et al., 1995; Sandaltzopoulos et al., 1995; Epplen et al., 1996; Chen
and Roxby, 1997; Kashi et al., 1997; Csink and Henikoff, 1998; Meloni et al., 1998; Miret et al., 1998; Okladnova et al., 1998; Xu and Goodridge, 1998; Gebhardt et al., 1999; Gebhardt et al., 2000; Liu et al., 2000; Martienssen and Colot, 2001; Fujimori et al., 2003; La Rota et al., 2005). Additional specific examples of microsatellite function, from a variety of organisms, include the following:

*Temperature compensation of circadian rhythm in Drosophila: Sawyer et al. (1997) reported that the clock gene *period (per)* in *Drosophila melanogaster* contains a hexanucleotide repeat encoding threonine and glycine (Thr-Gly) where the shorter allele (Thr-Gly)$_{17}$ results in a clock period closer to 24 hours, whereas the longer allele (Thr-Gly)$_{20}$ results in better temperature compensation so that temperature fluctuations have a lesser impact on circadian cycle. Population sampling across Europe and North Africa indicated that the two alleles had a significant latitudinal cline, with the longer allele predominant in colder regions.*

*Social behavior in voles: Hammock and Young (2004; 2005) showed that alternate alleles of a compound SSR in the 5’ regulatory region of the vasopressin receptor gene (*avpr1a*) could dictate partner preferences in monogamous (social) and non-monogamous (non-social) voles. This relationship was confirmed in a transgenic mouse containing the social species’ version of the SSR, where gene expression patterns and behaviors in response to experimental vasopressin injection were more like those of the social species than those of the wild-type mouse (Young et al., 1999).*
Morphological differences in domestic dog: While studying 92 breeds of domestic dog, Fondon and Garner (2004) found multiple lines of evidence that strongly suggests that SSRs are at least partially responsible for phenotypic differences among dogs within the same breed and for morphological differences between dog breeds. First, these investigators noted that the length ratio of two adjacent SSRs in the runt-related transcription factor Runx-2, that encodes 18-20 glutamines followed by 12-17 alanines, was correlated with facial shape characteristics across breeds. Secondly, the condition of polydactyly (extra toes) in Great Pyrenees was associated with a 51 bp contraction of a hexanucleotide repeat in Alx-4. Finally, and perhaps most interestingly, the authors noted that 17 other genes known to influence morphological traits had only a few silent SNPs (single nucleotide polymorphisms), yet had “extraordinary levels of tandem repeat variation” with a polymorphism in almost every gene. This led the authors to state that “frequent length mutations in gene-associated tandem repeats generate copious robust morphological variation… The extraordinarily high length mutation rates of tandemly repeated sequences can thus result in abundant variation upon which selection may act”. So in this case, the strand-slippage mutational mechanism of SSRs may outpace the single-nucleotide mutation rates allowing organisms to use microsatellite mutation as a “mechanism to achieve evolutionary agility”.
Unstable repeat-related diseases in humans: Allelic differences in the repeat number of SSRs in humans are known to cause a wide range of hereditary neurological disorders and disease susceptibilities. The most common repeat number neurological disorders/diseases are known as the ‘triplet repeat diseases’, however the list of these diseases/disorders has expanded to include tetra- and pentanucleotide repeats leading to a new designation as ‘diseases/disorders of unstable repeat expansion’ (Gatchel and Zoghbi, 2005). Primary neurological diseases caused by unstable repeat expansion include Fragile X syndrome, myotonic dystrophy, Huntington’s disease, various spinocerebellar ataxias, and Friedreich’s ataxia, while additional less common neurological diseases/disorders also have root causes related to unstable repeat expansion (Mirkin, 2006). In most cases the mechanisms for disease causation involve expansion of various SSRs beyond critical repeat lengths, resulting in loss of protein function, altered protein function, or altered RNA function as molecular features of the disease/disorder (Gatchel and Zoghbi, 2005). The exact molecular mechanism of a few neurological diseases is still unknown, but in all cases individuals exhibiting symptoms carry allelic variants with excessively long repeat lengths (Gatchel and Zoghbi, 2005). Additional human diseases linked to microsatellite repeat instability include higher incidence of hereditary nonpolyposis colorectal cancers and occasional sporadic tumors of the colon (Redston, 2001), as well as some association with certain lung cancers (Zienoddiny et al., 1999), endometrial cancers (Vassileva et al., 2002), gastric cancers (Yamada et al., 2002), and prostate cancers (Suuriniemi et al., 2007).
**Starch properties in rice:** Amylose content, an important determinant of rice starch quality, has been found to be primarily controlled by the *waxy* gene which encodes granule-bound starch synthase. Ayers et al. (1997) found seven (CT)$_n$ microsatellite alleles in the 5’ untranslated region of the *waxy* gene, which together explained more than 82% of the variation in the apparent amylose content of important long-, medium- and short-grain ‘non-waxy’ US rice cultivars. Later research conducted by Bao et al. (2002) identified four (CT)$_n$ alleles in the *waxy* gene, three (CT)$_n$ alleles in the starch branching enzyme (*SBE*) locus, and multiple di- and trinucleotide SSR alleles in the soluble starch synthase (*SSS*) locus of ‘waxy’ rice accessions, all of which appeared to be correlated to amylose or amylopectin properties.

**Control of somatic and zygotic cell cycles in maize:** Dresselhaus et al. (1999) identified ribosomal proteins involved in ribosome biosynthesis and translation that were differentially expressed in maternal somatic cells and embryonic cells shortly after fertilization in maize. A key difference between the two cell types was the presence of a GCC triplet repeat that was present in the 5’ UTR of ribosomal proteins stored in egg cells, indicating that the SSR motif may play a role in the transition from maternal to embryonic control of development.

**Population divergence in barley and wheat based on microclimate:** Nevo et al. (2005) studied the distribution of 19 nuclear and four chloroplast microsatellite
loci in wild barley (*Hordeum spontaneum*) populations of the Evolution Canyon in Israel. This site has significant microclimate contrast on south-facing versus north-facing slopes with biotic/ecological zones separated by only 100 to 400 meters at various locations within the canyon. The authors identified seven subpopulations of wild barley distributed across the north and south-facing slopes with significant interslope variation of SSR allele frequencies. Similarly, Fahima et al. (2002) found ecogeographic correlation with SSR allele frequencies in wild emmer wheat (*Triticum dicoccoides*) populations across divergent soil and climatic conditions. Both research groups indicated that the statistical analysis showed evidence of adaptive significance of the SSR alleles themselves, but because the exact function of the SSRs remains unknown, conclusive evidence that SSRs are being selected, rather than segregating with adaptive alleles, requires further study.

**Expressed sequence tagged simple sequence repeats (EST-SSRs):**

Expressed sequence tags (ESTs) are DNA sequences generated by sequencing of complimentary DNA (cDNA) clones that correspond to a mRNA or a part thereof (Gupta et al., 2002). ESTs are created by a multistep process: 1) selective isolation of mature mRNA (introns spliced and ends capped) from a target organism; 2) production of a cDNA strand from a mRNA template using reverse transcriptase; 3) creation of double stranded cDNA using DNA polymerase; 4) insertion of cDNAs into a vector library; 5) and end sequencing of cDNA inserts using universal primers. A cDNA library is a ‘snapshot’ representation of the messenger RNA population within an organism,
genotype, tissue, cell type, developmental stage, following specific biotic or abiotic challenges, or under a certain set of conditions at a particular point in time (Rudd, 2003; Dong et al., 2005). The impetus for developing ESTs stemmed from the early belief that initial sequencing efforts should target the discovery of genes in the human genome, essentially focusing on transcriptome sampling (Brenner, 1990; Adams et al., 1991; Schuler, 1997). Nobel laureate Sydney Brenner (1990) stated, “If something like 98% of the [human] genome is junk, then the best strategy would be to find the important 2%, and sequence it first”. Early opponents of this approach pointed out that valuable information from intronic and intergenic regions, such as regulatory sequences and chromosome organizational control, would be missed (Committee on Mapping and Sequencing the Human Genome, National Research Council, 1988). This debate, as well as debate over patent issues and public versus private accumulation of sequencing data continued during the early stages of the Human Genome Project (Boguski, 1995). Interestingly, as some scientists argued over procedure and policy, work on the generation of ESTs from numerous organisms continued, and it gradually “became apparent to most biologist, even former skeptics, that this type of incomplete and inaccurate, but rapidly and cheaply obtainable data was indeed quite useful: for ‘phylum hopping’ among organisms; for discovering new members of gene families involved in human disease; for the identification of exons in vast expanses of genomic DNA; and as a plentiful source of gene-based mapping reagents with which to populate physical maps” (Boguski, 1995). To a certain degree these debates have become moot, as major funding initiatives and the development of improved sequencing and bioinformatics technologies eventually allowed for whole genome sequencing and annotation for the human genome
project and a few other model organisms. Additionally, EST projects are now considered a complementary resource for genome-scale experiments within as yet un-sequenced genomes (Mayer and Mewes, 2001; Rudd, 2003). This resource is quite extensive, as currently there are just over 54.5 million EST sequences, from hundreds of organisms, in public databases (Boguski et al., 1993 - dbEST 18 July 2008).

ESTs by themselves are not molecular markers, as they simply represent sequence information generated from cDNA. After development, EST collections can be searched and used to develop both hybridization based (RFLP) or PCR based markers (Gupta and Rustgi, 2004). The PCR based markers could be developed using standard methods such as the design of locus specific primers that flank potential polymorphisms (sequence tagged site – STS), or developed in silico. The former could then be used for traditional analysis of AFLPs, SSRs, INDELS, SNPs etc., while the latter requires sequence information for all of the material to be compared for polymorphisms.

Searching for SSRs in EST collections has a number of advantages over traditional genomic SSR development protocols. First, EST-SSRs can be quickly obtained by electronic sorting, preventing the need for costly and time-consuming benchwork associated with library construction, enrichment and screening (Bouck and Vision, 2007). In fact, EST-SSRs are a free byproduct of the currently expanding publicly available EST databases (Scott, 2001; Gupta et al., 2002; Varshney et al., 2005a). Numerous examples of electronic mining of SSR markers out of EST data sets from a variety of organisms has been published in recent years (Sasaki et al., 1994; Miyao et al., 1996; Delseny et al., 1997; Yamamoto and Sasaki, 1997; Wang and Bowen, 1998; Cho et al., 2000; Eujayl et al., 2000; Henry et al., 2000; Scott et al., 2000; Scotti et
al., 2001; Cordeiro et al., 2001; Eujayl et al., 2002; Holton et al., 2002; Kantety et al.,
2002; Rohrer et al., 2002; Gupta et al., 2003; Jany et al., 2003; Bhat et al., 2005;
Varshney et al., 2002; Decroocq et al., 2003; Thiel et al., 2003; Gao et al., 2004;
Khlestkina et al., 2004; Nicot et al., 2004; Saha et al., 2004; Yu et al., 2004; Feingold et
al., 2005; Varshney et al., 2005b; Wang et al., 2005; Zhang et al., 2005; Parida et al.,
2006; Tang et al., 2006; Aggarwal et al., 2007; Becher, 2007; Shanker et al., 2007; Zhang
et al., 2007). Second, EST-SSRs have been shown to be more widely transferable among
closely, or sometimes distantly, related species and genera than SSRs generated from
genomic libraries (Cordeiro et al., 2001; Arnold et al., 2002; Holton et al., 2002; Rossetto
et al., 2002; Besnard et al., 2003; Decroocq et al., 2003; Gupta et al., 2003; Saha et al.,
2003; Thiel et al., 2003; Woodhead et al., 2003; Bandopadhyay et al., 2004; Chagne et
al., 2004; Eujayl et al., 2004; Liewlaksaneeyanawin et al., 2004; Rungis et al., 2004;
Boches et al., 2005; Coulibaly et al., 2005; Fraser et al., 2005; Gutierrez et al., 2005; Ng
et al., 2005; Varshney et al., 2005b; Zhang et al., 2005; Hayden et al., 2006; Pashley et
al., 2006; Poncet et al., 2006; Barbara et al., 2007; Ellis and Burke, 2007; Portis et al.,
2007; Slate et al., 2007; Yu and Li, 2008). This is likely due to the fact that EST-SSRs
are derived from the conserved coding regions of the genome of the original organism in
which they were found (Gupta et al., 2002; Zhang et al., 2005; Pashley et al., 2006;
Bouck and Vision, 2007; Slate et al., 2007). High transferability of EST-SSRs offers two
key features: 1) low cost and ease of marker development in taxa with limited molecular
遗传资源；和 2) 标记集用于比较物种之间的图谱或作为
anchor markers for comparison of model genomes to orphan or other crops for the
purposes of studying synteny and collinearity among related genomes (Gupta and Rustgi,
A third advantage of EST-SSRs over genomic SSRs stems from the fact that
EST-SSRs are derived from transcripts. As such, EST-SSRs can be used to assess
functional identity of a marker through homology searches in sequence and protein
databases (Bouck and Vision, 2007), could be used as a short-cut to identify candidate
genes associated with particular phenotypes (Vasemagi et al., 2005), or could more
broadly be used to assess biological/functional diversity in transcribed and known-
function genes (Ayers et al., 1997; Eujayl et al., 2002; Hackauf and Whehling, 2002;
Thiel et al., 2003; Varshney et al., 2005b). This latter use has led some authors to
distinguish EST based markers versus non-EST based markers as functional versus
random or anonymous markers, respectively (Hackauf and Whehling, 2002; van
Tienderen et al., 2002; Thiel et al., 2003; Gupta and Rustgi, 2004). Andersen and
Lubberstedt (2003) further defined functional markers as being derived from
polymorphic sites within genes causally involved in phenotypic trait variation. This latter
definition clearly encompasses the ultimate goal of some of the uses of molecular
markers, such as association mapping and QTL analysis, and is more likely to be
achievable when using EST derived markers than random or anonymous markers.

**Single Nucleotide Polymorphisms (SNPs):**

Single nucleotide polymorphisms (SNPs – read as “SNiPs”) are broadly defined
as a single base change in a DNA sequence (Vignal et al., 2002) or a DNA position at
which two alternative bases occur at appreciable frequency within a population (Wang et
al., 1998). A more formal definition was provided by Brookes (1999): “SNPs are single
base pair positions in genomic DNA at which different sequence alternatives (alleles)
exist in normal individuals in some population(s), wherein the least frequent allele has an abundance of 1% or greater”. In principle, this would exclude single base insertion/deletions (INDELS), but because single (or even multiple) base INDELS are often discovered during SNP searches many researchers loosely categorize some of these polymorphisms as SNPs (Brookes, 1999).

Theoretically, any one of the four nucleotides of DNA can represent a SNP at a particular locus, however, in practice SNPs are usually considered to be bi-allelic where only two variants predominantly occur. The principle reason behind this is the inherent bias in transition versus transversion mutations. Transition mutations involve purine to purine (A ↔ G) or pyrimidine to pyrimidine (C ↔ T) changes, while transversions involve purine to pyrimidine or pyrimidine to purine (A ↔ C; A ↔ T; G ↔ C; G ↔ T) changes. Based on chance alone, there should be twice as many transversions as transitions (8 possibilities versus 4 possibilities, respectively) (Brookes, 1999), however experimental data from several taxa has shown that transitions occur 1.4 to 2.8 times as often as transversions (Vignal et al., 2002; Khlestkina and Salina, 2006). The principle underlying reason for this experimentally observed bias is thought to be related to 5-methylcytosine spontaneous deamination reactions to thimidine at CpG dinucleotides resulting in high levels of C ↔ T transitions (seen as G ↔ A transitions on the reverse strand) (Cooper and Krawczak, 1989; Holliday and Grigg, 1993; Wang et al., 1998; Vignal et al., 2002). This C ↔ T transition is estimated to represent about 2/3 of the SNPs in humans (Brookes, 1999).

The transition mutation bias in SNPs represents an intriguing dilemma in relation to the utility/genetic information of SNPs relative to other markers. For instance, does
the bi-allelic nature of SNPs make them less informative as a molecular marker than multi-allelic SSRs (Vignal et al., 2002)? At a single locus the answer is probably yes, however, there is an extraordinary abundance of SNPs across genomes that offsets the bi-allelic nature of an individual SNP marker (Kruglyak, 1997; Gupta et al., 2001; Brumfield et al., 2003; Evans and Cardon, 2004). SNPs are by far the most abundant molecular marker known so far (Gupta et al., 2001), and may represent up to 90% of the variation in the human genome (Brumfield et al., 2003). The sheer number of SNPs allows for the creation of dense coverage of genetic maps that is not possible with other markers. As an example, early estimates of SNP frequency in the human genome suggested 1 SNP / 1 kb (Li and Sadler, 1991; Harding et al., 1997; Kruglyak, 1997; Lai et al., 1998; Nickerson et al., 1998; Taillon-Miller et al., 1998). The first release of the draft sequence of the human genome estimated that a SNP occurred on average once every 1.25 kb (Venter et al., 2001). In a review of SNP frequency reported in a number of different taxa (including dolphins, human, flycatchers, murrelets, parrots, chicken, fruit fly, and sugar beat), Brumfield et al. (2003) estimated non-coding SNP frequency at 1 / 200-500 bp and coding SNP frequency at 1 / 500-1000 bp. Even higher average frequencies have been reported in wheat, where the SNP density was estimated at 1 / 370 bp (Procnnier et al., 2003), and rice, where the SNP density was estimated at 1 / 232 bp (Nasu et al., 2002). Additionally, it should be noted that certain regions and/or loci in genomes of various taxa have variable SNP frequencies. For example, SNP frequency averages 1 / 1000 bp in the human genome as a whole, but Horton et al. (1998) found SNPs as frequently as 1 every 10 bp in the HLA-DQB1 locus. Similar observations of hypervariable regions containing SNPs have been made in the domestic cattle amyloid
precursor protein (one SNP every 104 bp – Vignal et al., 2002), four loci in the Mediterranean fruit fly (one SNP every 52 bp – Villablanca et al., 1998), 37 loci in sugar beet (one SNP every 130 bp – Schneider et al., 2001), three starch related genes in wheat (one SNP every 20 bp – Wolters et al., 2000), and 20 loci in maize (one SNP every 70 bp – Bhattaramikki et al., 2000).

As well as being the most predominate type of DNA marker identified to date, SNPs have more methods of discovery and/or analysis than any previously developed DNA marker (Khlestkina and Salina, 2006). Therefore, SNPs are not only the most common type of polymorphism in most taxa, but are also useful over a broad range of research projects where technical and financial levels are variable. The following section provides brief summaries of the various methods of SNP detection:

**Sanger dideoxy-sequencing:** The most straightforward SNP discovery method is by Sanger dideoxy-sequencing. In this method PCR amplified products are sequenced and sequence differences are used for discovery of new SNPs. SNP discovery through de novo sequencing of genetic libraries is time-consuming, labor intensive, and expensive. Locus specific sequencing of a priori sequences cuts out initial library scanning, but still requires relatively high outlays in labor, time, and reagents. Additionally, Sanger sequencing generates significantly more information than is necessary for single SNP discovery. As a result, other methods for initial SNP detection have been developed that seek to offset some of the negative aspects of blind sequencing by employing more economical initial scanning techniques or improving scale-up and automation.
Alignment of overlapping sequences: This method relies on the availability of databases of genomic and EST sequences for SNP discovery by comparison of sequences from a heterozygote or from more than one individual of the same species. Alignment of database sequences and SNP detection is automated through computer software (Taillon-Miller et al., 1998; Garg et al., 1999; Marth et al., 1999; Picoult-Newberg et al., 1999; Irizarry et al., 2000; Dawson et al., 2001; Kota et al., 2001; Lehnert et al., 2001; Usecche et al., 2001; Suliman-Pollatschek et al., 2002; Barker et al., 2003; Kota et al., 2003; Batley et al., 2003; Schmid et al., 2003; Cheng et al., 2004; Feltus et al., 2004; Shen et al., 2004; Savage et al., 2005; Huntley et al., 2006; Lee et al., 2006; Mona et al., 2006; Panitz et al., 2007). SNP discovery can be made by comparison of overlapping genomic or EST sequences. This method is extremely cost-effective assuming that sequences are already available for analysis, however, costs would rise with any needed de novo sequencing.

PCR-RFLP or Cleaved Amplified Polymorphic Sequence (CAPS): In the PCR-RFLP method (also referred to as CAPS (Konieczny and Ausubel, 1993)) a target DNA segment is amplified by PCR and the product is incubated with appropriate restriction enzymes. The PCR product is “cut” if it contains the appropriate SNP variant that corresponds to the recognition site of the restriction enzyme. SNP polymorphisms can then be detected by fragment size analysis.
**Allele specific PCR** (method based on Ugozzoli and Wallace, 1991; Bottema and Sommer, 1993): In allele specific PCR a target DNA region is amplified with oligonucleotide primers designed for specific SNP sites. Different fluorophores or fluorophore/quenching systems can be attached to the 5’-end while polymorphic nucleotides are attached to the 3’-end. Allele specific primers perfectly match only the corresponding SNP alleles (and do not match the alternative allele) and differential fluorophores (or other appropriate labeling) can be used as the reporter system for allele discrimination.

**Single strand conformational polymorphism (SSCP)** (Orita et al., 1989): This method relies on differences in secondary structure of single-stranded DNA from PCR products that differ by one or more internal nucleotides. SNP detection is accomplished by denaturing the target PCR products and electrophoretic separation of the single-stranded DNA. Single-stranded DNA will take on unique secondary structure (conformation) according to nucleotide composition. The resulting changes in secondary structure cause changes in electrophoretic mobility. Differences in electrophoretic mobility of single-stranded amplicons between wild-type and mutant PCR products are therefore evidence of nucleotide changes (SNPs or INDELs). An advantage of this technique is that prior knowledge of a SNP or INDEL is not required.
**Heteroduplex analysis** (Gerrard and Dean, 1998): Similar to SSCP, heteroduplex analysis relies on conformational changes in DNA. As mentioned above, alterations in nucleotide sequence can cause changes in the structure of DNA which alters mobility of DNA in a sieving matrix. A nucleotide mismatch in duplex DNA can also cause changes in structural conformation, so that heteroduplex DNA mobility in a sieving matrix will differ from homoduplex DNA mobility. Heteroduplex analysis involves mixing PCR-amplified mutant and wild-type DNA followed by denaturation and reannealing. The resultant reannealed products will be a mixture of hetero- and homoduplex DNA. SNPs and INDELs can be inferred from differences in the electrophoretic mobility of the hetero- versus homoduplex PCR products.

**Cleavage Modifications of heteroduplex analysis – chemical or enzymatic cleavage:** As an alternative to migration of intact hetero/homoduplexes, methods have been developed that cleave the conformational structures (loops, bulges, kinks) of DNA nucleotide mismatches. Chemical cleavage of mismatches (CCM) (Smooker and Cotton, 1993) involves formation of heteroduplexes, followed by modification of mismatched thymine residues with osmium tetroxide and modification of mismatched cytosine residues with hydroxylamine. The modified bases are then cleaved by piperidine, and subsequent fragment size detection of cleavage products localizes the sites of the mismatch. Unprotected nucleotides in loops resulting from insertions and deletions are also detected.
Enzymatic cleavage is also possible. As an example, Hsu et al. (1994) demonstrated that PCR product duplex mismatches could be cleaved with E. coli mismatch repair enzymes. In another example, Till et al. (2004) showed that purified extracts of single-strand specific nucleases (SSS nucleases) extracted from a variety of plant material cleaved bulges caused by INDELS and single-base mismatches in heteroduplex DNA.

Denaturing high performance liquid chromatography (DHPLC) (Oefner and Underhill, 1998): DHPLC is yet another modification of heteroduplex analysis. This method uses ion-pair reverse phase chromatography to detect single nucleotide mismatches/indels of hetero-/homoduplex PCR products by injecting them into a preheated mobile phase that results in partial denaturation of the mixed DNA PCR products. The method exploits the differential retention of homoduplex and heteroduplex DNA species under conditions of partial thermal denaturation, where heteroduplex PCR products with a reduced melting temperature will elute earlier from the column than homoduplex PCR products.

Denaturing Gradient Gel Electrophoresis (DGGE) (Fisher and Lerman, 1983): In DGGE, PCR products are separated based on sequence differences that results in differential denaturing characteristics of the DNA. During DGGE analysis, PCR products encounter increasingly higher concentrations of chemical denaturant as they migrate through a sieving gel. Upon reaching a threshold denaturant
concentration, the weaker melting domains (based on sequence differences) of the
double-stranded PCR product will begin to denature at which time migration
slows dramatically. Sequences that differ by a single nucleotide or greater will
denature at different denaturant concentrations resulting in a pattern of bands
present in the sieving matrix.

Temperature Gradient Gel Electrophoresis (TGGE) (Riesner et al., 1989): The
principle of TGGE is similar to DGGE, the difference being that PCR products
encounter increasingly higher temperatures (as opposed to increasingly higher
concentrations of a chemical denaturant) as they migrate through a sieving gel.
At a certain temperature, the PCR products start to denature, resulting in a
conformational change in DNA structure. As temperature induced
conformational changes occur, electrophoretic mobility is slowed compared to a
double-stranded DNA fragment. Since the melting temperature strongly depends
on the nucleotide sequence, DNA fragments of the same size but different
sequence can be separated by fragment analysis.

Oligonucleotide ligation assay (OLA) (Landergren et al., 1988): In the OLA
method two independent oligonucleotide probes, one 5’ biotinylated and the other
3’ fluorescent-labeled (or other appropriate reporter), are used for hybridization to
denatured target PCR product. DNA ligase joins the two probes if they both
perfectly match the target PCR product. The ligation products can be captured on
a solid streptavidin-coated matrix due to the biotinylation at its 5’ end. Capture of the ligated probes with the 3’ reporter indicates presence of wild-type DNA, while absence of the 3’ reporter indicates no ligation due to a probe/DNA mismatch because of a SNP or INDEL.

Oligonucleotide hybridization/Sequencing by hybridization on DNA chips and microarrays (Lemieux et al., 1998): DNA chips and microarrays have been used for expression analysis, polymorphism detection, DNA resequencing, and genotyping on a genomic scale. The key features of these tools are the potential for massive data acquisition, at a fast rate through automation, followed by rapid analysis using computer technology. The basic principle behind this method for SNP discovery is hybridization of oligonucleotides with complimentary sample DNA on the chip/array surface. A specific example of SNP detection would be as follows (Gupta et al., 2001): Four oligonucleotide probes are arranged in a column of an array, differing by a single internal SNP site. Only fully homologous sample DNA will hybridize at the correct array site corresponding to the base composition of one of the four oligonucleotides. Detection can be achieved through autoradiography, enzymatically, or through fluorescence visualization.

Dynamic allele-specific hybridization (DASH) (Howell et al., 1999): DASH genotyping discriminates SNPs based on differences in the melting temperature of
perfectly matched DNA duplexes versus mismatched DNA duplexes. DASH genotyping can be divided into four main steps: In the first step, a genomic DNA segment is PCR amplified and prepared for attachment to a solid support by using a biotinylated primer. In the second step, the PCR product is denatured and washed leaving only the biotinylated single strand. In the third step, an oligonucleotide probe containing the SNP site is added in the presence of an intercalating dye that fluoresces when bound to double-stranded DNA. The fluorescence intensity is proportional to the amount of double-stranded DNA (matched versus mismatched probe-target hybridization). In the final step, temperature is increased steadily to denature (melt) the probe-target duplex. A sudden drop in fluorescence signal indicates the melting temperature (Tm) of the duplex. Single-base mismatches or INDELs can be detected when there is dramatic lowering of the homoduplex melting temperature.

Taqman (Livak et al., 1995) and molecular beacons (Tyagi and Kramer, 1996): The Taqman and molecular beacon methods are similar to OLAs in that they employ hybridization probes, and similar to each other in that both systems use allele specific fluorogenic probes with reporter dyes at the 5’ end and a quencher dye at the 3’ end. In the Taqman system, the 5’ reporter dye fluorescence is quenched by the close proximity of the 3’ dye quencher. Upon successful hybridization of the probe to target DNA, the probes are degraded at the 5’ end by the 5’ to 3’ exonuclease activity of Taq polymerase. The exonuclease activity
releases (“cuts”) the fluorescent tag from the probe which increase the fluorescent signal. In the presence of a SNP the Taqman probe will not hybridize with the target DNA, is not exposed to the 5’ to 3’ exonuclease activity of Taq polymerase, and there is no release of fluorescence signal. The same principle is used for assays employing molecular beacons, however, molecular beacon probes are quenched because the ends form a hairpin structure due to self-annealing of the two ends. If the molecular beacon probe successfully anneals to the target DNA (no mismatch) the probe becomes linearized and releases the fluorescent reporter from close proximity to the quencher dye, which in turn increases fluorescence signal generation. In the presence of a SNP, the molecular beacon probe will not hybridize with the target DNA, the hairpin structure containing the reporter and quencher dyes are not linearized, and there is no release of fluorescence signal.

Pyrosequencing (Ronaghi et al., 1996; Ronaghi et al., 1998; Ronaghi, 2001): Pyrosequencing is a DNA sequencing technique, typically of short length nucleotides (20 nucleotides to possibly up to 100 nucleotides), that detects and analyzes the release of pyrophosphate (PPI) during DNA synthesis. The method relies on a series of enzymatic reactions, releasing visible light that is proportional to the quantity and identity of incorporated nucleotides. The reaction starts with the incorporation of a nucleotide that is complimentary to a template strand, which in turn releases inorganic PPI. The released PPI is used to convert amino phosphosulfurate (APS) to ATP by the enzyme ATP sulfurylase. The release of ATP provides the energy to the enzyme luciferase to oxidize luciferin and
generate visible light. The chemiluminescence is proportional to the ATP produced, which is directly proportional to the specific dNTP added to the growing strand. PCR primers are designed close to previously identified SNP loci to give short sequences of target sites for comparison of SNP alleles.

*Minisequencing* (Sokolov, 1990; Syvanen et al., 1990; Kuppuswamy et al., 1991): The minisequencing procedure is based on Sanger dideoxynucleotide sequencing, where the sequencing primer is designed to anneal one or a few basepairs upstream from the target SNP. A mixture of ddNTPs and dNTPs allows discrimination of the SNP by comparison of sequencing extension products that do or do not terminate at the SNP site by inclusion of a chain terminating ddNTP or chain elongating dNTP, respectively. Detection methods include traditional fluorescently labeled ddNTPs, radioactivity, MALDI-TOF (Matrix Assisted Laser Desorption/Ionization Time Of Flight – mass spectrometry method that discriminates based on molecular weight differences of extension products), or chip based assays.

*Invasive cleavage assay* (Lyamichev et al., 1999): The invasive cleavage assay is based on hybridization of two sequence specific probes (“invader” oligonucleotide and “probe” oligonucleotide) that anneal to the target DNA SNP site with an overlap of one nucleotide. The probe oligonucleotide forms a duplex with the target DNA and is then subsequently ‘invaded’ by the invader probe at the overlap site. The duplex plus invasion probe forms a unique secondary
structure that is recognized and cleaved by flap endonucleases (FENs). Presence of cleavage products indicates that the allelic base is complimentary to the base in the probes, while absence of cleavage products indicates non-complimentarity between the allelic base and probe. Detection methods include FRET (Mein et al., 2000), fluorescence polarization (Hsu et al., 2001), or MALDI-TOF.

‘Anchored’ primers for comparative genetics/mapping – EPIC-PCR (Exon Primed Intron Crossing), ILP markers (Intron Length Polymorphisms), COS markers (Conserved Ortholog Set), CATS (Comparative Anchor Tagged Sequences), CISPs (Conserved Intron Spanning Primers):

Comparative genetics and mapping are an important foundation for studying structural conservation among genomes of diverse genera and for establishing and testing hypotheses about the relationship between gene structure and function in a wide range of organisms (Van Deynze et al., 1998). On a more practical level, comparative genetics/mapping can be used to understand the evolution of plant genomes for phylogenetic assessment (Bennetzen and Freeling, 1993; Gale and Devos 1998), help validate quantitative trait loci from one species to another (Paterson et al., 1995), or be a useful strategy for map-based cloning of orthologous genes (Foote et al., 1997; Kilian et al., 1997). This is possible because early comparison of molecular markers among different taxa revealed remarkable conservation of gene content and gene order over millions of years of evolution (Gale and Devos, 1998). These findings established that “…the evolution of the small but essential portion of the genome that actually encodes the organisms genes has proceeded relatively slowly; as a result, taxa that have been
reproductively isolated for millions of years have retained recognizable intragenic DNA sequences as well as similar arrangements of genes along the chromosomes.” (Paterson et al., 2000). The cross-genome retention of similar DNA sequences allowed for early use of hybridization based (RFLP) comparative mapping procedures, which, with variable stringency conditions could be cross-mapped to similarly matched DNA sequences in closely to even distantly related genomes (Ahn and Tanksley, 1993; Devos et al., 1993a; Melake-Berhan et al., 1993; Devos et al., 1994; Grivet et al., 1994; Kurata et al., 1994a; van Deynze et al., 1995a; van Deynze et al., 1995b; Devos et al., 1998). One of the most successful results of this approach was the development of a core set of ‘anchor probes’ for comparative mapping of grass genera using rice as a reference genome (van Deynze et al., 1998). These anchor probes, along with similarly designed RFLP probes in genera other than grasses, have been widely used on numerous organisms, but still suffer from the drawbacks discussed in previous sections concerning the use of RFLPs. As a result, numerous investigators have sought alternative ‘anchor’ markers that are based solely on the polymerase chain reaction (PCR).

An obstacle to designing PCR based ‘anchor’ markers involves finding genomic sequences that are conserved across many divergent taxa that can serve as PCR primer site locations, yet also have DNA level differences that can reflect diversity at the ecotype, species, genera, or family level for genetic investigations. Feltus et al. (2006) pointed out that the relatively high level of conservation of the location of exon-intron interfaces, but not the sequences of introns, provides a potential resolution to this dilemma. This research group designed sets of conserved intron spanning primers (CISPs), that have primers located within relatively conserved exons near exon-intron
borders to scan introns for variation suitable for DNA marker identification. The marker types envisioned with this approach included SNPs and INDELS. Primer design was facilitated by aligning EST sequences from buffelgrass (*Pennisetum ciliare*), pearl millet (*Pennisetum glaucum*), and sorghum (*Sorghum bicolor/Sorghum propinquum*) with fully annotated target segments of rice. Therefore, this approach uses rice as a reference genome, and searches ESTs from other Poaceae family members for conserved exonic sequences for primer design. Of 384 designed primer sets 32% amplified successfully in *Oryza*, *Sorghum*, *Pennisetum*, and *Cynodon*, 88% in *Zea*, 50% in *Eragrostis*, 47% in *Triticum*, and 50% in *Hordeum*. These genera represent a wide range of lineages within Poaceae, illustrating the utility of these markers outside of the lineages used for primer design.

In addition to the CISP primers described above, other marker types that target exon-intron borders for primer design and subsequent intron scanning have been developed. This includes intron length polymorphisms (ILPs) using rice as a reference genome (Wang et al., 2005), conserved ortholog set (COS) markers using Arabidopsis as a reference genome (Fulton et al., 2002), and comparative anchor tagged sequences (CATS) (Lyons et al., 1997) and exon-primed intron-crossing PCR (EPIC-PCR) (Palumbi and Baker, 1994) using humans as a reference genome. The great utility of all of these types of markers stems from their potential to provide basic genetic information in ‘orphan’ taxa lacking primary molecular genetic resources.
Applications of DNA markers in plant breeding

As mentioned previously, much of the success of plant breeding has been accomplished without fundamental knowledge of plant biology and genetics. Plants were domesticated and significantly improved prior to Mendal’s discovery of the gene, Darwin’s articulation of his theory of natural selection, and even the elucidation of the structure of DNA by Watson and Crick. Furthermore, as Lee pointed out in 1995, “most important biological phenomena incorporated into plant breeding programs (e.g. heterosis, epistasis, host-pest interaction, response to abiotic stress) are described in abstract or anecdotal concepts”. This is probably still true today for almost all plant breeding programs except for a select few well characterized model crops and/or examples of detailed molecular genetic dissection of certain specific traits and characteristics. It is obvious that plant breeding as a science can make significant gains without complete elucidation of complex molecular genetic problems, but “…basic plant biology will be a source of much new information about genomes, genes, pathways, and interactions of direct relevance to crop improvement” (Lee, 1995). Modern molecular biology techniques, such as DNA markers, provide new tools that can facilitate development of improved plant breeding methods and enhance our knowledge of complicated and heretofore poorly understood plant genetic phenomena. DNA markers have been, and will continue to be a vital link between fundamental biological processes and classical plant breeding endeavors (Lee, 1995).

According to Ribaut et al. (2002) the application of molecular markers in plant breeding programs can be divided into three broad categories: “1) the characterization of germplasm, known as fingerprinting; 2) the genetic dissection of the target trait – actually
the identification and characterization of genomic regions involved in the expression of
the target trait [genetic mapping; linkage analysis; QTL analysis]; and 3) following the
identification of the genomic regions of interest, crop improvement through marker-
assisted selection (MAS)”. Numerous subcategories exist within each of these three
broad categories. The following sections review specific applications of DNA markers in
plant breeding programs within the context of the three main categories proposed by
Ribaut et al.

**Use of Molecular Markers for Germplasm characterization**

The foundation of plant breeding is largely dependant upon the availability and
knowledge of the current genetic resources available within the plant breeding program
(Lee, 1995). Almost every collection of crop germplasm has been categorized to provide
organization, structure, standards, context, and direction in plant breeding programs (Lee,
germplasm collections within breeding programs should be characterized. These include
the following:

1). Manage germplasm collections (acquisition, maintenance, characterization)
2). Gain in depth knowledge of current genetic content
3). Determine to what extent the gene pool of the collection overlaps with natural
variation
4). Avoid costly maintenance of duplicate or similar accessions
5). Establish ecogeographic distributional patterns to guide the development of future sampling/collections (data driven acquisition)
6). To establish genetic relationships between existing material
7). Identify elite germplasm
8). Develop rationale for parent selection (characterization of relationships and diversity management for the express purpose of genetic gain)
9). Germplasm protection (Plant breeders rights)

A simpler, and broader, treatment of reasons why crop germplasm should be characterized is provided by Paterson et al. (1991): germplasm identification is carried out to 1) determine relatedness and 2) to portray diversity. Historically, germplasm characterization has been achieved through careful phenotypic evaluation of agricultural traits (yield, disease reaction, abiotic stress tolerance, etc.), reproductive biology, morphology, breeding behavior, and detailed record keeping of crop pedigrees. However, morphological variability is often limited (especially among closely related accessions), characters may not be obvious at all stages of plant development, and the appearance of the characters may be affected by environment (Tanksley and McCouch, 1997). DNA molecular markers can be used in breeding programs to streamline and/or augment the difficult challenge of accurate and thorough germplasm characterization, and can be considered an ideal tool to determine relatedness and portray diversity.

Genetic diversity among organisms is a result of variations in DNA sequences. In fact, molecular genetic variation, such as nucleotide sequence variation, provides the
fundamental description of genetic variation in an organism (Hedrick, 2000). Molecular
genetic variation is substantial (Paterson et al., 1991) and for numerous organisms each
individual of a species possesses a unique DNA sequence (Jasieniuk and Maxwell, 2001).
DNA variations in their simplest form are mutations resulting from substitution of single
nucleotides (single nucleotide polymorphisms – SNPs) or allelic variance at a single
nucleotide position. Additional forms of DNA variation include insertion or deletion of
DNA fragments of various lengths (from a single to several thousand nucleotides), or
duplication or inversion of DNA fragments.

Fingerprinting and genetic profiling provide a vast abundance of information
derived from DNA ideal for population studies and species identification (Morgante and
Olivieri, 1993). The characterization of marker profiles provides reliable, direct and
efficient strategies for assessing genetic variation of germplasm. Therefore, DNA
variations can be used to define taxonomic structure – i.e. determining differences
between individual genotypes, populations, species, genera etc. Ultimately, molecular
genetic diversity is measuring similarity-dissimilarity at the DNA level, and is a measure
of “that difference between two entities that can be described by allelic variation”. This
latter definition was provided by Nei (1973) and termed genetic distance (GD). Genetic
distance is therefore a statistical model for estimating the accumulated number of codon
differences per locus and the divergence time between individuals (Nei, 1973). Nei
(1987) later elaborated this definition as “the extent of gene differences… between
populations or species that is measured by some numerical quantity”. A more
comprehensive definition of genetic distance is “any quantitative measure of genetic
difference, be it at the sequence level or the allelic frequency level, that is calculated between individuals, populations, or species (Beaumont et al., 1998).

Genetic distances can be calculated in various ways depending on the kind of data. Thus, genetic variability within a population can be measured as 1) the number (and percentage) of genes in the population that are polymorphic; 2) the number of alleles for each polymorphic gene; and 3) the number (and percentage) of genes per individual that are polymorphic. The genetic and mathematical properties of various genetic distance measures have been reviewed extensively by Reif et al. (2005). This review showed that genetic and mathematical properties of dissimilarity measures are of crucial importance when choosing a genetic dissimilarity coefficient for analyzing molecular diversity data. In another review, Mohammadi and Prassana (2003) presented some salient statistical tools and considerations that need to be taken into consideration when analyzing genetic diversity in crop plants. For reasonably accurate and unbiased estimates of genetic diversity, adequate attention has to be devoted to 1) sampling strategies; 2) utilization of various data sets on the basis of the understanding of their strengths and constraints; 3) choice of genetic distance measure(s), clustering procedures, and other multivariate methods in analyses of data; and 4) objective determination of genetic relationships.

According to Weising et al. (2005), the analysis of plant cultivars with molecular markers often follows a common theme. In the first stage of the investigation, different types of markers are tested for their ability to identify a cultivar [or individual] unequivocally, and/or to discriminate between closely related cultivars [or individuals]. The second phase then usually attempts to ascertain the relatedness between these
cultivars [or individuals]. After the DNA profiles of each individual are recorded, a genetic distance matrix is usually calculated. This matrix is then used for a cluster analysis [or other appropriate grouping analysis], and the results are depicted in a tree, dendrogram, or other appropriate schematic to simplify visualization of genetic relationships.

A vast number of statistics parameters/methods are widely used for pattern analyses of DNA genotypes in plant diversity studies. In general, statistical analyses of DNA profiles are performed by constructing genetic similarity or distance matrices among entities. The appropriate choice of a genetic distance measure is an important component in the analysis of genetic diversity among a set of genotypes. For analysis of molecular marker data, the amplification products may be equated to alleles. Allele frequencies can then be calculated and the data generated can be used to create a binary matrix for statistical analysis. For dominant markers, simple matching coefficients (Sokal and Michener, 1958), Jaccard’s coefficient (Jaccard, 1908), Nei and Li’s coefficient (Nei and Li, 1979), and Modified Rogers’ distance (Rogers, 1972) are commonly used genetic similarity measures where the data is in binary form (0-1; a-b; present-absent). Additional statistics have been developed for co-dominant markers that allow recognition of both homozygous states as well as recognition of the heterozygous state. Additionally, models have been proposed for cases where mixed variable types are measured (e.g., mixtures of binary, ordinal, categorical, and continuous variables) (Gower, 1971). These types of coefficients can be used to combine dominant (binary) and co-dominant (categorical) DNA markers or discrete genotypic and continuous phenotypic variables. Numerous other similarity/dissimilarity statistics have been
reported in the literature, but the above measures have been used more widely than others, or are generally considered as starting points during an initial investigation of a new species/group of species where molecular genetic resources are limited.

Cluster analysis (CA), Principal Component Analysis (PCA), Principal Coordinate Analysis (PCoA) and Multidimensional Scaling (MDS) are among some of the common multivariate methods widely used for genetic diversity studies. These methods seek to uncover hidden patterns among objects on which two or more independent variables have been measured. Cluster analysis aims to group items, in this case genotypes, based on the characteristics that they posses so that individuals with similar descriptions are mathematically gathered into the same cluster. Clustering methods usually lead to a graphical representation such as tree or dendrogram in which clusters may be visually identified (Mohammadi and Prasanna (2003). PCA and PCoA are multivariate techniques used to produce two or three dimensional scatter plots of items so that the geometrical distances among items in the plot reflect the genetic distances among them with little distortion. Grouping of items in scatter plots will reveal sets of genetically similar individuals. MDS is a procedure that represents a set of individuals or genotypes in a few dimensions using a similarity/distance matrix between them such that the inter-individual proximities in the map nearly match the original similarity/distances (Johnson and Wichern, 2002).

There are few, if any, recognized statistical rules for deciding how many genetic markers are needed to accurately classify accessions, describe genetic relationships, or accurately estimate genetic distance across all taxa and in all investigations. Additionally, the number of genetic markers used in an analysis may be dictated by
numerous non-statistical factors (e.g., costs, access to equipment, availability of molecular genetic resources in the study organism). While it is certainly true that the estimates of genetic similarity/dissimilarity between two entities is affected by the number and characteristics of the genetic markers utilized, at some point increasing the number of markers (short of a completely sequenced genome) produces a diminishing return in terms of answering the particular research question under investigation. Nonetheless, under ideal circumstances, genetic markers used for assessing diversity, determining relatedness, protecting intellectual property and classifying unknown genetic materials should be highly polymorphic and dispersed across the genome.

**Germplasm characterization for the protection of plant breeders rights**

A unique aspect of germplasm characterization in plant breeding endeavors is the protection of intellectual property developed within breeding programs. This protection of breeding material is more commonly referred to as plant breeders rights (Weising et al., 2005). The goals of plant breeding are to genetically modify plants in a way that makes them better suited to the needs of mankind. This requires extensive crossing programs and sophisticated selection strategies, making the breeding of new crop varieties time-consuming, expensive, and labor intensive. As an example, Troyer and Rocheford (2002) stated that the development of a new maize hybrid may take up to 5 years and require significant monetary investments in research. They further stated that “Germplasm ownership encourages commercial plant breeding research. Protection is necessary to prevent a competitor from selling the same product without incurring
comparable research expenses”. Additional benefits of plant breeder’s rights include assurance from the earliest stage of development through to farmers, processors, consumers etc. that varieties offered for sale are of the correct genotype and specified origin (Smith, 1997).

One of the earliest attempts at intellectual property protection specifically devoted to plants in the United States, used a trademark approach (Janis and Kesan, 2002). A legislative proposal in the US House of Representatives, introduced by Representative Allen in 1906, would have afforded to “any person who has discovered, originated, or introduced any new variety of plant, . . . and gives and applies thereto a name,” the opportunity to register the name, thereby securing for a twenty-year term “the exclusive right to propagate for sale and vend such variety of horticultural product under the name so registered” (Janis and Kesan, 2002). This bill was rejected based on the fact that the use of a trademark would have extended unintended patent protection to horticultural products (Janis and Kesan, 2002). According to Janis and Kesan (2002), after initial failure of plant variety protection using trademark law, Representative Allen immediately attempted to introduce legislation to amend the existing utility patent statute as to include protection of plant varieties. This attempt failed, and any form of plant variety protection would have to wait for more than two decades.

The first formal protection of plant varieties in the United States began in 1930 with the Townsend-Purnell Plant Protection Act (Crocker, 2003). This act granted patent protection to plant breeders for asexually reproduced (cloned) plant varieties. In order to qualify for this patent, a variety had to be novel, distinct and non-obvious (Crocker, 2003). One of the unique aspects of this law, relative to a utility patent, was that there
was no requirement for a written description of the ‘invention’ (a utility patent required a written description which must enable others to ‘make and use’ the invention after the protection expired) (Crocker, 2003). In lieu of a written description, a plant specimen could be used as a reference for the patent (Crocker, 2003). A plant patent under this act gave the patent holder protection for 20 years against others ‘making, selling or reproducing the variety’ (Crocker, 2003). Although this act was a milestone for plant breeders rights, it afforded protection only to asexually propagated material. Janis and Kesan (2002) suggest that the distinction between asexual and sexual plant material was not only made due to the difficulty of determining the exact genetic make-up of some sexually reproducing plants, but also due to political considerations at the time: Nurserymen, whose products were asexually propagated, had a strong lobbying influence, while seed companies at that time acted primarily as brokers rather than inventors of new varieties. Additionally, as pointed out by Louwaars et al. (2005), the exclusion of sexually reproducing plants essentially ignored potential food security issues (i.e. food security should not be the subject of commercial monopolies) with many staple food crops. Further evidence that food security was a concern at the time that the legislation was enacted was the specific exclusion of sexually or asexually propagated edible tubers (Louwaars et al., 2005).

The first attempts to develop a uniform set of procedures and protection of plant breeder’s rights at the international level began in Europe with a conference of thirteen European countries and representatives of the Food and Agricultural Organization of the United Nations (FAO) in 1961. This conference resulted in the signing of the International Convention for the Protection of New Varieties of Plants in 1961. The
convention is more commonly referred to as the UPOV convention, which stands for ‘Union internationale pour la protection des obtentions végétales’. The goal of UPOV, as established by the convention, is ‘to provide and promote an effective system of plant variety protection, with the aim of encouraging the development of new varieties of plants, for the benefit of society’ (UPOV, 2008a). The UPOV Convention sets minimum worldwide recognized principles/guidelines for protection of plant breeder’s rights which member countries account for when developing their own legislation (Smith, 1997; UPOV, 2008b).

The UPOV convention entered into force in 1968, and was amended in 1972, 1978 and 1991 (UPOV, 2008b). The purpose of the 1961 Act was to ‘recognize and ensure’ the rights of breeders of new plant varieties (UPOV, 1961). The act allowed for member countries to grant protection for a plant variety or to patent a plant variety, however, only one form could be granted for “one and the same botanical genus or species” (UPOV, 1961). The term plant variety in the 1961 Act referred to ‘any cultivar, clone, line, stock or hybrid which is capable of cultivation (UPOV, 1961). Therefore, this was the first act to declare protection for both sexually and asexually reproduced varieties. According to article 6 of the 1961 Act, in order to qualify for protection, a new variety must meet the following criteria: 1) “Whatever may be the origin, artificial or natural, of the initial variation from which it has resulted, the new variety must be clearly distinguishable by one or more important characteristics from any other variety whose existence is a matter of common knowledge at the time when protection is applied for.”; 2) “The new variety must be sufficiently homogeneous, having regard to the particular features of its sexual reproduction or vegetative propagation.”; and 3) “The new variety
must be stable in its essential characteristics, that is to say, it must remain true to its
description after repeated reproduction or propagation or, where the breeder has defined a
particular cycle of reproduction or multiplication, at the end of each cycle.”. A more
common vernacular that is used to describe these criteria is that a new variety must be
distinct, uniform, and stable (DUS) (Smith, 1997). The 1961 act specified that protection
of varieties must be for a minimum period of 15 years for most plants, while vines, fruit
trees and their rootstocks, forest trees and ornamental trees be protected for a minimum
period of 18 years.

An interesting exception to protection of plant breeder’s rights was included in the
1961 Act. Article 5 states: “Authorization by the breeder or his successor in title shall
not be required either for the utilization of the new variety as an initial source of variation
for the purpose of creating other new varieties or for the marketing of such varieties.
Such authorization shall be required, however, when the repeated use of the new variety
is necessary for the commercial production of another variety”. Although variety
improvements through plant breeding practices can sometimes be dramatically distinct or
novel, most plant breeding innovations are, in practice, considered incremental (Bijman
and Joly 2001). Generally, newly improved cultivars are used as a foundation to create
additional improvements in still other cultivars. The ‘breeders exception’ was included
so that all breeders would have access to newly improved plant material to create
continual improvements from the best material available at that time.

The 1972 revision of UPOV concerned procedural changes related to council
voting and methods for financial contributions by member nations (UPOV, 1972). The
1978 revision provided a few more significant changes (UPOV, 1978): It gave UPOV
limited legal status as an intergovernmental organization in member countries, and an article was added that allowed member states the ability to apply for an exception to the prohibition of dual protection (plant protection and patent).

The 1991 UPOV revision was more significant than the previous two revisions, in that additional protection for breeders was provided (UPOV, 1991). The prohibition of dual protection was removed from the act. This allowed all member states the ability to protect plant varieties by patents as well as plant breeders’ rights. The DUS concept was specifically written into the language of the act in Chapter III, Articles 5-9 (UPOV, 1991): “The breeder’s right is granted where the variety is (i) new, (ii) distinct, (iii) uniform and (iv) stable”. Distinctness was defined as: “The variety shall be deemed to be distinct if it is clearly distinguishable from any other variety whose existence is a matter of common knowledge at the time of the filing of the application”; uniformity was defined as: “The variety shall be deemed to be uniform if, subject to the variation that may be expected from the particular features of its propagation, it is sufficiently uniform in its relevant characteristics”; and stability was defined as “The variety shall be deemed to be stable if its relevant characteristics remain unchanged after repeated propagation or, in the case of a particular cycle of propagation, at the end of each such cycle” (UPOV, 1991).

Additionally, the ‘rights of the breeder’ (Chapter V, Article 14 – UPOV, 1991) were expanded and more clearly defined than in previous versions. In respect to the protected variety, authorization of the breeder is required for the following activities under the 1991 Act:

i. production or reproduction (multiplication),
ii. conditioning for the purpose of propagation,

iii. offering for sale,

iv. selling or other marketing,

v. exporting,

vi. importing,

vii. stocking for any of the purposes mentioned in (i) to (vi), above

Exceptions to the breeders right’s were still granted for “(i) acts done privately and for non-commercial purposes, (ii) acts done for experimental purposes and (iii) acts done for the purpose of breeding other varieties. This was a strong reaffirmation of the ‘breeders exception’, in that the language was not removed from the act (Janis and Smith, 2007).

In fact, UPOV has stated that… “This exception, for the purpose of breeding other varieties, is a fundamental aspect of the UPOV system of plant variety protection and is known as the "breeder's exemption". It recognizes that real progress in breeding—which, for the benefit of society, must be the goal of intellectual property rights in this field—relies on access to the latest improvements and new variation. Access is needed to all breeding materials in the form of modern varieties, as well as landraces and wild species, to achieve the greatest progress and is only possible if protected varieties are available for breeding”. (UPOV, 2008c). However, in respect to (iii) above [known as the ‘breeder’s exception’] the 1991 act introduced the concept of essentially derived varieties (EDV). Plant breeder’s rights were extended to EDV’s with the following language from Article 14 of the act:
The concept of EDV placed limitations on the ‘breeder’s exception’ from previous versions of UPOV. The EDV provision provided a means to manage and restrict commercialization of close copies or plagiarized versions of initial varieties (Smith, 1997). Additional protection for breeder’s in the 1991 act included extension of minimum time periods for protection to 25 years for vines and trees and 20 years for other species (UPOV, 1991).

In response to the creation of the UPOV convention in Europe, the United States passed legislation known as the Plant Variety Protection Act of 1970 (PVP) (Crocker, 2003). This act provided patent-like protection for plant varieties through certificates administered by the USDA instead of by the Patent and Trademark office (Crocker, 2003). Just as indicated in the UPOV acts, new plant varieties need to be distinct,
uniform, and stable (Crocker, 2003). A written description of the invention is not required. In place of a written description the breeder can deposit a seed sample to provide evidence for novelty (Crocker, 2003). The PVP act was the first plant breeder’s rights legislation in the US that included protection for sexually reproducing plants. Protection was extended to sexually reproducing plants as long as they produced the same characteristics over several generations (Crocker, 2003). Plant variety protection was granted to the breeder for a period of twenty years, which prevented others from selling, offering for sale, reproducing, importing or exporting the variety (Crocker, 2003).

As was the case for early versions of UPOV, The Plant Variety Protection Act of 1970 provided an exemption for research that allowed other breeders to use the protected variety for the development of new varieties (Crocker, 2003). Another exemption allowed farmers to save seed for replanting and to sell seed locally if State laws permitted such sales (Crocker, 2003). Both these exemptions were modified with a revision to the PVP act in 1994: farmers were restricted to saving seeds to replant on their own acres; and the research exemption was altered to include the provision that essentially derived varieties were considered an infringement on the PVP protection granted to the original variety (Crocker, 2003).

An obvious question at this point is why was there a need for protection of plant varieties under a system different than patent protection offered through the Patent and Trademark office. Two issues were primarily responsible for this situation (Bai, 1997): The first issue was that at the time of the PVP act, living organisms were not regarded as patentable under the guidelines for utility patents in the United States; The second issue dealt with the qualification process for utility patents. To qualify for a utility patent, the
invention must be new, non-obvious, useful and meet the written description requirement. The requirement of the written description had been an obstacle for traditional plant breeding ‘inventions’ (Bai, 1997).

Three separate legal actions, all dealing with slightly different issues, eventually led to a situation where plant breeders could apply for utility patents for their inventions. The first case, *Diamond v. Chakrabarty*, was decided by the US Supreme Court in 1980, and granted the right to a utility patent for a genetically engineered bacterium capable of degrading crude oil. This was the first example of a utility patent conferring intellectual property protection for a living organism (Crocker, 2003). The second case, *Ex parte Hibbard*, was brought before the US Patent and Trademark Office in 1985 and dealt specifically with plant material (tryptophan overproducing maize) (Kelmelyte, 2005). The decision stated that the protection afforded under the Plant Protection Act and the Plant Variety Protection Act did not exclude plants from obtaining utility patents (Kelmelyte, 2005). The third case was *J.E.M. Ag Supply v. Pioneer Hi-Bred* in 1995, where the US Supreme Court confirmed that plants could receive utility patents (Bai, 1997). All three cases potentially strengthen plant breeders rights, as utility patents offer greater legal protection than PVP certificates due to the extensive case law supporting the rights of holders of utility patents (Crocker, 2003).

The criteria for protection of a plant variety under PVP and UPOV centers around the DUS rules applied to phenotypic characteristics (Janis and Smith, 2007). To standardize the application of protection, UPOV has developed guidelines for relevant phenotypic characteristics for over 250 species for ‘distinctiveness’ assessment (Janis and
Smith, 2007). The guidelines specify measurement of characteristics in both subtle and minute detail (Janis and Smith, 2007), and typically require anywhere from 15 to 50 measurements depending upon the plant species under investigation (van Eeuwijk and Baril, 2001). Lists of ‘example varieties’ which demonstrate the various states of a characteristic are also available through UPOV (Cooke and Reeves, 2003). Typically, for DUS testing, the characteristics are field-based observations of morphological aspects of the phenotype (greenhouse or laboratory testing can be used where appropriate) (Cooke and Reeves, 2003). This process requires significant resource input in land use, labor, time and analysis.

Cooke and Reeves (2003) state that there is significant interest in the use of molecular markers for this process, the objective being to produce a system that is rapid, cost effective, objective as possible, using varietal characteristics that can be readily assessed and are not unduly influenced by the environment. To this end, various research groups have investigated the use of molecular markers specifically for DUS or EDV testing, using a number of molecular marker systems in a number of different plant species. Examples of these types of investigations include molecular markers available for wheat (Law et al., 1998; Law et al., 1999; Donini et al., 2000; Vosman et al., 2001), barley (Koebner et al., 2003), oilseed rape (Lee et al., 1996; Lombard et al., 2000; Tommasini et al., 2003), Calluna vulgaris (Borchert, 2008) sugarbeet (De Riek et al., 2001), soybeans (Giancola et al., 2002), rice (Choudhury, 2001; Singh et al., 2004), tomato (Vosman et al., 2001), and potato (Corbett et al., 2001).

Although the above examples show that attempts have been made to use molecular markers for DUS and EDV testing, Cooke and Reeves (2003) point out that
while molecular markers hold promise for variety protection, there are a number of issues that have and may continue to affect implementation of such a system, such as:

1). The number of markers that should be used

2). Whether or not the distribution of the markers within the genome is important

3). Whether or not it is important that the markers are mapped

4). Are standardized methods of marker analysis available?

5). Are the suggested markers publicly available?

6). While it is relatively well documented that markers can discriminate between varieties (and thus might be able to demonstrate ‘D’), what about the ‘U’ and ‘S’ aspects?

7). Would the use of markers inevitably reduce the ‘minimum distance’ between distinct varieties?

Other authors, industry groups, or governing authorities have raised similar or additional concerns. One of the main concerns is related to standardization and the development of protocols (UPOV, 2000; Vosman et al., 2001; Cooke et al., 2003). The International Seed Federation (ISF) has expressed a preference for continuing phenotypic assessment, and has also pointed out that DNA marker profiles are not yet predictive of most phenotypic characteristics due to a lack of genetic linkage information or to the relatively complex genetic control of quantitative phenotypic traits (ISF, 2003). On the other hand, in some instances it has been demonstrated that phenotypic characteristics using traits that were specifically selected by UPOV for DUS testing do not necessarily provide reliable estimates of genetic distance (Burstin and Charcosset, 1997; Dillman et al., 1997; Rebourg et al., 2001).
With strong arguments on either side of the issue, it may turn out that the implementation of molecular marker data for DUS testing will be decided on a case by case basis (certain species under certain circumstances). As technology and methodology continue to improve, and molecular genetic resources in orphan crops continue to increase, the adoption of molecular markers may be thought of as an additional tool to be used in DUS and EDV testing. This last point was the position of ASSINSEL (ASSINSEL, 2000) before forming the ISF, as well as additional scientists (e.g. Lombard et al., 2001; Bernet et al., 2003).

Use of Molecular Markers for Genetic mapping

The process of crop improvement begins with the selection of elite parental lines that possess the desired traits of interest to the plant breeder. The next step in the process involves making crosses between the elite parents to generate a population that is segregating for the trait(s) of interest. Progenies that exhibit the desired trait(s) are then used for additional crosses as improved breeding material is advanced from one generation to the next. Traits of interest may be controlled by single (or a few) genes with major effects. These types of traits are referred to as qualitative traits, and have the potential to be easily recognizable by the breeder and separated into discrete classes of phenotypic characters. Qualitative, recognizable phenotypic traits simplify the selection process, and improvements through traditional breeding methods can be relatively straightforward. Unfortunately, many important agronomic traits (such as yield or response to environmental stresses) are controlled by several to many genes at different loci, all making a small contribution to the observed phenotype, and are not inherited in
simple Mendelian terms [discrete characters] (Sleper and Poehlman, 2006). These traits are referred to as quantitative traits, and can be difficult to study because the observed phenotype does not give direct insight into the underlying genotype. Additionally, the expression pattern of genes controlling quantitative traits can be greatly influenced and complicated by the environment (Lynch and Walsh, 1998). Therefore, plant breeders are often confronted with the problem of trying to improve a trait that is controlled by many genes through traditional breeding practices. As a result, improvement of quantitative traits through traditional breeding can be time consuming, and practical genetic gains are harder to achieve through the slow process of combining numerous genes and/or gene combinations into elite breeding material.

Due to the difficulties of breeding for quantitative traits by conventional means, geneticists and breeders have long considered the potential use of genetic markers to identify chromosomal regions containing genes that are involved in quantitative (and qualitative) inheritance. As mentioned previously, genetic markers are genetic differences between individual organisms, species etc. Generally, genetic markers do not represent the target gene(s) themselves but rather act as ‘signs’ or ‘flags’ (Collard et al., 2005). Genetic markers that are located in close proximity to the gene(s) of interest may be referred to as gene ‘tags’ (Collard et al., 2005). The markers, or tags, do not necessarily affect the phenotype of the trait(s) of interest because they are located only near or ‘linked’ to genes controlling the trait (Collard et al., 2005). The concept of using genetic markers in plants to track inheritance of traits is not new. Sax (1923) found that seed weight in Phaseolus was “associated or linked with one or both of the eye factors [seed coat pattern], with eye pattern factors, and with factors which determine the color
of the pigment”. Sax concluded that because certain size factors [seed weight] could be found linked with factors for qualitative characters it should be possible to study independently the size factor or factors within each linkage group”.

The fact that ‘factors’ [genes] can be ‘linked’ is a violation of Mendel’s principle of independent assortment for the joint behavior of alleles at two loci (Borecki and Province, 2008). This violation was a fortuitous discovery by T. H. Morgan (1911) in that it laid the foundation for linkage analysis and genetic mapping (Borecki and Province, 2008). Morgan (1911), when studying sex-linked traits in Drosophila melanogaster stated:

“In place of attractions, repulsions and orders of precedence, and the elaborate systems of coupling, I venture to suggest a comparative simple explanation based on results of inheritance of eye color, body color, wing mutations and the sex factor of femaleness in Drosophila. If the materials that represent these factors are contained in the chromosomes, and if those factors that “couple” be near together in a linear series, then when the parental pairs (in the heterozygote) conjugate like regions will stand opposed. There is good evidence to support the view that during the strepsinema stage homologous chromosomes twist around each other, but when the chromosomes separate (split) the split is in a single plane, as maintained by Janssens [Janssens, 1909]. In consequence, the original materials will, for short distances, be more likely to fall on the same side of the split, while remoter regions will be as likely to fall on the same side as the last, as on the opposite side. In consequence, we find coupling in certain characters, and
little or no evidence at all of coupling in other characters; the difference depending on the linear distance apart of the chromosomal materials that represent the factors. Such an explanation will account for all the many phenomena that I have observed and will explain equally, I think, the other cases so far described. The results are a simple mechanical result of the location of the materials in the chromosomes, and of the method of union of homologous chromosomes, and the proportions that result are not so much the expression of a numerical system as of the relative location of the factors in the chromosomes. Instead of random segregation in Mendel’s sense we find “associations of factors” that are located near together in the chromosomes”.

According to the above description by Morgan, genetic linkage can be used as a means to determine the position of two loci relative to each other. This situation can be resolved based on observed recombination between marker loci in an experimental cross. More specifically, the products of meiosis (progeny genotypes) can be compared to parental genotypes to calculate a recombination frequency between two marker loci (Suzuki et al., 1989). According to Mendelian independent assortment, genetic markers on different chromosomes segregate independently, and the two recombinant classes always make up 50% of the progeny (Suzuki et al., 1989). Conversely, two marker loci that are close to each other on the same chromosome will tend to cosegregate; i.e. these markers will be inherited together (Weising et al., 2005). Markers that co-segregate (are always present or absent together) must be linked; i.e. they must be located in each others vicinity on the genome, and they will usually be inherited together in the majority of the progeny.
(Weising et al., 2005). In some cases, however, due to crossing-over events, the linkage between the markers will be lost, and these markers will not be inherited together in a certain percentage of the progeny. The frequency with which the linkage between co-segregating markers is broken increases with increasing distance between the markers on the chromosome. The recombination frequency for two markers on the same chromosome can range from 0% to 50%. In the case of genetic linkage, the recombination frequency will be lower than 50%, where a low percentage indicates relatively close proximity between the marker loci (strong linkage), while a high percentage (closer to 50% recombination) indicates greater distance between the marker loci. The distance is expressed in centimorgans (cM), where one cM corresponds to 1% recombination. An extensive analysis of the linkage between a large number of genetic markers yields information on their arrangement on the entire genome. Such analysis can finally result in the construction of a genetic linkage map, on which all markers are arranged in separate linkage groups (which are representations of the individual chromosomes).

All three major types of genetic markers (morphological, biochemical, and DNA or molecular markers) have been used in genetic mapping experiments. Morphological markers were the first generation of markers to be used for identification and selection of qualitative and quantitative traits. Although these early morphological maps were important for locating regions on chromosomes corresponding to phenotypic traits of interest, the usefulness of these types of markers was limited by a number of factors, including low or few numbers of morphological markers available to breeders resulting in poor coverage of the genomes under investigation, and the fact that expression patterns of
morphological markers can be influenced by the environment (Mohan et al., 1997). With the advent of protein analysis, many investigators believed they had found a more reliable source of markers in the allelic variants of enzymes, called isozymes (Winter and Kahl, 1995). Unfortunately, isozyme markers were also found to be limited in number, and their expression was found to be associated with specific tissues or only at specific developmental stages (Winter and Kahl, 1995). Unlike morphological and biochemical markers, DNA markers are practically unlimited in number and are not affected by the environment and/or tissue type and developmental stage of the plant (Collard et al., 2005).

All the data accumulated from scoring a mapping population sequentially with a series of DNA markers are used to construct a linkage map (Weising et al., 2005). Simple statistical tests such as chi-square analysis can be used to test the randomness of segregation, and therefore linkage for a few markers, however, it is not feasible to manually determine linkage analysis between large numbers of markers used to construct detailed maps; computer programs are required for this purpose (Collard et al., 2005). Two hypotheses have to be discriminated: no linkage where recombination frequency = ½ (null hypothesis) and linkage at a recombination frequency < ½ (alternative hypothesis) (Weising et al., 2005). The statistical criterion for linkage between two loci (in this case DNA markers) is based on an odds ratio $L$ (likelihood odds ratio), which represents the ratio of linkage versus no linkage (Collard et al., 2005). This ratio is more conveniently expressed as the logarithm of odds, and is referred to as the LOD value or LOD score (Collard et al., 2005). A LOD score of 3 (odds ratio of 1000:1) indicates that linkage is 1000 times more likely than the null hypothesis of no linkage, and is normally accepted as the least acceptable probability that two loci are linked (Weising et al., 2005). Higher
LOD scores (e.g. 4 or 5) are an indication that linkage is indeed reliable (Weising et al., 2005). If linkage is indicated by chi-square analysis of progeny segregation, then the potential for linkage between loci and map distances can be tested mathematically by an algorithm such as maximum likelihood or least squares regression methods (Weising et al., 2005). Various software packages for linkage analysis (e.g. MapMaker; JoinMap) make use of these statistical procedures (Collard et al., 2005; Weising et al., 2005). Mapmaker performs analysis of many linked loci, using an algorithm that groups markers rapidly into linkage groups and another algorithm for suggesting the best possible order of the markers (Weising et al., 2005). Once a likely order is established, another algorithm compares the strength of evidence for that order as compared with possible alternatives (Weising et al., 2005).

**QTL analysis and mapping**

Genetic factors that are responsible for a part of the observed phenotypic variation for a quantitative trait can be called quantitative trait loci (QTLs). Ideally a breeder would want to positively identify all the genes involved in a particular phenotype of interest; however, due to the sheer number of possible factors involved, epistatic interactions, and other sources of variation, this level of discrimination is usually not possible (Doerge, 2002). As an alternative, a breeder can attempt to identify a QTL, which represents a region on the genome that is partially responsible for the phenotype and could be comprised of one or more functional genes (Falconer & Mackay, 1996). In a process called QTL-mapping, the association between observed phenotypic traits and the presence/absence of alleles of DNA markers that have been mapped onto a linkage
map is analyzed. When it is proven statistically that the correlation between the observed phenotype and the DNA markers did not result from random chance, it is proclaimed that a QTL is detected.

The statistical methods used to identify QTLs include single marker mapping, interval mapping, composite interval mapping, and multiple trait mapping (Doerge, 2002). Single marker tests utilize t-tests, ANOVA and simple linear regression to associate the segregation of a phenotype with respect to a marker genotype, indicating which markers are associated with the quantitative trait of interest (Doerge, 2002). Although straightforward and relatively simple, single marker mapping may not accurately estimate the recombination that occurs between the DNA marker and the putative QTL causing the magnitude of the effect of the QTL to be underestimated (Collard et al., 2005). Interval mapping methods make use of linkage maps and analyze intervals between adjacent pairs of linked markers along chromosomes simultaneously, instead of analyzing single markers (Lander and Botstein, 1989). This method compensates for recombination between the markers and the QTL resulting in a statistically more robust test than single marker methods (Lander and Botstein, 1989; Liu, 1998; Doerge, 2002). The results of the test statistic for interval methods are typically reported as a logarithm of odds (LOD) score (Collard et al., 2005). The position on the map with the highest LOD score indicates the most likely position of the QTL (Collard et al., 2005). A LOD score of 3 is usually chosen as the significance threshold to indicate the presence of a QTL (Collard et al., 2005).
Marker-assisted selection

Marker-assisted selection (MAS), sometimes referred to as marker-aided selection, is a relatively new tool in plant breeding programs. When a DNA marker can be associated with a particular phenotype, the breeder can connect the gene action underlying the phenotype with the distinct regions of the genome in which the gene resides (Winter and Kahl, 1995). In its simplest form MAS may greatly increase the efficiency and effectiveness of plant breeding programs (Collard et al., 2005). Once markers that are tightly linked to genes or QTLs are identified, breeders may use the markers as a diagnostic tool to identify plants carrying the gene or QTL of interest (Collard et al., 2005). According to Collard et al. (2005), the advantages of utilizing MAS include:

1.) time saving from the substitution of complex field trials with molecular tests

2.) elimination of unreliable phenotypic evaluation associated with field trials due to environmental effects

3.) selection of genotypes at seedling stage

4.) gene ‘pyramiding’ or combining multiple genes simultaneously

5.) avoid the transfer of undesirable or deleterious genes

6.) selecting for traits with low heritability

7.) testing for specific traits where phenotypic evaluation is not feasible
Genetic Marker Technology in Turfgrass Science

The development of molecular and genomic tools in turfgrasses has lagged behind similar research efforts in model plants or important food crops. According to Zhang et al. (2006) there are four principle reasons why this has occurred: 1) many turfgrasses are genetically complex, being out-crossing polyploids, and thus represent complicated research targets; 2) many traits of interest to turfgrass breeders and scientists are under complex control of a number of genes (quantitative traits); 3) there is a lack of well characterized genetic materials and absence of repeatable phenotyping protocols for many target traits which limits the application of genomics technologies; and 4) the turfgrass research community has failed to attract major funding for molecular and genomics research from governmental agencies or the private sector. Gresshoff et al. (1998) have described turfgrass researchers as being in a “catch-up” mode in relation to molecular genetic technologies. Additionally, Zhang and Mian (2003) summarized the situation by stating that “…many of the fundamental tools required for forage and turf to benefit fully from the revolution in genomics do not exist or are incomplete”. Despite the lack of significant molecular genetic tools available to turfgrass scientists, there has been some progress in developing limited molecular genetic resources for improvement of these grasses. The following section will provide an overview of research efforts in this area to date.

Genetic and molecular markers to assess turfgrass diversity

An understanding of the genetic diversity present in the cultivated germplasm of a turfgrass species, as well as the location of new sources of genetic variability is important
for the optimal utilization of turfgrass genetic resources (Warnke et al., 1997). Plant breeders must have an understanding of the genetic variability of elite germplasm because continued reselection within this germplasm can narrow the genetic base of elite material and ultimately increase the potential vulnerability to pests and abiotic stresses (Warnke et al., 1997). Information about the location of new sources of genetic variability can help broaden the genetic base of elite material and maintain long-term improvement strategies (Warnke et al., 1997). Additional reasons to assess genetic diversity in turfgrasses include characterization of evolutionary relationships and origin of species (Caetano-Anolles, 1998), to ascertain the genetic diversity within and among natural and cultivated populations (Caetano-Anolles, 1998), to identify ‘off-types’ in turfgrasses that reproduce vegetatively (Caetano-Anolles, 1998) or through apomictic reproduction (Huff and Bara, 1993), to develop practical cultivar blending strategies for end users (Murphy et al., 1997; Bonos et al., 2000), and to establish protocols for variety registration and the protection of plant breeders rights (Curly and Jung, 2004).

**Morphological, protein, and isozyme markers**

Prior to the widespread use of molecular tools, turfgrasses have been the subject of conventional systematic studies using comparative morphological and ecological characters. As an example, Gould and Shaw (1983) classified species used as turfgrasses into 3 subfamilies, 20 genera and more than 30 species based on morphological characteristics alone. Numerous morphological characteristics are still used for varietal registration of turfgrasses in the US and Europe. Examples of morphological traits listed in the USDA’s Plant Variety Protection Application Form include plant height, panicle
height, flag leaf dimensions, subtending leaf dimensions, and rhizome spread and length (Bonos et al., 2000).

The first use of molecular tools for turfgrass genetic diversity studies was conducted by Wilkinson and Beard (1972). These researchers were able to distinguish cultivars of creeping bentgrass and Kentucky bluegrass on the basis of electrophoretic separation of total leaf protein extracts. The first use of isozymes in turfgrass involved the use of peroxidase markers for cultivar identification of creeping bentgrass and Kentucky bluegrass (Moberg, 1972). A few years later, Wehner et al. (1976) expanded this work with a more detailed analysis of peroxidase isoenzyme banding patterns in Kentucky bluegrass cultivars. Work with isozymes for cultivar identification continued for a number of years with additional isoenzymes of esterases, phosphoglucomutases and others in creeping red fescue (Villamil et al., 1982), Kentucky bluegrass (Gilliland et al., 1982; Wu et al., 1984; Weeden and Emmo, 1985; Wu and Jampates, 1986), perennial ryegrass (Hayward and McAdam, 1977; Hayward et al., 1978; Lallemand et al., 1991; Booy et al., 1993), tetraploid ryegrass (Nielsen, 1980; Ostergaard and Nielson, 1981), and creeping bentgrass (Yamamoto and Duich, 1994). Work with isozymes has also been conducted for the characterization of natural populations of perennial ryegrass (Charmet et al., 1993; Balfourier and Charmet, 1994) and red fescue (Livesey and Norrington-Davies, 1991). Additional work with isozymes in turfgrass included the use of isozyme markers to distinguish different species of ryegrass (Charmet and Balfourier, 1994a), to study the influence of ecological factors on population structure and genotypic frequency in perennial ryegrass (Charmet and Balfourier, 1994b), for differentiating selfing versus interspecific hybrids of perennial and annual ryegrasses (Arcioni and Mariotti, 1983), and
for evaluating somaclonal variation in tissue culture derived plants of tall fescue (Dahleen and Eizenga 1990; Eizenga and Cornelius, 1991; Humphreys and Dalton, 1991).

One of the most systematic and detailed evaluation of isozyme polymorphisms in turfgrasses involved multiple loci (phosphoglucose isomerase, phosphoglucomutase, glutamate oxaloacetate transaminase, and triosphosphate isomerase) which generated unique banding patterns for numerous creeping bentgrass cultivars (Warnke et al., 1997). This study was one of the first in turfgrasses to utilize the isozyme banding patterns to create a genetic distance matrix and subsequent Unweighted Pair-Group Method of Analysis (UPGMA) to infer phylogenetic relationships among the cultivars. Interesting results from this study included the finding that much of the US germplasm that was studied had a narrow genetic base and could be divided into two main groupings, while a plant introduction of European origin was clearly genetically distant to the US cultivars.

**RFLP markers**

Genetic characterization of organisms has been dramatically improved by the use of DNA-based markers (Chai and Sticklen, 1998). Genetic polymorphisms revealed by DNA analysis are much more informative than protein-based markers because they are variations in the actual genetic material (Chai and Sticklen, 1998). Additionally, DNA-based markers are also present in greater number, and unaffected by environmental and physiological variability (Chai and Sticklen, 1998). Early DNA-based approaches to detect DNA polymorphisms included RFLPs (Botstein et al., 1980) and detection of low-copy number repeats through Southern hybridization (Wyman and White, 1980). Early use of RFLPs in turfgrass species, from both nuclear and chloroplast DNAs (cpDNA),
were utilized to provide accurate identification and clarification of phylogenetic relationships (Ohmura et al., 1993; Yaneshita et al., 1993a and 1993b). The RFLP work by Yaneshita et al. (1993b) was interesting in that it supported the traditional broad taxonomic treatment of turfgrasses as cool-season (C3 type) and warm-season (C4 type) species. Additional early research characterizing genetic diversity and phylogenetic relationships using RFLPs was conducted in tall fescue and fine fescues (Xu and Sleper, 1994), centipedegrass (Sayavedra-Soto et al., 1997), and a detailed analysis of the Lolium-Festuca complex (Xu et al., 1991; Xu et al., 1992; Xu et al., 1994; Charmet et al., 1997). More recent utilization of RFLPs for characterizing genetic diversity of turfgrasses has been conducted in paspalum species (Jarret et al., 1998), bentgrass (Caceres et al., 2000), perennial ryegrass, meadow fescue, and tall fescue (Yamada and Kishida, 2003), tall fescue (Busti et al., 2004), and buffalograss (Gulsen et al., 2005). Additional utilization of RFLPs for genetic diversity of turfgrasses is likely to decline due to simpler and less time-consuming methodologies and analysis now possible with the polymerase chain reaction.

**AP-PCR markers**

The advent of the Polymerase Chain Reaction (PCR) (Mullis and Faloona, 1987; Saiki et al., 1988) has enabled the development of powerful genetic markers for the measurement of genotypic variation. The advantages of PCR-based genetic markers over other markers such as isozymes and RFLPs has been previously discussed in this review as well as widely detailed in the literature (Arnheim et al., 1990; Weber, 1990b; Rafalski and Tingey, 1993; Morell et al., 1995). Briefly, advantages of PCR include the following
list as adapted from Morell et al. (1995): 1) PCR only requires small amounts of DNA which can be readily recovered from seeds, seedlings or adult plant material; 2) PCR normally involves fewer steps than other methods such as RFLP; 3) PCR is technically less demanding than other methods such as RFLP, and is relatively straightforward once the PCR conditions have been established; 4) PCR does not require the use of radioactivity; 5) PCR can be automated at most stages from DNA extraction to data collection and analysis; and 6) the vast range of potential primer sequences gives the technique great diagnostic power.

Through the mid- to late 1990s, the most frequent application of PCR-based genetic markers in turfgrass had been for the purpose of species and cultivar identification and for defining genetic relationships (Peakall, 1997). With the exception of one study of *Paspalum*, which utilized SSR markers (Liu et al., 1995), all studies through 1997 utilized one of the arbitrary-primed PCR methods (AP-PCR) (Peakall, 1997). By virtue of their applicability to any species, the AP-PCR methods provided a useful entry point for the investigation of turfgrass species that were poorly or completely unknown genetically (Peakall, 1997). The following sections will review the early research to assess genetic diversity in turfgrasses using AP-PCR methods, and then discuss the more recent use of additional marker systems in various turfgrass species.

A few studies have utilized RAPD technology to investigate phylogenetic relationships or turfgrass diversity at the species or genus level. One of the earliest studies to investigate the utility of RAPDs for species identification in turfgrasses was a study by Stammers et al. (1995) that analyzed genetic relationships among the closely related genera *Lolium* and *Festuca*. An interesting finding in this work was that RAPD
analysis showed strong correspondence with the traditional classical taxonomy in these
genera. A notable exception was the placement of meadow fescue (*Festuca pratensis*)
within the genus *Lolium*. The authors indicated that this result could be supported by the
previous research of Jauhar (1976) that showed that this species readily hybridized with
*L. perenne* and *L. multiflorum*. Additional studies utilizing RAPDs at the species or
genus level include work by Charmet et al. (1997) and Polok et al. (1998). Kollik et al.
(1999) looked at both species and cultivar level differences in *Festuca pratensis* Huds.,
*Lolium perenne* L., and *Dactylis glomerata* L. utilizing RAPDs.

The early use of AP-PCR techniques for the identification of turfgrass cultivars
included work in buffalograss (Wu and Lin, 1994), *Agrostis* (Golembiewski et al., 1997),
centipedegrass (Weaver et al., 1995), *Cynodon* (Caetano-Anolles et al., 1995; Caetano-
Anolles et al., 1997), and perennial ryegrass (Huff, 1997). Additional work was
carried out to determine the genetic diversity using RAPDs in natural populations of
seashore paspulum (Liu et al., 1994) and buffalograss (Huff et al., 1993; Peakall et al.,
1995). A few additional interesting early studies with AP-PCR methods included the use
of RAPDs to distinguish biotypes of annual bluegrass from golf course fairways versus
putting greens (Sweeney and Danneberger, 1995), a combination of flow cytometry and
RAPD analysis to distinguish ‘off-types’ or aberrant progeny of Kentucky bluegrass from
the usual apomictic progeny (Huff and Bara, 1993), the use of RAPDs to determine the
stability of tissue culture regenerated meadow fescue (Valles et al., 1993), and the
characterization of dry turfgrass seeds of several species using RAPD analysis (Sweeney
et al., 1996). More recent investigations of turfgrass genetic diversity using AP-PCR
methods include DNA amplification fingerprinting (DAF – variant of RAPD technique
with higher stringency PCR conditions) in *Cynodon* (Anderson et al., 2001; Roodt et al., 2002; Yerramsetty et al., 2005), and RAPD analysis of Zoysiagrass (Weng et al., 2007), creeping bentgrass (Casler et al., 2003), Kentucky bluegrass (Johnson et al., 2002; Curley and Jung, 2004), *Poa trivialis* (Rajasekar et al., 2005), perennial ryegrass (Bolaric et al., 2005a; 2005b), annual ryegrass (Vieira et al., 2004), and *Paspalum dilatatum* (Garcia et al., 2007). The use of AP-PCR methods in these recent investigations was likely chosen as a method of ‘initial analysis’ of new germplasm, as a method of analysis in a species that had limited molecular genetic resources, or as a means of rapidly assessing/comparing germplasm resources in unique ecogeographic regions.

RFLP and RAPD analysis in forages [and turfgrasses] has given valuable information on the genetic structure of populations, optimum bulk sizes and the number of markers required for population discrimination (Forster et al., 2001). However, RAPD analysis is now relatively disfavored due to problems with reproducibility (Jones et al., 1997; Forster et al., 2001). RFLP analysis is highly reproducible and informative but logistically complex for large numbers of samples (Forster et al., 2001). Due to these issues with RFLP and RAPDs, Forster et al. stated in 2001 that analysis based on AFLPs and SSRs will become more important.

**AFLP markers**

The earliest use of AFLPs for turfgrass genetic diversity studies began in the late 1990’s. Two of the earliest analyses were conducted by Roldan-Ruiz et al. (1997) to distinguish species differences in ryegrasses and the work by Zhang et al. (1999) that characterized 27 bermudagrass genotypes into three major clusters that agreed with
known breeding histories. Since these early studies a number of additional investigations utilizing AFLPs have been conducted in several different turfgrass species:

*Cynodon*:

- Karaca et al. (2002) used AFLPs to characterize genetic diversity in forage bermudagrass species.
- Ho et al. (2003) utilized AFLPs to distinguish ‘off-types’ and differences in Malaysian golf course samples of *Cynodon dactylon* versus Australian Cynodon samples.
- Wu et al. (2004) used AFLPs to study the diversity of Cynodon dactylon accessions from Africa, Asia, Australia, and Europe.
- Wu et al. (2005) studied the genetic diversity and relationships among and between Cynodon transvaalensis, *C. dactylon*, hexaploid *C. dactylon* and their interspecific hybrids.
- Wu et al. (2006) used AFLPs to study the diversity of *C. dactylon* accessions collected from China.
- Kang et al. (2008) used AFLPs to study the diversity of Cynodon accessions collected from South Korea.
- Ludovic et al. (2005) utilized AFLPs to track dinitroaniline-induced genetic changes in six bermudagrass cultivars. This work was interesting in that it proposed a hypothesis about the origination of bermudagrass vegetative or somatic mutation ‘off-types’ from putting green cultivars – chronic exposure
to dinitroaniline herbicides (pendimethalin and oryzalin), which have a mode of action that disrupts normal microtubule formation during meiosis.

*Agrostis:*

- Vergara and Bughrara (2003) conducted an extensive investigation of 14 species of *Agrostis* from twenty countries utilizing AFLPs. This work was significant in that it clarified a number of confusing taxonomic classifications among and between species that had been based solely on cytological or morphological characters.
- Vergara and Bughrara (2004) utilized AFLPs to assess relationships between creeping bentgrass cultivars and PI accessions as well as a limited number of hexaploid redtop cultivars and accessions.
- Zhao et al. (2006) used AFLPs to determine the genetic diversity of colonial bentgrass cultivars and accessions from the USDA germplasm collections and plant material collected from northern Spain.

*Lolium:*

- Roldan-Ruiz et al. (2000a) established that AFLPs could be useful for discriminating perennial ryegrass cultivars in an initial screening of different primer combinations in two cultivars.
- Roldan-Ruiz et al. (2000b) showed that AFLPs could be used as a basis to establish ‘initial variety’ and ‘essentially derived variety’ concepts among twelve accessions of perennial ryegrass with known breeding lineages. This
work was interesting in that AFLP results were consistent with previous
studies that had established varietal relationships based on morphological
data.
- Cresswell et al. (2001) utilized AFLPs to assess genetic diversity of 127 plants
of Lolium perenne, Lolium multiflorum, and hybrids between the two species
collected from Portugal
- Skot et al. (2002) utilized AFLPs to assess genetic diversity of 47 wild
populations of perennial ryegrass across temperature clines in Europe; this
work was interesting in that regression analysis showed that multiple bands
generated through AFLP had a positive relationship with cold temperature
tolerance
- Ghesquiere et al. (2003) used AFLPs to assess the relatedness of 162
populations of ryegrasses from European gene banks
- Kalb et al. (2003) characterized 92 L. multiflorum varieties, breeding lines,
accessions, and inbred lines and four L. perenne varieties using AFLPs
- Muylle et al. (2003) genotyped 105 L. perenne plants with AFLPs and SSRs

Poa:
- Bertoli et al. (2003) characterized genetic diversity of sixteen indigenous
populations of Kentucky bluegrass from northern and central Italy using
AFLPs
- van Treuren (2008) utilized AFLPs to characterize 11 Dutch reference
cultivars and 59 genotypes of Kentucky bluegrass from undisturbed Dutch
grasslands. An interesting result of this study was the finding that natural collections of Kentucky bluegrass showed greater genetic diversity than the reference cultivars indicating that novel genetic variation can be obtained for breeding purposes from native populations.

- Renganayaki et al. (2001) used a combination of AFLPs and RAPDs to assess the genetic diversity in 28 Texas bluegrass (*Poa arachnifera* Torr.) genotypes.

**Festuca:**
- Mian et al. (2002) characterized a total of 36 tall fescue genotypes with AFLPs.
- Fjellheim and Rognli (2005) investigated the relatedness Norwegian meadow fescue populations and cultivars using AFLPs.
- Majidi et al. (2006) characterized 34 Iranian fescue accessions, representing four species utilizing AFLPs.

**Paspalum:**
- Espinoza et al. (2006) utilized AFLPs to characterize 31 accessions of bahiagrass (*Paspalum notatum* Flugge).

**SSR markers**

The use of SSRs for the study of turfgrass diversity began in the mid- to late 1990s. The first studies were conducted by Liu et al. (1995) in seashore paspalum and Kubik et al. (1999) in perennial ryegrass. As these were some of the first studies to
utilize SSRs in turfgrass species, a significant effort was put forth to characterize the SSR abundance and type of polymorphisms present in these two turf species. Liu et al. (1995) also included a phylogenetic assessment of a group of seashore paspalum ecotypes that had been previously assessed using RAPD markers. In general, the phylogenetic analysis of SSRs in seashore paspalum was in agreement with the previously obtained RAPD analysis. The ryegrass work was later supplemented with a study of genetic diversity in seven perennial ryegrass cultivars (Kubik et al., 2001) utilizing some of the SSRs characterized in the original work (Kubik et al., 1999). Taken together, these papers showed that SSRs could be valuable markers for determining genetic diversity in turfgrass species.

Work with SSRs has continued in turfgrasses, but has been somewhat limited compared with other major crops. This is likely due to the expense associated with creating genomic SSR libraries. Another obstacle has been that while some additional genomic SSRs have been identified in some turfgrass species there have been limited numbers that have been made publicly available to the research community (Mian et al., 2005). A few notable exceptions include 718 unique genomic SSR clones for perennial ryegrass (Jones et al., 2001), 28 SSR primers from *Lolium perenne* genomic libraries (Jensen et al., 2005), 60 SSRs from a genomic library of a cross between *Lolium multiflorum* x *Festuca glaucescens* (Lauvergeat et al., 2005), 511 primer pairs developed from a genomic SSR library of tall fescue (Saha et al., 2006) and two sets of SSR primers for *Zoysia japonica* (32 from Tsuruta et al., 2005; 30 from Ma et al., 2007). Momotaz et al. (2004) utilized the perennial ryegrass SSRs from Jones et al. (2001) for identification of cultivars and accessions of *Lolium, Festuca, and Fetulolium* hybrids. Jensen et al.
(2007) used the previously available perennial ryegrass SSR primers from Kubik et al. (1999; 2001), Jones et al. (2001), Jensen et al. (2005), and Lauvergeat et al. (2005) in 23 different species of Poaceae. Saha et al. (2006) showed that the tall fescue genomic SSRs had high transferability rates across multiple grass species.

Additional sources of SSRs for turfgrass genetic diversity studies have come from data mining of EST libraries (EST-SSRs) of turfgrasses or other Poaceae crops. Saha et al. (2004) developed 157 EST-SSRs from tall fescue ESTs. These EST-SSRs were initially used to characterize 11 grass genotypes of Festuca arundinacea, tetraploid fescue, Lolium perenne, Lolium multiflorum, Festuca pratensis, Triticum, and Oryza (Saha et al., 2004), and later used to characterize 54 genotypes of 12 grass species (turfgrasses included Lolium perenne, Lolium multiflorum, Festuca arundinacea, Festuca pratensis, and Festuca rubra) (Mian et al., 2005). Wang et al. (2006) investigated the transferability of 130 SSRs from wheat, maize, and sorghum to Paspalum species and found a transfer rate of 67.5, 49.0 and 66.8%, respectively. Asp et al. (2007) developed a set of 955 non-redundant EST-SSRs from three genotypes of perennial ryegrass.

Additional genomic or EST-SSRs have been developed for various turfgrass species, but have not been utilized for genetic diversity studies. These SSRs have been utilized for the development of genetic linkage maps. Specific examples of these SSRs will be discussed later in the section pertaining to turfgrass genetic mapping.
Other markers used in turfgrass molecular genetic studies

A few additional marker systems have been employed in turfgrasses. Budak et al. (2004a; 2004b; 2005) have had success characterizing divergent buffalograss (Buchloe dactyloides (Nutt.) Englem) germplasm with sequence related amplified polymorphisms (SRAPs). Additional work on buffalograss germplasm characterization was conducted utilizing peroxidase gene polymorphisms (Gulsen et al., 2007). Taylor et al. (2001) developed sequence tagged site (STS) markers for perennial ryegrass, where polymorphisms were based on restriction digest of PCR products, from primer sets developed from barley, oats, and Phalaris. Muylle et al. (2003) and Lem and Lallemand (2003) also developed STS markers for ryegrasses, however, polymorphisms for these markers were based on length differences of PCR products that were primed in exons and spanned the intervening intron sequence. Inoue and Cai (2004) developed STS markers that amplified in perennial ryegrass, meadow fescue and tall fescue by primer development of end-sequenced RFLP probes from Lolium multiflorum. Budak et al. (2004b) utilized ISSR markers (along with RAPDs, SSRs, and SRAPs) in their buffalograss work, while Pasakinskiene et al. (2001) used ISSRs in the Lolium-Festuca complex and Ghariani et al. (2003) utilized ISSRs to characterize Tunisian perennial ryegrass accessions.

Molecular markers for mapping and QTL analysis in turfgrasses

Genetic linkage maps developed from molecular markers are available for multiple ryegrass species (Lolium spp.), tall fescue (Festuca arundinaceae), meadow fescue (Festuca pratensis L.), zoysiagrass (Zoysia japonica), creeping bentgrass (Agrostis
stolonifera L.), colonial bentgrass (*Agrostis capillaris* L.), Texas bluegrass (*Poa arachnifera* L.), Kentucky bluegrass (*Poa pratensis* L.), bahiagrass (*Paspalum notatum* Flugge) and bermudagrass (*Cynodon* spp.). A number of different marker systems have been utilized in the development of the genetic maps for the above mentioned species. In general, early studies that created these maps first utilized RFLPs, RAPDs, and AFLPs, while SSRs and additional markers have been utilized as these maps have been further developed. The following section summarizes the current status of linkage map construction for these species.

Among all turfgrass species, the ryegrasses have received far more attention than any other turfgrass species in the development and utilization of genetic linkage maps (Fei, 2008). This has occurred primarily for two reasons (Fei, 2008): 1) the genetic maps developed in the ryegrasses studied thus far are from diploid species; and 2) ryegrasses, particularly annual and perennial ryegrasses are pasture grasses that play an important role in animal grazing in Europe, Australia, New Zealand, Japan and other parts of the world. These two factors are important because diploids are easier research targets than many of the other complex polyploid turfgrasses, and the important role of the forage ryegrasses as animal feed has resulted in greater interest and investment in genomic studies (Fei, 2008).

The first genetic map in ryegrasses was developed by Hayward et al. (1994) utilizing an interspecific hybrid of annual ryegrass (*Lolium multiflorum* L.) x perennial ryegrass (*L. perenne* L.). The map was constructed utilizing 61 markers of isozymes, RAPDs, and RFLPs to create 13 linkage groups and locate 10 QTLs involved in flowering time and yield (Hayward et al., 1994). An interesting finding associated with
the QTL analysis was that up to 80% of the variation in inflorescence production in the establishment year could be accounted for by one region of the genome (Hayward et al., 1994). The original map was later enhanced with additional markers (Hayward et al., 1998): the 1998 map giving rise to seven linkage groups comprising 17 isozyme markers, 48 RAPD markers, and 41 RFLP markers. Within the RFLP markers, 34 were produced with cDNA and genomic *Lolium* probes and seven were produced with heterologous probes from oat and wheat (Hayward et al., 1998).

The first predominantly PCR based linkage map in perennial ryegrass was created by Bert et al. (1999) using a one-way pseudo-testcross reference population of the International Lolium Genome Initiative (p150/112) between a highly heterozygous perennial ryegrass genotype and a doubled haploid perennial ryegrass genotype. The map was created with 463 AFLP markers (from 17 primer pairs), three isozymes and 5 EST markers, and covered 930 cM in seven linkage groups (Bert et al., 1999). This same reference mapping population (p150/p112) was later enhanced by Jones et al. (2002a) by adding 124 codominant RFLP markers, of which 109 were heterologous anchor RFLP probes from wheat, barley, oat and rice. The use of heterologous RFLPs as anchor probes allowed the researchers to compare the perennial ryegrass genetic map with those of the Triticeae grasses, oat, and rice. Although map coverage was not sufficient to observe small scale duplications and rearrangements, each ryegrass linkage group contained segments that were syntenic and colinear on a broad scale with the corresponding homoeologous groups of the Triticeae species (Jones et al., 2002a). An interesting finding in this work was that a lower level of synteny and colinearity was observed between perennial ryegrass and oat compared with the Triticeae grasses, despite
the closer taxonomic affinity between these species (Jones et al., 2002a). This reference mapping population has also been used to map an additional 93 SSR loci (Jones et al., 2002b), used in a comparison study of mapping families representing backcross (this reference population) versus F₂ type mapping populations (Armstead et al., 2002), used to map the ryegrass gametophytic two locus system of self-incompatability (S and Z Loci) (Thorogood et al., 2002), used to map two casein protein kinase 2 α-subunit genes (CK2) implicated in light regulation response and circadian rhythm control (Shiozuka et al., 2005), used to map a C-repeat binding factor gene cluster (CBF) implicated in cold acclimation (Tamura and Yamada, 2007), and used to identify QTL of morphological, developmental, and Winter hardiness-associated traits (Yamada et al., 2004), QTL of herbage quality traits (Cogan et al., 2005), QTL of pathogenesis-related (PR) protein-encoding and reactive oxygen species (ROS)- generating gene classes (Dracatos et al., 2008) , and QTL for seed set (component of seed yield) (Armstead et al., 2008).

Additional mapping populations have been developed for perennial ryegrass. Dumsday et al. (2003) developed a mapping population of perennial ryegrass that segregated for resistance to crown rust (Puccinia coronata f. sp. lolii). Utilizing 105 previously map assigned SSR loci and bulked segregant analysis, these researchers were able to identify a QTL for crown rust resistance on ryegrass linkage group 2 (LG2) (Dumsday et al. 2003). The same bulked segregant analysis approach was later used by Muylle et al. (2005a) with previously mapped AFLP markers on an additional perennial ryegrass population segregating for crown rust resistance. In this work two separate regions of ryegrass LG2 were identified as being important for crown rust resistance (Muylle et al., 2005a). Faville et al. (2004) constructed a linkage map for perennial
ryegrass based on a two-way pseudo-testcross mapping population using EST-RFLP and EST-SSR markers. This was the first map in a ryegrass species that was based solely on expressed (gene-derived) markers. This mapping population was later used for gene associated SNP discovery (Cogan et al., 2006), association mapping and linkage-disequilibrium analysis of herbage nutritive quality genes (Ponting et al., 2007), and SNP mapping of candidate genes for pathogen defense response (Dracatos et al., 2008).

Armstead et al. (2002; 2004) developed an F2 *L. perenne* mapping population by selfing an F1 hybrid obtained by crossing individuals from partially inbred lines of the cultivars ‘Perma’ and ‘Aurora’. This mapping population was later used to map 376 novel SSR markers (Gill et al., 2006), and also used to identify QTL for heading date (Armstead et al., 2004), fructan metabolism (Turner et al., 2006), and seed set (Armstead et al., 2008).

Jensen et al. (2005) developed an F2 *L. perenne* mapping population from a cross between a single genotype of the cultivar ‘Veyo’ and a single genotype of the ecotype ‘Falster’. This map was used to identify 5 QTL for vernalization response, and clone a putative orthologue of the wheat *VRN1* gene (Jensen et al., 2005). Plants from this mapping population were later used to test the amplification of 306 novel EST-SSR primer pairs (Studer et al., 2008), for mapping 12 ISSR markers (Pivoriene et al., 2008), and for locating QTL associated with resistance to powdery mildew (*Blumeria graminis* DC) and resistance gene analogues (Schejbel et al., 2008). Muylle et al. (2005b) developed a mapping population segregating for resistance to crown rust and mapped 227 loci (AFLP, SSR, RFLP, and STS markers) that covered 744 cM. This research located four QTL involved in crown rust resistance, two each on LG1 and LG2 (Muylle et al., 2005b). King et al. (2008) recently developed new genomic SSR markers and mapped
them utilizing an F$_1$ population from a cross of one amenity and one forage perennial ryegrass. This work identified two QTL associated with heading date (King et al., 2008). Byrne et al. (2008) developed a two-way pseudo-testcross F$_1$ mapping population using a late flowering genotype ‘J51’ and an early flowering sibling genotype ‘J53’ as parents. This work mapped 54 publicly available SSRs and 6 EST markers and identified QTL for days to heading, spike length and spikelets per spike (Byrne et al., 2008).

Another ryegrass species that has received significant attention for genetic mapping and QTL studies is annual ryegrass or Italian ryegrass (Lolium multiflorum). This species is considered a valuable cool-season forage grass in much of the temperate regions of the world (Studer et al., 2006). The earliest genetic linkage map in L. multiflorum was reported by Fujimori et al. (2000) utilizing a cytoplasmic male sterile (CMS) mapping population and AFLP markers. An early high density linkage map of L. multiflorum was constructed by Inoue et al. (2004a) using RFLPs, AFLPs, and telomeric repeat associated sequence markers. The map covered 1244.4 cM, with an average of 3.7 cM between markers, on seven linkage groups and utilized Poaceae RFLP anchor probes (predominantly from Oat) for the purpose of comparative mapping (Inoue et al., 2004a). This mapping population was later expanded and utilized to assess QTL for lodging resistance and related traits (Inoue et al., 2004b). An additional L. multiflorum mapping population was developed by Studer et al. (2006). This population was used to create a genetic linkage map utilizing AFLPs and SSRs, and to identify QTL for bacterial wilt resistance caused by Xanthomonas translucens pv. graminis (Studer et al., 2006) and crown rust resistance (Studer et al., 2007). Miura et al. (2005) used the CMS reference mapping population previously developed by Fujimori et al. (2000) along with bulked
segregant analysis to map EST-derived CAPS markers and AFLP markers linked to a 
gene for resistance to ryegrass blast (*Pyricularia* sp.) on LG5. This same CMS mapping 
population was later used to map 218 genomic SSR loci (Hirata et al., 2006) and 69 EST-
CAPS markers (Miura et al., 2007).

An interesting approach to mapping and/or QTL analysis has been the creation of 
interspecific hybrids in ryegrasses or the use of chromosome introgressions in the 
*Lolium*/*Festuca* complex. Warnke et al. (2004) reported the use of an annual x perennial 
ryegrass interspecific hybrid mapping population to generate a genetic map consisting of 
235 AFLP markers, 81 RAPD markers, 16 comparative grass RFLPs (from oat and 
barley), 106 SSR markers, 2 isozyme loci and 2 morphological characteristics. This map 
was later expanded to include 120 RFLP markers from cDNA clones of barley, oat, and 
rice, allowing comparison of syntenic regions of ryegrass chromosomes relative to the 
*Tritaceae*, oat, and rice (Sim et al., 2005). This annual x perennial ryegrass mapping 
population was also utilized to identify QTL for seedling root fluorescence (Warnke et 
al., 2004), resistance to gray leaf spot (*Magnaporthe grisea*) (Curley et al., 2005), and 
fiber and protein characteristics (Xiong et al., 2006).

Chromosome introgressions of the *Lolium*/*Festuca* complex were carried out in 
the work by King et al. (2002). In this work a single chromosome of *Festuca pratensis* 
was introgressed into a *Lolium perenne* background for both physical mapping using 
genomic *in situ* hybridization (GISH) and genetic mapping utilizing RFLPs and AFLPs 
(King et al., 2002). This work allowed the researchers to assess differences in genetic 
distances in centimorgans versus physical distances in base pairs along the *F. pratensis* 
substituted chromosome. This study showed that the interrelationship between physical
and genetic maps varies from one part of the chromosome to another (King et al., 2002). This result was in accordance with research in other plants where the frequency of recombination (i.e. genetic distance) along the length of a chromosome has been shown to vary (e.g. Gustafson and Dille, 1992; Werner et al., 1992; Hohmann et al., 1994, 1995; Chen and Gustafson, 1995; Delaney et al., 1995a, b; Mickelson-Young et al., 1995; Gill et al., 1996a, b; Kunzel et al., 2000). Similar work was conducted by Roderick et al. (2003), except that a chromosome segment from Festuca pratensis, conferring crown rust resistance, was introgressed into an Italian ryegrass background. The locus conferring resistance was physically mapped utilizing GISH (Roderick et al., 2003), and later genetically mapped using AFLPs and STS markers (Armstead et al., 2006).

Although not as extensive as the work in ryegrasses, a significant number of linkage maps have been developed for the fescues (Festuca spp.). The earliest linkage map of tall fescue (Festuca arundinaceae Schreb.) was developed by Xu et al. (1995) utilizing 108 RFLP markers, and covered 1274 cM on 19 linkage groups with a marker density of 17.9 cM per marker. The first PCR based linkage map of tall fescue was developed by Saha et al. (2005) utilizing AFLPs, EST-SSRs and genomic SSR markers. The integrated map (both male and female parents) covered 1841 cM on 17 linkage groups and had an average marker density of 2.0 cM per marker (Saha et al., 2005). The first linkage map of meadow fescue was developed by Chen et al. (1998) using 66 heterologous RFLP probes from tall fescue. The map covered 280.1 cM on seven linkage groups (Chen et al., 1998). Comparison of 33 of the markers that mapped to meadow fescue in this study, and tall fescue in the work by Xu et al. (1995) detected highly conserved linkage groups between the two species and supported the hypothesis that
meadow fescue was one of the diploid progenitors of hexaploid tall fescue (Chen et al., 1998). A more thoroughly developed map of meadow fescue was developed by Alm et al. (2003) utilizing RFLPs, AFLPs, isozymes and SSR markers. The integrated map (male and female parents) in the work of Alm et al. (2003) covered 658.8 cM with 466 markers and had an average marker density of 1.4 cM per marker. An interesting finding in this work, based on comparative mapping using heterologous RFLPs, was that chromosome 4F of meadow fescue was completely orthologous to rice chromosome 3 (Alm et al., 2003). This is in contrast to the Tritaceae where rice chromosome 3 is distributed over homoeologous group 4 and 5 chromosomes, indicating that the meadow fescue genome had a more ancestral configuration than any of the Tritaceae genomes (Alm et al., 2003). Later work with the meadow fescue mapping population created by Alm et al. (2003) identified QTL for vernalization requirement, heading time, and number of panicles (Ergon et al., 2006).

A genetic linkage map has recently been constructed for creeping bentgrass, comprised of 169 RAPD, 180 AFLP, 39 heterologous RFLP markers from cereal crops and 36 homologous RFLP markers from bentgrass cDNA (Chakraborty et al., 2005). The map covered 1,110 cM across 14 linkage groups, representing the expected 7 pairs of homeologous chromosomes (Chakraborty et al., 2005). The mapping population was derived from a cross between two highly heterozygous vegetative clones, exhibiting dramatic differences in leaf color, shoot density, root depth, and disease response to snow mold (*Typhula* spp.) and dollar spot (*Sclerotinia homoeocarpa* F. T. Bennett) (Chakraborty et al., 2005), and was later used to identify a single major QTL for dollar spot resistance on linkage group 7.1 (Chakraborty et al., 2006).
Additional work in bentgrasses includes the creation of a framework linkage map of colonial bentgrass (Rotter, 2009). The map was created by utilizing a hybrid backcross population of creeping x colonial bentgrass, and mapping 212 AFLPs and 110 gene-based markers utilizing a dideoxy polymorphism scanning technique (marker development based on SNP detection utilizing manual cycle sequencing) (Rotter, 2009). The colonial bentgrass map generated covered 1157 cM, and was utilized to identify regions of the colonial bentgrass genome potentially involved in dollar spot resistance (Rotter, 2009).

A few reports of linkage mapping in bluegrass species are available in the literature. Porceddu et al. (2002) developed a framework linkage map of Kentucky bluegrass by crossing an apomictic and sexual genotype. The maps consisted of 70 AFLP markers and 161 SAMPL markers, and covered 367 cM on the paternal map and 338.4 cM on the maternal map. Later work by this group added a few additional SSR and microsatellite-AFLP (AFLP derived marker system, in which markers are anchored to the 5’-end of microsatellite loci) markers (Albertini et al., 2003). Renganayaki et al. (2005) developed AFLP linkage maps of maternal and paternal plants of dioecious Texas bluegrass. The maternal map contained 126 single-dose AFLP markers, 31 linkage groups, and covered 1744 cM with an average marker spacing of 13.8 cM (Renganayaki et al., 2005). The paternal map contained 210 single-dose AFLP markers, 46 linkage groups, and covered 2699 cM with an average marker spacing of 12.9 cM (Renganayaki et al., 2005). The mapping population in this study was also used to identify QTL associated with the dioecy locus on the paternal map (Renganayaki et al., 2005).
Genetic linkage mapping in warm-season turfgrasses has been conducted on three genera/species. The earliest maps of a warm-season grass were developed by two separate research groups (Ebina et al., 1999; Yaneshita et al., 1999) utilizing an interspecific cross between *Zoysia japonica* and *Zoysia matrella*. The map by Yaneshita et al. (1999) was based on 115 genomic and cDNA RFLP markers, covering 22 linkage groups and a total map distance of 1506 cM. The map constructed by Ebina et al. (1999) utilized a combination of RFLPs and AFLPs to create a map with 23 linkage groups covering 1320 cM with an average marker distance of 8 cM. Later, Cai et al. (2004) developed an AFLP based linkage map in *Zoysia japonica*, based on 364 markers, grouped into 26 linkage groups covering 932.5 cM. This map was later enhanced by the addition of 161 genomic SSR markers (Cai et al., 2005).

Additional genetic mapping work in warm-season grasses includes the development of a framework map for a triploid interspecific hybrid between the tetraploid *Cynodon dactylon* and the diploid *Cynodon transvaalensis* (Bethel et al., 2006), a genetic linkage map of diploid *Paspalum notatum* (Ortiz et al., 2001), and a genetic linkage map of tetraploid *Paspalum notatum* (Stein et al., 2007). All of these maps were based on RFLPs. With the exception of the tetraploid *Paspalum notatum* map, which was used to identify QTL implicated in the trait for apospory, there have been no other reports of QTL analysis utilizing the maps generated in the warm-season grasses.
Kentucky Bluegrass

Kentucky bluegrass (Poa pratensis L.) is one of the most popular, and hence, widely propagated turfgrass species for amenity uses in the northern United States and Canada (Huff, 2003). According to Funk (2000), Kentucky bluegrass is lauded as the ‘king of grasses’ and the ‘final triumph of nature’, and is acclaimed as the premiere lawn grass, dominating or contributing to the turf of 40 million lawns throughout Canada and all but the warmest regions of the United States. Kentucky bluegrass is attractive, hardy, aggressive, and persistent, and its adaptation to a wide range of soils and climates makes it highly useful for soil preservation, forage, and turf (Bashaw and Funk, 1987; Funk, 2000). Of all the cool-season turfgrasses, Kentucky bluegrass has the most extensive rhizomes which contribute significantly to its ability to produce an excellent sod and survive stressful environments/conditions (Meyer and Funk, 1989). Its overall aesthetic appeal as a lawn grass sets the standard against which most other turfgrasses are compared due to its combination of softness, medium- to fine- leaf texture, high shoot density, dark green color, and persistence (Huff, 2003). Additionally, as a highly palatable and nutritious forage, Kentucky bluegrass is used for grazing on over 15 million hectares of pastures throughout the northeastern and north-central U.S. and substantial areas of Canada and Europe (Bashaw and Funk, 1987; Funk, 2000).

Kentucky bluegrass is native to the Old World and occurs naturally throughout the temperate and cooler regions of Europe and Asia (Bashaw and Funk, 1989). This fits with the center of origin of the genus Poa, which based on morphological, cytological, and species diversity is considered to be Eurasia, though many species have worldwide distribution (Huff, 2003). It is interesting to note that in modern times Kentucky
bluegrass has a worldwide, circumpolar distribution ranging from about 30°N lat to above 83°N lat (Huff, 2003), because there is some debate as to how this has occurred. Bashaw and Funk (1987) state that early colonists brought the majority of the North American Kentucky bluegrass ecotypes with them in seed mixtures, hay, and bedding. Once introduced, Kentucky bluegrass was rapidly disseminated by birds and other animals to millions of hectares without seeding by humans (Bashaw and Funk, 1987). Despite this, occasional references to Kentucky bluegrass being indigenous to the United States can be found in the literature – Gray (1908) pointed out that it is difficult to determine whether or not some strains found in northern and western North America are indigenous, while Soreng (1985) postulated that certain ecotypes of Kentucky bluegrass may be native to the Rocky Mountains, possibly originating through interspecific hybridization. Along these lines, Funk (2000) postulated that circumpolar Poas, including Kentucky bluegrass, may have migrated south into the U.S. via the Rocky Mountains. Evidence for this stems from the fact that Kentucky bluegrasses collected from old meadows throughout the coastal range in western Oregon and stream banks and rangelands in the inter-mountain west are frequently very different from bluegrasses collected in the eastern U.S. and Europe (Funk, 2000).

**Apomictic breeding behavior**

Kentucky bluegrass is a facultative apomict with complex cytological and embryological characteristics (Meyer and Funk, 1989). The species has a highly variable chromosome number, with the basic number being x = 7, and a complex series of polyploidy and aneuploidy ranging from 2n = 28 – 154 (Muntzing, 1933; Akerberg, 1939;
Tinney, 1940; Nielson, 1946; Grazi et al., 1961; Love and Love, 1975; Meyer and Funk, 1989; Huff, 2003). Additionally, Kentucky bluegrass has the ability to hybridize with and absorb entire genomes of other species of *Poa* (Clausen, 1961). This fact, coupled with the ability of apomictic reproduction to restore fertility to superior hybrid genotypes probably helps to explain the widespread distribution and great genetic diversity of this species (Dale et al., 1975; Bashaw and Funk, 1987).

Apomixis in plants is an asexual method of reproduction that mimics sexual reproduction in that seed develops in the ovule of the flower, but without the union of the sperm and egg (Bashaw and Funk, 1987). The embryo of the seed of an apomictic plant is formed from an unreduced and unfertilized nucleus, either from a somatic (vegetative) cell of the ovule (or integuments) or from a megaspore mother cell that fails to undergo meiosis (Bashaw and Funk, 1987). The progeny of an obligate apomictic plant receives its full chromosome compliment entirely from the maternal plant (Bashaw and Funk, 1987).

There are three mechanisms of apomixis that occur in higher plants: apospory, diplospory, and adventitious embryony. Adventitious embryony does not occur in grasses, but does occur in other plant families (Bashaw and Funk, 1987). Adventitious embryony is most commonly associated with citrus (Bashaw and Funk, 1987). This mechanism of apomixis is distinguished by the development of bud-like structures that differentiate into embryos, originating from somatic tissue of the ovule or integuments (Bashaw and Funk, 1987). In diplospory, the megaspore mother cell, which would normally undergo meiosis in a sexual plant, instead undergoes mitotic division to produce an unreduced embryo sac (Bashaw and Funk, 1987). In apospory, embryo sacs arise
from one or more of the somatic cells in the ovule (Bashaw and Funk, 1987). Facultative apomixis is possible with both diplospory and apospory. Facultative diplosporous plants are capable of either apomictic or sexual reproduction (Bashaw and Funk, 1987). In facultative aposporous plants, ovules are capable of both apomictic and sexual reproduction, in that a reduced megaspore mother cell or embryo sac may fail to abort and a normal sexual embryo sac is formed along with one or more aposporous embryo sacs that originate from the nucellus (Bashaw and Funk, 1987).

The type of apomixis in Kentucky bluegrass is somatic apospory, which is the development of embryo sacs from unreduced somatic cells (Bashaw and Funk, 1987; Funk 2000). The typical sequence of events in this type of apomixis involves the death of the reduced meiotic egg mother cell, or a situation where the reduced egg mother cell is outcompeted by an unreduced apospory embryo sac or by a rapidly developing apomictic proembryo (Huff, 2003). In the case of the unreduced apospory embryo sac, the embryo sac usually initiates from the nucellar region (Huff, 2003). As a facultative apomict, Kentucky bluegrass can form either sexual or aposporous embryo sacs (Bashaw and Funk, 1987). In fact, in many instances both sexual and aposporous embryo sacs can occur within the same ovule (polyembryony), and can give rise to multiple types of seedlings (Bashaw and Funk, 1987; Funk, 2000). Therefore, reproductive behavior in Kentucky bluegrass can range from completely sexual to nearly 100% apomictic (Funk, 1966; Bashaw and Funk, 1987).

The relative frequency of apomictic versus sexual reproduction can vary for a given genotype when grown in different locations or environments (Bashaw and Funk, 1987). Only a few of the causes for such variability in the expression of the apomictic
trait in Kentucky bluegrass are understood (Bashaw and Funk, 1987). Hovin et al. (1976) observed a greater frequency of sexually produced progeny of field grown plants in field locations promoting a longer flowering period, while Grazi et al. (1961) and Han (1969) observed a greater tendency toward sexuality when plants where grown under greenhouse conditions versus field conditions. Perhaps the best understood factor, from a mechanistic point of view, is that of pollination timing. Tinney (1940) as well as Akerberg and Bingefors (1953) observed that the egg cell of Kentucky bluegrass initiates division to form a proembryo before or shortly after the flowers open, which can effectively limit or prevent fertilization and sexual reproduction. Thus, pollination at or immediately following anthesis could increase the frequency of sexual reproduction (Bashaw and Funk, 1987).

In facultative apomicts, plants that differ genetically from their seed-bearing maternal parent are referred to as ‘off-types’ or ‘aberrants’ (Bashaw and Funk, 1987; Huff, 2003). Grazi et al. (1961) have described five general classes of progeny in Kentucky bluegrass (with classes (2) – (5) being considered aberrants): (1) apomorphic, where progeny are identical to the maternal parent; (2) haploid, from the pseudogamous development of a reduced (n) egg – results in progeny with only half the chromosomes of the maternal parent and none of the chromosomes of the paternal parent; (3) diploid, from the fertilization of a reduced (n) egg with a reduced (n) male gamete (normal sexual reproduction) – results in progeny with half the chromosomes from each parent; (4) triploid, from fertilization of unreduced (2n) egg by reduced (n) male gametes or, rarely, vice versa – results in progeny with all the chromosomes of the maternal parent and half the chromosomes of the paternal parent; and (5) tetraploids, from the rare fusion of two
unreduced (2n) gametes – results in progeny with the full chromosome compliment from each parent. The four general classes of aberrant progeny have come to be referred to as B_I individuals (haploid or polyhaploid), B_{II} hybrids, B_{III} hybrids (maternal or paternal, although paternal is rare), and B_{IV} hybrids, respectively (Funk, 2000; Huff, 2003). It should be noted that due to variable parental ploidy levels and irregular meiosis, aberrant progeny would only approximate the expected haploid, diploid, triploid, and tetraploid chromosome number (Bashaw and Funk, 1987). In other words, there can be a high percentage of fertilized or unfertilized aneuploids, that are capable of surviving because of the apomictic breeding system (Grazi et al., 1961; Huff, 2003). Additionally, it is possible for aberrants to arise by other methods, including mutation and androgenesis (Bashaw and Funk, 1987).

Although the progeny classification system developed by Grazi et al. (1961) and adopted by others describes four general types of aberrants in Kentucky bluegrass, not all types would be expected to be a source of elite germplasm in a breeding program. Haploids, or B_I individuals, are generally small and weak, and would not be expected to survive long in nature (Bashaw and Funk, 1987; Funk, 2000), while B_{IV} hybrids would be rare due to the fact that diploid pollen is very rare in Kentucky bluegrass (Grazi et al., 1961; Funk and Pepin, 1971). As a result, Funk and Pepin (1971) stated that B_{II} and B_{III} aberrant progeny would be of most interest to a breeding program, however, early experimental evidence suggested that the majority of hybrids are B_{III} hybrids (Pepin and Funk, 1971; Funk et al., 1973; Hintzen, 1979). Although it is not completely known why B_{III} hybrids are the most prevalent, Huff (2003) speculated that this may reflect the higher frequency of this type of hybridization event (particularly in highly apomictic lines where
maternal diploid eggs are common), or it may reflect the preferential selection of this
type of hybrid due to gigas characteristics (associated with increased chromosome
numbers/higher ploidy level) that confer greater ‘fitness’.

As stated previously, plant breeding consists of the creation of genetic variability,
selection of elite types from that variability, and synthesis of a stable cultivar from those
elite selections (Dudley, 2002). The apomictic breeding behavior of Kentucky bluegrass
makes realizing these goals of plant breeding both a challenge and a tremendous
opportunity. On the one hand, apomixis is an excellent means of maintaining the genetic
purity of a superior cultivar from one generation to the next (Meyer, 1982). Apomixis, as
it exists in Kentucky bluegrass, combines the genetic advantages of asexual reproduction
with the agronomic advantages of seed production (Meyer and Funk, 1989). Additional
benefits of apomixis include the ability to fix and maintain stable hybrid vigor, as well as
the ability to maintain high levels uniformity and stability required for certified seed
production and commercialization (Huff, 2003). While these benefits offer a great
opportunity to develop and release improved germplasm, the challenges of breeding
apomictic Kentucky bluegrass are significant. First and foremost, hybridization, or the
ability to create genetic variability, is possible only to the extent that sexuality is present
(Huff, 2003). Furthermore, in order to maintain high levels of apomixis in the hybrid
progeny that will later be used for commercialization, the parent plants of Kentucky
bluegrass should be highly apomictic (Huff, 2003). Bashaw and Funk (1987) stated that
most breeders would prefer an elite, highly apomictic plant as the male parent, while the
frequency of hybrids would be increased with increasing levels of sexual reproduction in
the female parent. However, they also stated that the probability of obtaining highly
apomictic segregates [aberrant progeny that retain the apomictic character] increases with higher levels of apomixis in both parents (Bashaw and Funk, 1987). This creates an obvious dilemma in that a balance must be obtained between the need to use sexual parents to create variability, and the need to use apomictic parents to fix the apomictic trait in the hybrid progeny. The practical result of this dilemma is that breeding progress can be slower and less predictable than in non-apomictic species (Funk, 2000). An excellent illustration of this fact is provided by Funk (2000) when he commented on the difficulties of making rapid improvements in apomictic Kentucky bluegrass by pointing out that ‘Midnight’ and ‘America’, developed in the 1970s, still remain among the top performers in the latest trials of the National Turfgrass Evaluation Program (NTEP).

Funk (2000) also stated that nearly every characteristic needed to develop an ideal lawn grass is present in the Kentucky bluegrass germplasm available to turfgrass breeders, however, available breeding techniques are insufficient to recombine all of the characteristics into a single cultivar. The practical application of these principles in a Kentucky bluegrass breeding program necessitates the creation and evaluation of large numbers of progeny in order to identify the striking of the balance between apomixis and sexual recombination. Funk (2000) likened this process as akin to ‘playing the lottery’.

A major first-step breakthrough in finding the balance between apomixis and sexual recombination occurred in the turfgrass breeding program at Rutgers University in the 1960’s. Funk and Han (1967) reported an increase in the frequency of F1 hybrids (aberrant progeny) of Kentucky bluegrass by applying foreign pollen between 1:00 and 4:00 A.M. to plants flowering in the greenhouse. This early morning time period coincides with anthesis, and ensures that pollen is applied at or shortly after the female
flower opens and the stigmas emerge (Funk and Han, 1967). This time period is critical because the egg of apomictic Kentucky bluegrass normally initiates division to form a proembryo before, at or very soon after anthesis (Tinney, 1940; Akerberg and Bingefors, 1953). Thus, the frequency of hybrids was increased by applying pollen as early as possible.

The above work was greatly expanded upon by Pepin and Funk (1970). These authors began by outlining three criteria for a successful hybridization program in Kentucky bluegrass: (1) crossing must be relatively easy and hybrids must be produced and identified in adequate numbers; (2) some hybrids must recombine desirable characteristics from both parents and also possess adequate vegetative and reproductive vigor; and (3) apomixis must be recovered in some of the hybrids (Pepin and Funk, 1970). Additional observations included that anthesis for Kentucky bluegrass occurred primarily between midnight and 5:00 AM, and was accelerated by a drop in temperature and an increase in humidity (Pepin and Funk, 1970). Refinement of greenhouse environmental conditions included subjecting plants to 17 hour day-lengths at 16°C to 28°C, and humid nights at 16°C, which was found to favor a high degree of sexual reproduction (Pepin and Funk, 1970). Plants crossed under these conditions produced 807 aberrant progeny out of 2113 total progeny (Pepin and Funk, 1970). Two additional significant contributions of this work included observations made in regards to the importance of the correct choice of parents and the predominant type of hybrid recovered from these crossing procedures. First, it was observed that using partially sexual plants as females was important, as the experimental selection Belturf (apomictic) produced only 7% aberrants as a female while the experimental selections Bellevue and P-102
(both only partially apomictic) averaged 42% and 25% aberrants as females, respectively (Pepin and Funk, 1970). Secondly, as previously stated, this work showed that an unexpectedly high percentage of the progeny were BIII hybrids (Pepin and Funk, 1970), which were likely derived from fertilization of an unreduced egg by reduced male gametes (Pepin and Funk, 1970; Bashaw and Funk, 1987).

Cultivars and cultivar classification systems

Bashaw and Funk (1987) developed an early broad classification system for Kentucky bluegrass cultivars that was based on (1) their type of use or (2) the breeding method used in their development. The first classification, usage type, was subdivided into two categories of (a) forage and conservation common-types and (b) specialized turf types (Bashaw and Funk, 1987). The forage and conservation common-types were originally defined as the erect, narrow-leafed bluegrasses typical of plants predominating in most old pastures throughout temperate regions of the world, and having characteristics that include early maturity, good stress tolerance, and the ability to survive dry summers in a dormant condition (Bashaw and Funk, 1987). In contrast to the forage and conservation common-types, the specialized turf type bluegrasses were defined as exhibiting a lower growth profile, being more responsive to higher maintenance, having greater resistance to leafspot, and being more tolerant of close mowing (Bashaw and Funk, 1987). Additionally, at the time of this publication specific cultivars of the specialized turf type bluegrasses had been released that exhibited improved shade tolerance, increased resistance to specific diseases, improved wear tolerance, a more
attractive appearance, and improved adaptation to specific regions and uses (Bashaw and Funk, 1987).

The second broad classification for Kentucky bluegrass cultivars described by Bashaw and Funk (1987), breeding method used in development, was subdivided into four categories of (a) selection of naturalized ecotypes; (b) selection of single, highly apomictic plants from old pastures or turfs; (c) blending of single plant progenies; and (d) selection of single, highly apomictic plants from breeding programs using intraspecific or interspecific hybridization. The earliest cultivars of Kentucky bluegrass fell primarily in the first two subcategories or occasionally the third subcategory, while recent breeding efforts have focused on the last category. Meyer (1982) stated that all cultivars of Kentucky bluegrass in the United States prior to 1970 were derived from composites collected from naturalized stands located in the Midwest or from individual apomictic clones found in old turf areas, while Bashaw and Funk (1987) stated that before 1970, selection among and within natural ecotype populations was the only breeding method that had produced commercially successful cultivars. In fact, as late as 1987, ecotypic selection was still considered to be the most efficient method of cultivar development in Kentucky bluegrass by Bashaw and Funk (Bashaw and Funk, 1987). Intraspecific hybridization as a method of breeding Kentucky bluegrass for turf was not proven to be a reliable method of cultivar development until 1970 (Pepin and Funk, 1970).

The earliest Kentucky bluegrasses available in the United States before the early 1950s were common types that represented naturalized ecotypes or single plants selected from old pastures or turfs (Bashaw and Funk, 1987). Examples of early common-type cultivars developed from selection of naturalized ecotypes or single plants from old
pastures included ‘Arboretum’, ‘Kenblue’ and ‘South Dakota Certified’ (Bashaw and Funk, 1987; Funk, 2000). Examples of early turf-type cultivars developed from selection of naturalized ecotypes or single highly apomictic plants from old pastures or old turfs included ‘Baron’, ‘Merion’, ‘Glade’, ‘Parade’, and ‘Sydsport’ (Bashaw and Funk, 1987; Funk, 2000). The cultivar ‘Merion’ was considered the ‘Landmark’ cultivar in Kentucky bluegrass. Merion was the first specialized turf type Kentucky bluegrass, exhibiting a true turf-type low growth habit and improved resistance to leaf spot disease (Drechslera poae [Baudys.] Shoem. or Bipolaris sorokiniana [Sacc.] Shoem.) (Meyer, 1982; Huff, 2003). This was in contrast to all other cultivars available until the late 1960’s that were highly susceptible to leaf spot when maintained at close mowing and high fertility (Meyer, 1982). Merion, a highly apomictic cultivar, discovered by golf course superintendent Joseph Valentine in 1936, originated as the progeny of a single plant growing on the Merion Golf Club in Ardmore, Pennsylvania (Meyer, 1982; Meyer and Funk, 1989). After thorough evaluation by agronomists at the USDA, USGA Green Section, and the Pennsylvania Agricultural Experiment Station, Merion was commercially released in 1947 (Meyer and Funk, 1989). Examples of early cultivars developed by blending progenies of two or more highly apomictic plants include ‘Park’ and ‘Voyager’ (Funk, 2000). Early cultivars developed through the groundbreaking intraspecific-hybridization work at Rutgers University include ‘Adelphi’, ‘America’, ‘Bristol’, ‘Bonnieblue’, ‘Eclipse’, ‘Galaxy’, ‘Majestic’, and ‘Midnight’ (Meyer, 1982; Bashaw and Funk, 1987; Funk, 2000).

Since the release of Merion, continued selection of natural ecotypes of Kentucky bluegrass, as well as intraspecific hybridization procedures have produced numerous new
The most recent NTEP Kentucky bluegrass trial (2005) lists 110 cultivars and experimental selections as entries. Additionally, university breeding programs and private seed companies have or continue to develop dozens if not hundreds more cultivars/experimental selections. Although many cultivars exhibit unique characteristics, the apomictic breeding behavior and the fact that ecotypic selection has played a significant role in past cultivar development has led to several types of cultivars having been introduced which can be grouped together based on similar agronomic and/or performance characteristics. The first description of a cultivar classification scheme was reported by Murphy et al. (1993) and was developed in order to “…increase the awareness about cultivar diversity in Kentucky bluegrass and provide a more user-friendly presentation of cultivar adaptation and use…”.

This initial classification system grouped cultivars according to their common features, and was primarily based on turf performance characteristics (Murphy et al., 1993). The following section provides an overview of the classification system as first described by Murphy et al. (1993). The classification system has been updated numerous times to include additional agronomic/performance characteristics or the inclusion of newly developed/released cultivars. Subsequent publications will be reviewed from the standpoint of how new data has been used to update the original classification types.

**Kentucky bluegrass cultivar classification types** (Murphy et al., 1993)

**Common Type or Midwest Ecotype Cultivars:**

This group of grasses was characterized as having an erect growth habit and narrow leaf blades and is frequently referred to as ‘common type’ Kentucky
bluegrass. These grasses are widely used for pastures, conservation purposes, and lower maintenance turf. They are commonly selected from naturalized ecotypes surviving within old pastures located in the Midwestern United States.

Common type cultivars were reported as generally producing seed early and exhibiting high seed yields. In general these cultivars were shown to exhibit good stress tolerance, in that they survive dry summers in a dormant state. As a turf, the major weakness of these cultivars was a tendency to be highly susceptible to leaf spot, however, many showed good to excellent resistance to dollar spot. Winter turf quality was reported as generally poor due to excessive purple coloration and damage from leaf spot. The most appropriate use of these grasses is as a utility turf grown under high mowing height and low maintenance where aesthetic quality is of little importance. Examples of cultivars in this category included ‘Arboretum’, ‘Kenblue’, ‘South Dakota Certified’, and ‘Ginger’.

**Aggressive Types:**

This category of cultivars was grouped primarily based on exhibiting aggressive lateral growth resulting in a high density turf. These cultivars were reported to dominate a turf stand when used in blends with other bluegrass types or when used in mixtures with other turf species. The aggressive growth habit was also shown to lead to excessive thatch production under certain environmental and management conditions. Turf performance was reported as variable from year to year. Other characteristics such as disease susceptibility and
spring greenup varied according to cultivar. Examples of cultivars in this category included ‘Princeton 104’, ‘A-34’, ‘Limousine’, and ‘Touchdown’.

**BVMG Types (Baron, Victa, Merit, Gnome Types):**

This group was classified as a widely used group of cultivars which forms a medium-low growing turf of medium density and medium-wide leaves. These cultivars generally showed good turf quality and high seed yield potential. They exhibited moderately good resistance to leaf spot, moderate resistance to dollar spot, stem rust, and leaf rust, but were reported to have high susceptibility to certain races of stripe smut. Another common characteristic included a stemmy appearance as a turf during the reproductive period of growth. Examples of cultivars in this category included ‘Baron’, ‘Victa’, ‘Merit’, and ‘Gnome’.

**Compact Types:**

These cultivars were characterized as exhibiting low compact growth habits, with many forming highly attractive turfs having good to excellent resistance to leaf spot. Other common characteristics included a slight tendency towards purple coloring during the winter and later spring greenup. Dollar spot and stem rust resistance as well as summer stress tolerance were reported as being cultivar dependant. Examples of cultivars in this category included ‘Blacksburg’, ‘Midnight’, ‘Unique’, ‘America’, ‘Able 1’, and ‘Nugget’.

**Mid-Atlantic Ecotypes:**
This group of cultivars was reported to form a vigorous turf of medium-high density with a deep and extensive rhizome system. Other common attributes included tolerance or good recovery from billbug damage, as well as excellent summer stress tolerance. Leaf spot resistance was reported as generally moderate, although some cultivars showed good resistance. Spring greenup and winter purple color varied considerably among the cultivars. Examples of cultivars in this category included ‘Preakness’, ‘Livingston’, ‘Baltimore City’, ‘Riverside Park’, Valsburg Park’, ‘Eagleton’, and ‘Wabash’.

**Bellevue Types:**

This group of cultivars was reported to form a turf of medium-low growth, medium-wide leaves, and medium density. As a group, they exhibited early spring greenup and excellent winter turf quality, showing little, if any, purple color. These cultivars showed good leaf spot, stem rust, and dollar spot resistance, and moderate resistance to summer patch and stripe smut. Tolerance to billbugs and summer stress was reported as being only moderate. The appearance of reproductive stems in late spring was reported as common with these cultivars. Examples of cultivars in this category included ‘Dawn’, ‘Classic’, ‘Banff’, and ‘Georgetown’.

**Other turf types:**

A large group of cultivars were reported as falling under the category of ‘other turf types’. This group was defined as a preliminary grouping of cultivars
requiring further study in order to accurately classify cultivars or assign to different groups.

The first changes to the classification system appeared in 1995 (Murphy et al., 1995). The changes included defining new types based on unique characteristics or defining sub-types within an already established category. The classification system was still based on growth and performance characteristics as evaluated in replicated turf trials. The following represents changes to the original classification scheme (Murphy et al., 1995):

**CELA type (Challenger, Eclipse, Liberty, Adelphi types):**

This new group was defined as turf-type cultivars that exhibited early spring green-up, moderate to good stripe smut resistance, good resistance to leaf spot, and variable winter color and dormancy.

**Cheri type:**

This new group was defined as cultivars that possess growth and performance characteristics similar to the cultivar ‘Cheri’. These cultivars were described as producing a turf of medium-low growth and medium density, with medium-wide leaves. These cultivars exhibited moderate resistance to leaf spot and dollar spot and possessed good resistance to stripe smut. Additional defining characteristics included relatively high seed yield and moderate winter dormancy.
**America type:**

This new group was originally defined as a sub-type within the compact type, and consisted of a number of cultivars that exhibited similar growth and performance characteristics to the cultivar ‘America’. Additional common characteristics included moderate winter dormancy with some purpling, moderate recovery from summer stress, and good resistance to dollar spot, leaf spot, and stripe smut diseases.

**Julia type:**

This new group was defined as cultivars that possess growth and performance characteristics similar to the cultivar ‘Julia’. These cultivars were found to form a turf of medium to dark green color with medium texture, and exhibited medium to good winter color and spring green up. Additional common characteristics included reports of good wear tolerance in European turf trials, good resistance to leaf spot disease, but poor resistance to dollar spot.

**Wabash type:**

This new group was originally defined as a sub-type within the Mid-Atlantic ecotype, and consisted of a number of cultivars that exhibited similar growth and performance characteristics to the cultivar ‘Wabash’. This sub-type was further defined as having a medium-green color, excellent heat tolerance, and above average resistance to billbugs. An additional defining characteristic
included moderate susceptibility to leaf spot, with the qualification that cultivars within this type exhibited good recovery from this disease.

Two additional groups appeared in a 1998 version of the classification system, while the Wabash type was removed. The additions/subtractions were still based on growth and performance characteristics as evaluated in replicated turf trials. The following additions to the original classification system are from Bonos et al. (1998):

**Midnight type:**

This new group was originally defined as a sub-type within the compact type, and consisted of a number of cultivars that exhibited similar growth and performance characteristics to the cultivar ‘Midnight’. This sub-type was further defined as cultivars having characteristically long winter dormancy with late spring green-up. Additional defining characteristics included dark green color, good turf quality, good heat tolerance, as well as susceptibility to powdery mildew.

**Shamrock type:**

This new group was initially defined as consisting of a number of cultivars that exhibited similar growth and performance characteristics to the cultivar ‘Shamrock’. This group was further defined as cultivars exhibiting moderate resistance to leaf spot, moderate susceptibility to billbug feeding, and having
moderate winter color. An additional defining characteristic included the potential for high seed yield production.

The development of the classification system took a new turn in 2000, in that for the first time measurements other than growth and performance characteristics evaluated in replicated turf trials were utilized. These measurements included typical parameters used during the PVP application process including plant height, panicle height, flag leaf height and length, subtending leaf length and width, rhizome spread, and longest extending rhizome (Bonos et al., 2000). The measurements were taken on 45 cultivars and selections, representing the major types of Kentucky bluegrass (as defined originally by Murphy et al., 1993), grown in a spaced-plant nursery trial. Principle component analysis of the above mentioned morphological measurements made on spaced-plants supported the classification types of the Common, Compact, Bellevue, Mid-Atlantic, and BVMG, but not necessarily the Aggressive type (Bonos et al., 2000). This discrepancy compared with groupings developed based on evaluation of replicated turf trials was attributed to distinct differences in other traits, not studied in this experiment, that exist between the major Kentucky bluegrass types – specifically, the Aggressive type cultivars did not form a distinct group based on growth characteristics measured in this study, but do share characteristics of high shoot density and aggressive lateral spread in replicated turf trials (Bonos et al., 2000).

Updates of the classification system appeared in the Rutgers Turfgrass Proceedings in 2003 and 2004 (Bonos et al, 2003; Shortell et al., 2004). These updates were based on growth and performance characteristics evaluated in replicated turf trials.
In addition to continuing to update and place newly released cultivars and selections into classification groups a few additional changes were made:

1. Compact-Midnight type defined as a subtype of the compact type
2. Compact-America type defined as a subtype of the compact type
3. Aggressive type also referred to as High-Density type
4. Texas x Kentucky bluegrass intraspecific hybrids added as a new type
5. Cheri type removed from the classification scheme

The most recent work on the classification system involved a molecular analysis of the grouping scheme utilizing RAPDs (Curley and Jung, 2004) and a second round of morphological measurements on spaced-plant nursery trials (Shortell et al., 2009). In the work by Curley and Jung, the authors used RAPD marker data collected from three seedling replicates each of 123 cultivars, plant introductions (PIs), experimental breeding lines, and interspecific hybrids between Kentucky bluegrass and Texas bluegrass. The RAPD data supported a correlation between genetic relatedness and morphological groupings only for the BVMG and Compact-Midnight types, and a weak correlation for a few of the Common type cultivars investigated in the study, while showing no correlation with any of the remaining morphological/trait based groupings. This is interesting in that the basis for the classification system has been based in part on pedigree (breeding history), and the PVP morphological analysis conducted by Bonos et al. (2000) should reflect genetic relatedness. There are a number of likely possibilities that could explain the discrepancy: 1) Performance characteristics and morphological measurements are not
accounting for all of the genetic variability at the DNA level captured in the RAPD analysis; 2) the RAPD analysis is not accurately reflecting the genetic relatedness of plant material within a classification group; or 3) the seedlings chosen for study in the RAPD analysis did not accurately reflect the apomictic character typical of Kentucky bluegrass cultivars. In regards to the first possibility, the classification scheme was originally developed based in part on turfgrass performance in replicated turf trials, and the performance of a cultivar is not necessarily an indication of genetic relatedness. That being said, the PVP morphological analysis work conducted by Bonos et al. (2000) supported the grouping system with the exception of support for the Aggressive type. In regards to the second possibility, the drawbacks of RAPDs as a marker system have already been discussed in detail. Of particular importance is the lack of repeatability of these markers within or between laboratories. As an example, the within-cultivar/seed source variability between the three seedling replicates in the study under discussion ranged from 0.05 to 0.50 with a mean value of 0.22 for all 123 genotypes. The authors speculated that this variability represented ‘off-types’ from seed, however, part of this variability could simply be due to the marker system chosen for the analysis. The final possibility relates to whether or not the seedling plant material accurately reflected the apomictic breeding behavior of Kentucky bluegrass. The work of Bonos et al. (2000), as well as the typical PVP work on Kentucky bluegrass, identifies the repeating apomictic maternal type plants prior to taking morphological measurements, whereas the RAPD work by Curley and Jung utilized plant material grown-out from 3 random seeds. Additional work will be needed to resolve the discrepancy described above.
The most recent morphological characterization of the classification system was conducted Shortell et al. (2009) utilizing the previously described PVP measurements on all 173 cultivars and selections in the 2000 NTEP Kentucky bluegrass test. As before, these measurement were conducted on plants grown in a spaced-plant nursery (Shortell et al., 2009). This work found significant differences in morphological and agronomic traits between the classification types of Kentucky bluegrass, which was consistent with previous research (Shortell et al., 2009). Additional support was found for grouping the Compact-Midnight and Compact-America types as subtypes within the Compact type (Shortell et al., 2009). This data supports the following scheme for the Kentucky bluegrass cultivar classification system:

- Compact Type
  - Compact-Midnight Type
  - Compact-America Type
- Julia Type
- Mid-Atlantic Ecotype
- Shamrock Type
- Aggressive or High Density Type
- CELA Type
- Bellevue Type
- BVMG Type
- Common Type
- Texas x Kentucky bluegrass Hybrids
Creeping Bentgrass

The genus *Agrostis*, or bentgrass, is a large genus of cool-season grasses (Poaceae) consisting of over 200 species. The center of origin for the genus is considered to be Western Europe (Harlan, 1992). The word *Agrostis*, derived from Greek, translates as ‘some kind of grass’ (Stubbendieck and Jones 1996). Taxonomically (based on morphological characteristics), the genus *Agrostis* belongs to the subfamily Pooideae, in the tribe Aveneae, and is therefore closely related to the genus *Avena* or oats (Kellogg 1998). Currently, five species of *Agrostis* are used as turfgrasses in the United States: colonial bentgrass (*Agrostis capillaris* L.), velvet bentgrass (*Agrostis canina* L.), dryland bentgrass (*Agrostis castellana* Boiss. and Reut.), redtop bentgrass (*Agrostis gigantea* Roth) and creeping bentgrass (*Agrostis stolonifera* L.) (Brilman, 2001; Warnke, 2003). These species are perennial, outcrossing grasses and used for lawns, athletic fields, and golf courses.

Historically the taxonomic relationships of the five species used as turfgrasses have been somewhat confusing. Vergara and Bughrara (2003) stated that this has occurred as a result of too few hybridization studies, difficulties in obtaining old botanical records, and inconsistencies in morphological characterizations, i.e., synonyms, subspecies, and re-classification of some species. Additional points of confusion likely have arisen as a result of new phenotypes with different number of chromosomes and/or distinct characteristics having been given new species names, or uncertainty of
classification of spontaneous natural hybrids where lineage and ploidy levels are not known (Vergara and Bughara, 2003). A few examples of classification problems include:

1) Inclusion of the cultivars ‘Highland’ and ‘Exeter’ in the species *Agrostis capillaris* in the US National Plant Germplasm System (Ruemmele, 2003), despite many sources recognizing these cultivars as belonging to the separate species of *A. castellana* (Hubbard, 1984; Hartmann et al., 1988; Brilman, 2001)

2) Creeping bentgrass being variously referred to as *Agrostis palustris* Huds., *Agrostis stolonifera* L., *Agrostis stolonifera* L. var. *palustris* (Huds.) Farw. in the literature (Warnke, 2003). Hitchcock (1951) considered *A. stolonifera* and *A. palustris* to be distinct species. Today, it has come to be accepted that *A. stolonifera* and *A. palustris* are synonyms for the same species, with *A. stolonifera* being more universally accepted as the correct nomenclature for creeping bentgrass (Warnke, 2003)

3) Velvet bentgrass, *A. canina*, was originally divided into two subspecies: ssp. *canina* or velvet bentgrass, and ssp. *montana* or brown bentgrass (Brilman, 2003). It was later determined that brown bentgrass was likely an autotetraploid of ssp. *canina*, and was therefore separated and classified as the species *A. vinealis* (Brilman, 2003).

The examples of classification problems listed above, as well as other inconsistencies in the literature, likely stem from the wide range of morphological variation found within
and among species of this genera. Brede and Sellman (2003) stated that many of today’s grass breeders are aware that breeding populations of a single *Agrostis* species can sometimes vary in nearly all morphological characters. It is likely that confusion will occasionally continue until an extensive morphological and molecular genetic investigation of both turf and non-turf species of *Agrostis* is conducted.

The most significant cytogenetic work in the *Agrostis* species used for turf was conducted by Jones (1956 a, b, c), where part of the work included evaluation of chromosome pairing behavior in hybrids between multiple species. Based on this work, the tetraploid turf species, *A. vinealis, A. capillaris,* and *A. stolonifera* were given the genome designation $A_1A_1A_1A_1$, $A_1A_1A_2A_2$, and $A_2A_2A_3A_3$, respectively. The diploid *A. canina* was determined to be the likely progenitor of the A1 genomes, while the A2 genomes originated with an unknown rhizomatous diploid progenitor and the A3 genomes originated with an unknown stoloniferous diploid progenitor (Jones, 1956c; Wipff and Fricker, 2001; Warnke, 2003). Taken as a whole, the work of Jones (1956 a, b, c) points to creeping bentgrass being an allotetraploid with two unknown diploid progenitors.

Today, creeping bentgrass is considered to be the most adapted species for use on golf course tees, fairways and putting greens because of its fine texture and adaptation to mowing heights as low as 3mm (Warnke, 2003). This is interesting in that this species did not comprise a high percentage of the grass species mixtures that were originally used to seed some of the earliest golf courses in the United States. Golf courses that were established in the late 19th century or early 20th century were planted with seed mixtures termed ‘South German Mixed Bentgrass’ (Duich, 1985; Warnke, 2003). The seeds of
South German Mixed Bentgrasses were collected from semi-wild pasture populations in Austria, Hungary and other areas of central Europe (Duich, 1985; Warnke, 2003). A significant bentgrass seed trade from Europe continued for many decades after the late 1800’s, with production becoming more concentrated from semi-wild pastures across numerous locations in Germany (Edler, 1930). Although percentages of different Agrostis species varied from seed lot to seed lot, it appears that South German Mixed Bentgrass generally contained up to 75% A. capillaris, 15-30% A. canina, up to 10% A. gigantea, and a trace to up to 5% A. stolonifera [approximate averages based on information in numerous publications] (Hillman, 1921; Piper and Oakley, 1922a; USGA, 1922; Oakley, 1923a; Henry, 1928; Hillman, 1930; Montieth, 1930; USGA, 1930). Thus, creeping bentgrass was originally the minority of this mixture, but slowly emerged to be the dominant species comprising the turf in most regions of the US, with the exception of the cool maritime Northeast where A. canina sometimes out-competed the other bentgrasses in the original seeding mixtures (Piper and Oakley, 1922a, b; Oakley, 1923a; USGA, 1923; Montieth, 1930; USGA 1930; Hurley, 1973; Skogley, 1973; Brilman and Meyer, 2000; Ruemmele, 2000; Rose-Fricker et al., 2004; Hollman et al., 2005).

Additional sources of bentgrass seed in the late 19th century and early 20th century were harvested from naturalized stands of various bentgrasses in New Zealand, Rhode Island (as well as other New England states), and the West Coast of the United States. In 1930, imports of bentgrass seed from New Zealand were greater than imports of South German Mixed Bentgrass by slightly over 200,000 lbs [294,000 lbs. from Germany versus 508,900 lbs. from New Zealand] (USGA, 1930). The majority of the seed from New Zealand at this time was apparently A. capillaris, with a small percentage of the
seed being *A. castellana* (Suckling and Forde, 1978; Karataglis, 1980; Karataglis, 1986; Rapson and Wilson, 1992a, b). Seed production in Rhode Island and neighboring states was originally production of seed from naturalized stands of *A. capillaris* (once referred to as Rhode Island bent, but later recognized as colonial bent), but was later shifted to production of *A. canina* and *A. stolonifera* as imports of *A. capillaris* increased in quantity from Germany and New Zealand (Odland, 1930; Hillman, 1930). Bentgrass seed production in the Pacific Northwest of the United States began in the early 1920’s, apparently after Lyman Carrier (Office of forage crops, Bureau of Plant Industry, USDA) encouraged the development of a seed producing industry after being informed about high quality naturalized stands of bentgrass growing in coastal regions of Oregon and Washington (Schoth, 1930). Seed production began in an area from southwestern Oregon to the Canadian border, mostly west of the Cascade Mountain range (Schoth, 1930). The largest bentgrass seed producing areas became Coquille valley and the lower Columbia river districts in Oregon, and the Puget Sound area in Washington (Schoth, 1930). Smaller areas of production included Reedsport and Gardiner near the mouth of the Umpqua river, Cushman near the mouth of the Siuslaw river, areas near Yoncalla, the Willamette valley, and in an area near Klamath Falls in Oregon, and a few locations in Washington near Chehalis and Raymond (Schoth, 1930). These regions produced either creeping bentgrass or colonial bentgrass (Schoth, 1930). Much of the creeping bentgrass from these regions collectively became known as ‘Seaside bentgrass’ (Schoth, 1930). A region near Coos Bay and areas of production near the lower Columbia River were recognized as being free from mixtures and of the purest quality stands (Schoth, 1930). Pure stands of creeping bentgrass from the Coos Bay region of Oregon became known as
‘Cocoos’ bentgrass, and were protected under a trademark from the US patent office and the Commissioner of Patents of Canada (Carrier, 1928).

The mixtures of bentgrasses seeded on early golf courses would eventually segregate under biotic and abiotic stresses, and higher quality patches would ‘stand-out’ as being better adapted to various management regimes or environmental conditions. Bentgrasses from these patches or superior plants from seeding mixtures were collected by personnel from the USDA and USGA, and evaluated for superior turf characteristics at the Arlington Turf Gardens as early as 1917 (Oakley, 1923b). The procedure was described by Oakley (1923b):

“In September, 1917, plots 8x8 feet in the grass garden at Arlington were prepared as for seeding, and runners of some of the best strains of both creeping bent and velvet bent were spread thinly but uniformly over the beds. They were then covered with about one-half inch of good compost, composed largely of topsoil, and rolled and watered carefully. The runners thus planted rooted quickly and sent up new plants from the joints… By September, 1918, the most rapidly spreading strains of creeping bent had covered a width of 4 to 6 feet; therefore, abundance of planting material was available for continuing the turf experiments started the fall previously”.

Another interesting aspect of this work was that evaluation of promising vegetative material began almost immediately on-site on various golf courses. Again, from Oakley (1923b):
“The most notable plantings of the fall of 1918 were made on the greens of the East Potomac Park Public Golf; in all, five greens were planted on this course between the dates of October 2 and 12. These were the first greens to be planted with bent runners and gave added evidence of the feasibility and advantages of the vegetative method. Plots of many different strains of both creeping bent and velvet bent were planted at Arlington, so that a good opportunity was afforded for studying the possibilities of the new way of getting putting green turf.

Additional nursery rows were planted at Arlington in September, 1918, and Dr. Walter S. Harban started a nursery at the Columbia Country Club later that fall. This was the first bent grass nursery to be established on a golf course. From this nursery the 9th green of the Columbia course was planted in the fall of 1919. This green attracted so much attention that it virtually “sold” the vegetative method to the golfing fraternity”.

This selection and evaluation work continued for a number of decades (Latham, 1958). Superior clonal plant material from this work became known as the ‘C-series’ of bentgrasses, and were used to vegetatively establish putting greens and other golf course turfs (Warnke, 2003). A description, of some of the C-series bentgrasses was provided by the USGA (1924) and Latham (1958). From the USGA (1924):
Washington Bent: A creeping bent selected in 1919 from No. 4 green at the Washington Golf and Country Club, near Rosslyn, Va. At the time the selection was made, the piece of turf selected was only slightly affected by brown-patch, although nearly the entire green was brown.

Metropolitan Bent: A creeping bent from a piece of turf sent in for identification in 1917 by a seed firm in New York City.

Virginia Bent: A creeping bent selected in 1919 from No. 18 green at the Washington Golf and Country Club, near Rosslyn, Va. The selection, at the time it was made, was found to be strongly resistant to brown-patch, but not immune.

Columbia Bent: A creeping bent selected in 1916 from Columbia Country Club, near Washington, D. C. This is the strain used to plant No. 9 green at the Columbia Country Club.

Vermont Bent: A creeping bent received in 1917, without any special data, from Middlebury, Vt. [from Ekwanok Country Club, Manchester, Vt. according to Latham (1958)].

Acme Bent: A velvet bent selected in 1919 from the grounds of the Department of Agriculture, Washington, D. C.

Revere Bent: A seaside bent from the wet salt meadows at Revere Beach, Mass., where it occurs in large areas.

A listing of the C-series bentgrasses, along with their origin and original collection date are presented in Table 1 (USGA, 1943).
Table 1.  C-series creeping bentgrasses (adapted from USGA, 1943)

<table>
<thead>
<tr>
<th>Strain</th>
<th>Found Originally</th>
<th>City</th>
<th>State</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>C 1</td>
<td>CC of Atlantic City</td>
<td>Atlantic City</td>
<td>NJ</td>
<td>1928</td>
</tr>
<tr>
<td>C 4</td>
<td>Arlington Turf Garden</td>
<td>Arlington</td>
<td>Va</td>
<td>1934</td>
</tr>
<tr>
<td>C 5</td>
<td>Arlington Turf Garden</td>
<td>Arlington</td>
<td>Va</td>
<td>1934</td>
</tr>
<tr>
<td>C 7</td>
<td>Pine Valley Golf Club</td>
<td>Clementon</td>
<td>NJ</td>
<td>1935</td>
</tr>
<tr>
<td>C 8</td>
<td>Baltimore CC</td>
<td>Baltimore</td>
<td>Md</td>
<td>1935</td>
</tr>
<tr>
<td>C 9</td>
<td>Washington Golf &amp; CC</td>
<td>Arlington</td>
<td>Va</td>
<td>1936</td>
</tr>
<tr>
<td>C 11</td>
<td>Washington Golf &amp; CC</td>
<td>Arlington</td>
<td>Va</td>
<td>1936</td>
</tr>
<tr>
<td>C 12</td>
<td>Los Angeles CC</td>
<td>Beverly Hills</td>
<td>Ca</td>
<td>1936</td>
</tr>
<tr>
<td>C 14</td>
<td>Toronto Golf Club</td>
<td>Long Branch</td>
<td>Ontario</td>
<td>1936</td>
</tr>
<tr>
<td>C 15</td>
<td>Toronto Golf Club</td>
<td>Long Branch</td>
<td>Ontario</td>
<td>1936</td>
</tr>
<tr>
<td>C 16</td>
<td>Rolling Green Golf Club</td>
<td>Springfield</td>
<td>Pa</td>
<td>1936</td>
</tr>
<tr>
<td>C 17</td>
<td>Manor CC</td>
<td>Norbeck</td>
<td>Md</td>
<td>1936</td>
</tr>
<tr>
<td>C 19</td>
<td>Congressional CC</td>
<td>Washington</td>
<td>D. C.</td>
<td>1936</td>
</tr>
<tr>
<td>C 27</td>
<td>Washington Golf &amp; CC</td>
<td>Arlington</td>
<td>Va</td>
<td>1937</td>
</tr>
<tr>
<td>C 28</td>
<td>Washington Golf &amp; CC</td>
<td>Arlington</td>
<td>Va</td>
<td>1937</td>
</tr>
<tr>
<td>C 32</td>
<td>Congressional CC</td>
<td>Washington</td>
<td>D. C.</td>
<td>1936</td>
</tr>
<tr>
<td>C 35</td>
<td>Manor CC</td>
<td>Norbeck</td>
<td>Md</td>
<td>1937</td>
</tr>
<tr>
<td>C 36</td>
<td>Manor CC</td>
<td>Norbeck</td>
<td>Md</td>
<td>1937</td>
</tr>
<tr>
<td>C 38</td>
<td>Arlington Turf Garden</td>
<td>Arlington</td>
<td>Va</td>
<td>1937</td>
</tr>
</tbody>
</table>

Commercial Strains:

- C 50 Washington bent
- C 51 Metropolitan bent
- C 52 Old Orchard bent
- C 60 Seaside bent
- C 61 Astoria bent
- C 65 Highland Colonial bent

Latham (1958) added that C1 was referred to as ‘Arlington bent’, C7 was referred to as ‘Cohansey bent’, and C15 was referred to as ‘Toronto bent’. Latham also notes that ‘Congressional bent’ came from Congressional CC in 1936 [possibly C19 or C32] and that ‘Collins bent’ came from Washington Golf & CC in 1937 [possibly C27 or C28].
Another named bentgrass, ‘Dahlgren bent’, was selected by Dr. Fred V. Grau on the Naval Proving Grounds, Dahlgren, Virginia (Latham, 1958).

The pinnacle of the evaluation work conducted on the C-series bentgrasses occurred with the planting of experimental greens on various golf courses across the US. The experimental putting green C-series evaluation research was described by Montieth (1939):

“During the past few years several strains of creeping bent have shown much promise in the Arlington tests, so it was decided to test them under actual putting green conditions. In the fall of 1937 through the co-operation of the Capital Golf and Country Club and T. S. Lumsden, Greenkeeper, the third green on that course was planted to different strains of creeping bent, some of them being popular and well known grasses, and others comprising selections which have stood up particularly well at the Arlington Turf Garden.

The green was divided into wedge-shaped sectors like a pie and each sector planted to a different strain. In this way it has been possible to try out various selections for durability under play under Washington conditions…

During the past spring similar experimental greens have been planted on golf courses in many districts in the United States and one in Canada. With one exception these plantings are on practice putting greens or on regular greens, so they will be exposed to the regular wear and tear of play. Planted under such widely different soil and climatic conditions, it will be possible to determine the
range of conditions which some of our more promising strains will tolerate under play”.

Locations of the experimental greens included the following (Montieth, 1939):

Denver Country Club, Denver, Colorado
Omaha Country Club, Omaha, Nebraska
Mohawk Park Municipal Golf Course, Tulsa Oklahoma
Swope Park Municipal Golf Course No. 2, Kansas City, Missouri
Algonquin Golf Club, Webster Grove, Missouri
Westwood Golf Club, Clayton, Missouri
Oak Park Country Club, Oak Park, Illinois
Olympia Fields Country Club, Olympia Fields, Illinois
Westmoreland Country Club, Wilmette, Illinois
Country Club of Detroit, Grosse Pointe, Detroit, Michigan
Meridian Hills Country Club, Indianapolis, Indiana
Canterbury Golf Club, Warrensville, Ohio
Kenwood Country Club, Madisonville, Ohio
East Lake Golf Course of Atlanta Athletic Club, Atlanta, Georgia
Capital Golf and Country Club, Glen Echo, Maryland
Allegheny Country Club, Sewickley, Pennsylvania
Baltusrol Golf Club, Springfield, New Jersey
Canoe Brook Country Club, Summit, New Jersey
Many of the vegetatively propagated C-series bentgrass cultivars were of very high quality (Warnke 2003), and were widely used to establish putting greens and other golf course turfs up until the 1970’s when a mysterious “decline and dying-out” began to affect Toronto C-15 bentgrass in the Midwestern United States (Couch, 1981; Roberts and Vargus, 1984). The cause of this disease was likely bacterial wilt caused by *Xanthomonas campestris* (Roberts and Vargus, 1984). The devastating affects of C-15 decline in the 1970s and 1980s signaled the beginning of the end of clonally established vegetative bentgrasses by illustrating vulnerabilities associated with planting genetically identical monocultures.

The first improved seeded cultivar of creeping bentgrass was ‘Penncross’, developed by Dr. H.B. Musser at Pennsylvania State University (Musser, 1959; Warnke, 2003). Penncross was released in 1955, and is considered to be the landmark cultivar of creeping bentgrass (Rogers, 1991; Warnke, 2003). Penncross is considered a first generation three clone synthetic, produced by maintaining and continually crossing three distinct vegetatively propagated clonal creeping bentgrass strains (Warnke, 2003).

Since the development of Penncross, plant breeding efforts have markedly improved creeping bentgrass varieties to withstand the increasing demands of the game
of golf, addressing the need for better turf quality, lower mowing tolerances, darker green color, improved shoot density, traffic tolerance, recuperative ability, and increased disease and environmental stress tolerance (Koch et al., 2007). Of particular importance was the need for improved tolerance to lower mowing heights as golfers began expecting tournament quality playing conditions at their local golf courses (Hurley and Murphy, 1996). Fast putting green speeds became the goal, and the easiest way to accomplish this was to lower cutting heights (Hurley and Murphy, 1996). As an example, it was common to cut putting greens at 1/4 inch in the 1920’s, 3/16 inch in the 1960’s, an 1/8 inch in the 1990’s, and an 1/8 inch or lower in the 2000’s (Hurley and Murphy, 1996). These lower cutting heights have served to increase susceptibility to traffic stress, increase disease pressure, increase weed invasion, and increase susceptibility to environmental stresses such as heat and drought. Turfgrass breeders have responded by continuously releasing seeded creeping bentgrass cultivars with incremental improvements in quality compared to Penncross. Examples of improved varieties that were released in the 1970’s through the 1990’s are listed in Table 2 (Hurley and Murphy, 1996).

Major improvents in quality occurred with the development and release of L-93 and the Penn A and G series in the mid- 1990’s. At the time of their release, the A and G series bentgrasses had markedly greater shoot densities than almost all other commercially available creeping bentgrass cultivars. Shoot densities of the A and G series bentgrasses ranged from the low- to mid- 2000’s to the mid- 3000’s shoots per square decimeter, whereas older bentgrass varieties generally had shoot densities below 2000 shoots per square decimeter (Beard and Sifers, 1997). L-93 also produced a finer, denser turf than most commercially available cultivars at the time, but was not quite as
dense as the A and G series creeping bentgrasses from Penn State (Hurley and Murphy, 1996). The greater density of these varieties allowed for lower cutting heights and improved putting quality compared to older cultivars (Hurley and Murphy, 1996).

Table 2. Examples of improved creeping bentgrass varieties developed in the 1970’s through 1990’s (adapted from Hurley and Murphy, 1996)

<table>
<thead>
<tr>
<th>Variety name</th>
<th>Year released</th>
<th>Developed by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penneagle</td>
<td>1979</td>
<td>Penn State University</td>
</tr>
<tr>
<td>Pennlinks</td>
<td>1986</td>
<td>Penn State University</td>
</tr>
<tr>
<td>Cobra</td>
<td>1987</td>
<td>Rutgers Univ. / International Seeds</td>
</tr>
<tr>
<td>SR 1020</td>
<td>1987</td>
<td>Univ. of Arizona / Seed Research Inc.</td>
</tr>
<tr>
<td>Providence</td>
<td>1988</td>
<td>Univ. of Rhode Island / Seed Research Inc.</td>
</tr>
<tr>
<td>Putter</td>
<td>1989</td>
<td>Washington State Univ.</td>
</tr>
<tr>
<td>Southshore</td>
<td>1991</td>
<td>Rutgers Univ. / Lofts Seed Inc.</td>
</tr>
<tr>
<td>Cato</td>
<td>1993</td>
<td>Texas A &amp; M Univ.</td>
</tr>
<tr>
<td>Crenshaw</td>
<td>1993</td>
<td>Texas A &amp; M Univ.</td>
</tr>
<tr>
<td>Penn A-1</td>
<td>1995</td>
<td>Penn State University</td>
</tr>
<tr>
<td>Penn A-4</td>
<td>1995</td>
<td>Penn State University</td>
</tr>
<tr>
<td>Penn G-2</td>
<td>1995</td>
<td>Penn State University</td>
</tr>
<tr>
<td>Penn G-6</td>
<td>1995</td>
<td>Penn State University</td>
</tr>
<tr>
<td>Lofts L-93</td>
<td>1995</td>
<td>Rutgers Univ. / Lofts Seed Inc.</td>
</tr>
</tbody>
</table>

Additional improved characteristics of L-93 and the A and G series bentgrasses, compared to most other commercially available creeping bentgrass varieties at the time, included overall improved turf quality (Morris, 2003), improved heat stress tolerance (Toubakaris and McCarty, 2000), improved traffic tolerance (Samaranayake et al., 2008),
and improved resistance to invasion of annual bluegrass (Samaranayake et al., 2008). L-93 has also displayed good resistance to dollar spot disease (Hurley and Murphy, 1996).

Although the above discussion points to the significant improvements made in overall turf quality of creeping bentgrass cultivars, one major obstacle that still remained at the time of the release of L-93 and the A and G series bentgrasses was dramatic improvements in resistance to dollar spot (Sclerotinia homoeocarpa F.T. Bennett) disease. Dollar spot is the most common and persistent disease of creeping bentgrass in the United States and throughout much of the world (Bonos et al., 2006). Cultivars of creeping bentgrass have been shown to differ in their response to this disease (Vincelli et al., 1997; Golembiewski and Danneberger, 1998; Abernathy et al., 2001; Morris, 2005).

In a 2003 study conducted at Rutgers University, only 12 of 265 creeping bentgrass clones collected from New Jersey, Illinois, and Arizona exhibited significant resistance to dollar spot (Bonos et al., 2003). Many of the newest generation of creeping bentgrass cultivars, released in the 2000’s, have been specifically bred for resistance to dollar spot. Newly released cultivars such as Declaration, Kingpin, Benchmark DSR, 007, and experimental selections 13M, HTM, and HTL, developed through an intensive dollar spot breeding program, typically sustain less disease than cultivars not intentionally selected for dollar spot resistance, such as Penneagle, Penn G-2, Penn A-2, Bengal, SR 1119, Backspin, Independence, and Eighteenth Green (Bonos, 2005; Weibel et al., 2005; Bonos et al., 2006).

All of the above mentioned improvements have come through traditional classical field breeding techniques, with little or no contribution from the study of molecular genetics in this important turfgrass species. To date, the lone exception has been the
development of a genetic linkage map in creeping bentgrass utilizing RAPDs, AFLPs, and RFLPs (Chakraborty et al., 2005). Although this work was a significant initial step in developing molecular genetic resources for creeping bentgrass, particularly by identifying QTL for dollar spot resistance, the molecular markers utilized in this study are of limited practical value for a turfgrass breeding program. The problems with RAPD markers have already been discussed (intra- and interlaboratory repeatability), while the laborious and demanding procedures associated with RFLPs are not conducive too rapidly or efficiently screening large populations of plants in breeding programs. A linkage map consisting of predominantly PCR based markers would prove more efficient and have more utility as a complement to traditional field based creeping bentgrass breeding programs.
Literature Cited


Lagercrantz, U. 1998. Comparative mapping between Arabidopsis thaliana and Brassica nigra indicates that brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics. 150:1217-1228.


from expressed sequence tag datasets derived from three cattle cDNA libraries. Journal of Biochemistry and Molecular Biology. 39:183-188.


by independent mutations at corresponding genetic loci. Science. 269:1714-1718.


Schneider, K., B. Weisshaar, D. C. Borchardt, and F. Salamini. 2001. SNP frequency and allelic haplotype structure of *Beta vulgaris* expressed genes. Molecular Breeding. 8:63-74.


Sokolov, B. P. 1990. Primer extension technique for the detection of single nucleotide polymorphisms in genomic DNA. Nucleic Acids Research. 18:3671-


Soreng, R. J. 1985. Poa in New Mexico, with a key to middle and southern Rocky Mountain species (Poaceae). Great Basin Naturalist. 45:395-422.


UPOV. 2000. Progress report of the technical committee, the technical working parties and the working group on biochemical and molecular techniques, and DNA-profiling in particular. UPOV Document C/34/10 July 24, 2000.


USGA. 1943. Identity of creeping bent strains planted on experimental greens. Timely Turf Topics. December. p. 3.


Molecular marker map in 'Kanota' × 'Ogle' hexaploid oat (Avena spp.) enhanced by additional markers and a robust framework. Genome. 46:28-47.


Wu, L. and H. Lin. 1994. Identifying Buffalograss [Buchloe dactyloides (Nutt.) Engelm.] cultivar breeding lines using random amplified polymorphic DNA


CHAPTER 2

Isolation and characterization of 88 polymorphic microsatellite markers in Kentucky bluegrass (*Poa pratensis* L.)

Abstract

Kentucky bluegrass (*Poa pratensis* L.) is an important facultative apomictic temperate perennial grass species utilized for both forage and cultivated turf. Through apomixis, this species is able to propagate diverse and odd ploidy levels, resulting in many genetically distinct phenotypes. A wide range of diverse cultivars and accessions of Kentucky bluegrass have been previously characterized based on common turf performance or morphological characteristics, as well as by random amplified polymorphic DNA (RAPD) markers. While previous characterization efforts have provided valuable information, the use of both morphological characteristics and RAPD markers for genetic diversity analysis has limitations. In the current report, we developed and characterized 88 novel microsatellite markers for Kentucky bluegrass. Polymorphism for each marker was assessed in 265 Kentucky bluegrass cultivars, experimental selections, collections, and hybrids. The number of alleles for individual microsatellites ranged from four to 81, with an average of 38.3 alleles per SSR. These polymorphic microsatellite markers would be useful tools for investigating genetic diversity, creation of genetic linkage maps, assessment of levels of apomixis in cultivars and experimental varieties, and identification of aberrant progeny in apomictic Kentucky bluegrass breeding programs.
**Introduction**

The bluegrasses, also commonly referred to as meadowgrasses, are one the most economically important genera of the *Poaceae* (Huff, 2010; Soreng and Barrie, 1999). Kentucky bluegrass (*Poa pratensis* L.) is the botanical-type species for the genus *Poa* (Soreng and Barrie, 1999), and is recognized as one of the most widely utilized temperate perennial grass species for both forage and amenity turf in the northern United States and Canada (Huff, 2003; Huff, 2010). Kentucky bluegrass is a facultative apomictic species, which has a highly variable chromosome number, creating a series of polyploidy and aneuploidy ranging from $2n = 28 – 154$ (Akerberg, 1939; Grazi et al., 1961; Huff, 2003; Love and Love, 1975; Meyer and Funk, 1989; Muntzing, 1933; Nielson, 1946; Tinney, 1940). While this complex polyploidy may sometimes present a challenge to breeding efforts in this species, the ability of Kentucky bluegrass to propagate diverse and odd ploidy levels through apomixis results in many genetically distinct phenotypes within the species (Huff, 2010). This wide range of diversity of cultivars and accessions of Kentucky bluegrass has been previously characterized based on common turf performance or morphological characteristics (Bara et al., 1993; Bonos et al., 2000; Murphy et al., 1997; Shortell et al., 2009), as well as by random amplified polymorphic DNA (RAPD) markers (Curley and Jung, 2004; Huff, 2001; Johnson et al., 2002).

Although currently used in many plant variety protection (PVP) schemes, the use of morphological characteristics to determine genetic diversity and distinctness, uniformity, and stability (DUS) of cultivars has numerous drawbacks including the maintenance of increasingly large reference collections for comparative analyses, a limited number of descriptors available to distinguish varieties, time-consuming field
based measurement of large numbers of samples and replicates, and the potential for the
determination of genetic diversity or for DUS testing, numerous researchers have proposed using
molecular markers for PVP (Bonow et al., 2009; Borchert et al., 2008; Cooke and
Reeves, 2003; Giancola et al., 2002; Gunjaca et al., 2008; Heckenberger et al., 2002;
Heckenberger et al., 2003; Heckenberger et al., 2005a; Ibanez et al., 2009; Kwon et al.,
2005; Roldan-Ruiz et al., 2001; Smith et al., 2009; Smykal et al., 2008; Tommasini et al.,
2003; van Eeuwijk and Law, 2004; Vosman et al., 2004). In contrast to morphological
characters, molecular markers offer a number of advantages including a nearly unlimited
number of characters, high degree of polymorphism, ease of scoring, and they are
unaffected by the environment (Lombard et al., 2001; Smykal et al., 2008; Tommasini et
al., 2003). In this report, we describe the development of the first polymorphic
microsatellite markers for ongoing molecular genetic research in Kentucky bluegrass.

Materials and Methods

Total genomic DNA was extracted from a single plant of the Kentucky bluegrass
cultivar ‘Cabernet’ (Bonos et al., 2004), using Sigma GenElute Plant Genomic DNA
Miniprep Kit (St. Louis, MO) according to the manufacturer’s instructions. DNA was
sent to Genetic Identification Services Inc. (GIS, Chatsworth, CA) for construction of
SSR libraries enriched for CA, GA, AAT, and CAG SSRs. Methods for DNA library
construction and enrichment were developed following Jones et al. (2002). Briefly,
genomic DNA was partially digested with a cocktail of seven blunt-end restriction enzymes (RsaI, HaeIII, BsrB1, PvuII, StuI, ScaI, EcoRV). Fragments in the size range of 300 to 750 bp were adapted with 20 bp oligonucleotides which contained a HindIII site at the 5’ end, and subjected to magnetic bead capture (CPG, Inc., Lincoln Park, New Jersey), using 5’-biotinylated CA(15), GA(15), AAT(12), and CAG(12) as capture molecules according to the manufacturer’s instructions. Captured molecules were amplified using a primer complimentary to the adaptor, digested with HindIII to remove the adaptor sequences, and ligated into the HindIII site of pUC19. Recombinant plasmids were then electroporated into E. coli DH5α. GIS delivered DNA libraries 50% enriched for CA and GA repeats, and libraries 15% enriched for the tri-nucleotides AAT and CAG.

Several 50 ul aliquots of each SSR library were plated out onto LB agar plates containing ampicillin, IPTG and Bluo-Gal. The plates were incubated at 37 °C overnight. Three thousand individual colonies were chosen from the CA and GA enriched libraries, and grown in 6ml LB broth containing ampicillin. Five hundred colonies were chosen from the AAT and CAG enriched libraries. DNA was isolated from the cultures using Qiagen QIAprep Spin Miniprep kit (Valencia, CA). Samples were sequenced using an ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA). A total of 2523 clones were sequenced. Sequence data from clones in the AAT and CAG libraries proved to be complex (long sequences of highly repetitive DNA), and the percentage appropriate for primer design was less than 1%. As a result, only the dinucleotide libraries were used for further analyses. Six hundred and seventy-four dinucleotide clones contained no repeat, while 546 clones had other problems that precluded primer
design (poor sequence, repeat too close to cloning site, etc.). One thousand and seventy-one clones contained either CA or GA repeat motifs suitable for primer design.

Sequence data from the clones containing dinucleotide SSRs were analyzed for primer selection, and polymerase chain reaction (PCR) primers were designed to flank regions surrounding the SSR motif using Primer3 software (Rozen and Skaletsky, 2000). The forward primer in the pair was elongated at the 5’ end using the M13(-21) 18 bp sequence (5’ – TGTAAAACGACGGCCAGT – 3’) for economic fluorescent labeling (Schuelke 2000), and synthesized by Integrated DNA Technologies (Coralville, IA). Of the 1071 clones containing dinucleotide repeats, PCR primer pairs were designed for 500 of these sequences, of which, 88 (17.6%) produced distinct, repeatable polymorphic bands in a panel of 265 Kentucky bluegrass cultivars, experimental selections, collections, and hybrids with Texas bluegrass (*Poa arachnifera* Torr.).

For genotyping, PCR reactions were conducted in 13 ul reactions using 50 ng genomic DNA, 1X Ramp-Taq PCR buffer (Denville Scientific, Metuchen, NJ), 2 mM MgCl₂, 0.25 mM each dNTP (Denville Scientific), 0.5 U Ramp-Taq DNA polymerase (Denville Scientific), 0.5 pmol forward primer with M13(-21) addition, 1 pmol reverse primer, and 1 pmol forward florescent dye-labeled M13(-21) primer (FAM, NED, PET or VIC). Thermal-cycling parameters were 94 °C for 5 min., followed by 30 cycles of 94 °C for 30 s, 55 °C 45 s, 72 °C 45 s, followed by 20 cycles of 94 °C for 5 min., 53 °C for 45 s, 72 °C for 45 s, ending with a final extension for 72 °C for 10 min. PCR reaction products were analyzed on an ABI 3130xl Genetic Analyzer and sized using Genemapper 3.7 software (Applied Biosystems, Foster City, CA) and LIZ 500(-250) size standard (Applied Biosystems, Foster City, CA).
Although microsatellites can generate codominant data, problems can arise during the identification of polyploid genotypes because it is difficult to determine the number of copies of an allele in heterozygotes (Liao et al., 2008; Markwith et al., 2006; Saltonstall, 2003). Thus, banding patterns observed at particular loci are referred to as ‘allele phenotypes’ (Becher et al., 2000). The individual alleles of the SSR markers utilized in the current study of 265 polyploid Kentucky bluegrasses and hybrids were treated as dominant markers, where the banding phenotypes were scored as band absence (0) or presence (1). Polymorphism Information Content (PIC) of each individual SSR allele was calculated according to the formula described by Weir (1990): $\text{PIC} = 1 - \sum P_i^2$, where $P_i$ is the frequency of the $i$th allele in the genotypes examined. For dominant markers, this formula can be simplified to $\text{PIC} = 2P_iQ_i$ where $P_i$ is the frequency of presence and $Q_i$ is the frequency of absence of a particular band (Tehrani et al., 2008). For dominant markers, the maximum PIC value is 0.50. PIC values for the 88 SSR markers in the current study are given as a range of PIC values for all alleles generated by a particular SSR marker (Table 3).

**Results and Discussion**

Characteristics of the SSR markers and their primer sequences are shown in Table 3. All SSR markers produced well-defined discrete alleles. The total number of alleles produced by the 88 SSR markers in the 265 cultivars and accessions was 3373. The number of alleles for individual SSR markers ranged from four to 81, with an average of 38.3 alleles per SSR. The high average number of alleles per SSR is likely due to the combination of high genetic diversity and high ploidy levels in the genotypes under
Table 3. Primer sequences and characteristics of 88 *Poa pratensis* (L.) microsatellite markers obtained from the cultivar ‘Cabernet’, tested on 265 *Poa pratensis* cultivars, experimental selections, collections, and hybrids.

<table>
<thead>
<tr>
<th>Marker ID</th>
<th>GenBank Accession no</th>
<th>Primers (5’-3’)</th>
<th>Repeat motif</th>
<th>N&lt;sub&gt;a&lt;/sub&gt;</th>
<th>Original cloned allele (bp)</th>
<th>Allele size range (bp)</th>
<th>Range of PIC values for N&lt;sub&gt;a&lt;/sub&gt; alleles</th>
</tr>
</thead>
</table>
| NJPpGA1   | HM136687             | F: AAGGCTCGTTGAGTTTCCAG  
  R: TTTGGAAGGGAGGAGCAGG | (CT)<sup>12</sup>(CG)(CT)<sup>16</sup> | 40 | 252 | 199-305 | 0.01 – 0.48 |
| NJPpGA6   | HM136688             | F: GGTTGGTCCTGCTCAAGAAGTGG  
  R: GTGACCCTTCTAGCCTGATAT | (GA)<sup>6</sup>A(GA)<sup>7</sup> | 22 | 279 | 181-314 | 0.01 – 0.45 |
| NJPpGA9   | HM136689             | F: GCCGTAAATAGTGAGAGACAGG  
  R: AAAATCCGGACTGTTGGAGAC | (CT)<sup>21</sup> | 40 | 208 | 142-275 | 0.01 – 0.50 |
| NJPpGA107 | HM136690             | F: TGTTGCTACAGATTTTTTATCC  
  R: AAGTTGAGTCCTCATTGAGAC | (GA)<sup>22</sup> | 39 | 238 | 187-268 | 0.01 – 0.50 |
| NJPpGA108 | HM136691             | F: GAGTGGAGACACACCAACACATG  
  R: CTTTCTTTAAGGGGACAG | (GA)<sup>20</sup> | 81 | 213 | 161-292 | 0.01 – 0.44 |
| NJPpGA111 | HM136692             | F: TGCAAGCAGTCTTATATGTACC  
  R: GAGAGGAGAAGTGGAGGACTC | (GA)<sup>19</sup> | 32 | 229 | 179-268 | 0.01 – 0.43 |
| NJPpGA115 | HM136693             | F: CCACCAGAGACATTGGACTG  
  R: CTTGCACTATAGCAAGGCC | (GA)<sup>18</sup> | 24 | 277 | 281-328 | 0.01 – 0.50 |
| NJPpGA124 | HM136694             | F: GCCCTCTCTCCAAAGGATAC  
  R: ATGTTGCTGTGTTGTTCC | (CT)<sup>26</sup> | 33 | 114 | 117-173 | 0.01 – 0.50 |
| NJPpGA125 | HM136695             | F: GCAAGAACAAAAAGTCATACTGG  
  R: AACCCCTTTAGCCCTACAC | (GA)<sub>6</sub>G(GA)<sub>18</sub> | 64 | 231 | 158-439 | 0.01 – 0.50 |
| NJPpGA128 | HM136696             | F: GAGGCGGAAGAGTAGATGC  
  R: AGATGGCGAGAACGAAGTCC | (CT)<sup>17</sup> | 30 | 156 | 131-199 | 0.01 – 0.50 |
| NJPpGA129 | HM136697             | F: CCACGCAACATCAAGAAGGACAC  
  R: TCCGGAAGAATGTTGTTGG | (CT)<sup>13</sup> | 27 | 296 | 272-322 | 0.01 – 0.48 |
| NJPpGA132 | HM136698             | F: TTTGCTTCACTTCACTACG  
  R: AGATTGCGAGAAAGATGCC | (CT)<sub>7</sub>T(T)<sub>2</sub> | 4 | 300 | 311-321 | 0.01 – 0.50 |
| NJPpGA134 | HM136699             | F: ACACCCCTTTGTAGATTTCG  
  R: CTTTGTCTGGTTCTTCATC | (CT)<sub>7</sub>GT(T)<sub>22</sub> | 15 | 150 | 129-172 | 0.02 – 0.48 |
| NJPpGA274 | HM136700             | F: GAGGACAAAGAATCTGAGACT  
  R: CAAAATGACGACAAACGGGAT | (CT)<sub>19</sub> | 35 | 298 | 191-443 | 0.01 – 0.50 |
| NJPpGA329 | HM136701             | F: TGCTATTTTCAATTGGAAAGCA  
  R: CGAATTTGGAACCTGACATC | (CT)<sub>23</sub> | 35 | 300 | 292-382 | 0.01 – 0.48 |
<table>
<thead>
<tr>
<th>Marker ID</th>
<th>GenBank Accession no.</th>
<th>Primers (5'-3')</th>
<th>Repeat motif</th>
<th>Original cloned allele (bp)</th>
<th>Allele size range (bp)</th>
<th>Range of PIC values for N&lt;sub&gt;a&lt;/sub&gt; alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJPpGA379</td>
<td>HM136702</td>
<td>F: AAATTTTGACCATGCAACCC</td>
<td>(GA)&lt;sub&gt;18&lt;/sub&gt;</td>
<td>42</td>
<td>295</td>
<td>286-370</td>
</tr>
<tr>
<td>NJPpGA393</td>
<td>HM136703</td>
<td>F: CATGAGAGACACGAAACAGGA</td>
<td>(CT)&lt;sub&gt;10&lt;/sub&gt;</td>
<td>43</td>
<td>248</td>
<td>239-335</td>
</tr>
<tr>
<td>NJPpGA403</td>
<td>HM136704</td>
<td>F: GTCAATTCGGCCGTGAGTC</td>
<td>(GA)&lt;sub&gt;15&lt;/sub&gt;</td>
<td>18</td>
<td>262</td>
<td>231-297</td>
</tr>
<tr>
<td>NJPpGA405</td>
<td>HM136705</td>
<td>F: GACGAAATGGATAACCTGACAGA</td>
<td>(GA)&lt;sub&gt;21&lt;/sub&gt;</td>
<td>38</td>
<td>300</td>
<td>293-379</td>
</tr>
<tr>
<td>NJPpGA412</td>
<td>HM136706</td>
<td>F: GCACCGTGAGACAAAGTTATT</td>
<td>(CT)&lt;sub&gt;17&lt;/sub&gt;</td>
<td>64</td>
<td>260</td>
<td>244-335</td>
</tr>
<tr>
<td>NJPpGA422</td>
<td>HM136707</td>
<td>F: TGCTGTCCCTCAACATCAAT</td>
<td>(CT)&lt;sub&gt;24&lt;/sub&gt;</td>
<td>43</td>
<td>296</td>
<td>267-419</td>
</tr>
<tr>
<td>NJPpGA424</td>
<td>HM136708</td>
<td>F: GGTCAAAAGTTAAGACGATTGG</td>
<td>(CA)&lt;sub&gt;20(GA)&lt;sub&gt;14&lt;/sub&gt;</td>
<td>41</td>
<td>284</td>
<td>283-336</td>
</tr>
<tr>
<td>NJPpGA426</td>
<td>HM136709</td>
<td>F: GGTGACATCCCTGTGCTAA</td>
<td>(GA)&lt;sub&gt;25(GT(GA)&lt;sub&gt;28&lt;/sub&gt;</td>
<td>32</td>
<td>294</td>
<td>242-313</td>
</tr>
<tr>
<td>NJPpGA434</td>
<td>HM136710</td>
<td>F: GAAGATAAATGACACATCAAAGACA</td>
<td>(CT)&lt;sub&gt;34(CG)&lt;sub&gt;31(CCC)&lt;sub&gt;31&lt;/sub&gt;</td>
<td>41</td>
<td>399</td>
<td>351-425</td>
</tr>
<tr>
<td>NJPpGA446</td>
<td>HM136711</td>
<td>F: TGCGCTCATAATGCTCAAAA</td>
<td>(CT)&lt;sub&gt;25&lt;/sub&gt;</td>
<td>79</td>
<td>285</td>
<td>153-409</td>
</tr>
<tr>
<td>NJPpGA448</td>
<td>HM136712</td>
<td>F: ATCGTCAAGGGAGAATC</td>
<td>(CT)&lt;sub&gt;37&lt;/sub&gt;</td>
<td>36</td>
<td>297</td>
<td>187-317</td>
</tr>
<tr>
<td>NJPpGA450</td>
<td>HM136713</td>
<td>F: CTGGTTGCTAAACCATTTGC</td>
<td>(CT)&lt;sub&gt;15(TT)(CT)&lt;sub&gt;7&lt;/sub&gt;</td>
<td>66</td>
<td>281</td>
<td>216-340</td>
</tr>
<tr>
<td>NJPpGA470</td>
<td>HM136714</td>
<td>F: TTGAAAGCCACCACCAATTA</td>
<td>(GA)&lt;sub&gt;15&lt;/sub&gt;</td>
<td>48</td>
<td>290</td>
<td>269-494</td>
</tr>
<tr>
<td>NJPpGA480</td>
<td>HM136715</td>
<td>F: AGCGGAGATCTCTACGAGTC</td>
<td>(CT)&lt;sub&gt;24&lt;/sub&gt;</td>
<td>52</td>
<td>187</td>
<td>161-242</td>
</tr>
<tr>
<td>NJPpGA748</td>
<td>HM136716</td>
<td>F: GCCTTTATATCGATCGAGAAA</td>
<td>(GA)&lt;sub&gt;25(GC(GA)&lt;sub&gt;24&lt;/sub&gt;</td>
<td>37</td>
<td>299</td>
<td>264-339</td>
</tr>
<tr>
<td>Marker ID</td>
<td>GenBank Accession no.</td>
<td>Primers (5'-3')</td>
<td>Repeat motif</td>
<td>N&lt;sub&gt;a&lt;/sub&gt;</td>
<td>Original cloned allele (bp)</td>
<td>Allele size range (bp)&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>NJPpGA749</td>
<td>HM136717</td>
<td>F: CTGGCAGAGTACGCAACATAGTTG R: AGGGAACTGATCTTGCTAT</td>
<td>(GA)&lt;sub&gt;2&lt;/sub&gt;TA(GA)&lt;sub&gt;17&lt;/sub&gt;</td>
<td>35</td>
<td>279</td>
<td>253-337</td>
</tr>
<tr>
<td>NJPpGA753</td>
<td>HM136718</td>
<td>F: TCAGGAGACTGAGCAAGATAATTG R: GAATGATTGAGTGTTATGG</td>
<td>(CT)&lt;sub&gt;17&lt;/sub&gt;</td>
<td>49</td>
<td>274</td>
<td>255-357</td>
</tr>
<tr>
<td>NJPpGA771</td>
<td>HM136719</td>
<td>F: TCAGGAGTACAGAGAAATTG R: CTATTTTCGATCTTGCTAT</td>
<td>(GA)&lt;sub&gt;13&lt;/sub&gt;</td>
<td>36</td>
<td>250</td>
<td>250-317</td>
</tr>
<tr>
<td>NJPpGA783</td>
<td>HM136720</td>
<td>F: TGCCCAAGAAATAATTATGA R: GGGCTCAGCTAAGCTCAAAT</td>
<td>(GA)&lt;sub&gt;17&lt;/sub&gt;AA(GA)&lt;sub&gt;25&lt;/sub&gt;</td>
<td>77</td>
<td>365</td>
<td>267-398</td>
</tr>
<tr>
<td>NJPpGA799</td>
<td>HM136721</td>
<td>F: TTGTTGAACTGCTTCTGAAATG R: GAATTTGGGAGGCAATGG</td>
<td>(CT)&lt;sub&gt;14&lt;/sub&gt;</td>
<td>29</td>
<td>300</td>
<td>231-335</td>
</tr>
<tr>
<td>NJPpGA806</td>
<td>HM136722</td>
<td>F: TCCCTCAAGTGACCTACGAA R: TGTAACCCAGAAGTTAGCTAT</td>
<td>(CT)&lt;sub&gt;21&lt;/sub&gt;</td>
<td>38</td>
<td>298</td>
<td>249-352</td>
</tr>
<tr>
<td>NJPpGA815</td>
<td>HM136723</td>
<td>F: CACTAAAGCCAAACACAGA R: AAATGTGAGGAGGATGGA</td>
<td>(GA)&lt;sub&gt;13&lt;/sub&gt;AA(GA)&lt;sub&gt;5&lt;/sub&gt;</td>
<td>41</td>
<td>286</td>
<td>179-365</td>
</tr>
<tr>
<td>NJPpGA832</td>
<td>HM136724</td>
<td>F: TCAGGAGCATACGAGATGAAAAC R: CTAATTTGGTGATCTCGCT</td>
<td>(GA)&lt;sub&gt;13&lt;/sub&gt;</td>
<td>45</td>
<td>250</td>
<td>236-338</td>
</tr>
<tr>
<td>NJPpGA836</td>
<td>HM136725</td>
<td>F: AACATAAAAGCCAAACACAGA R: AAATGTGAGGAGGATGGA</td>
<td>(GA)&lt;sub&gt;19&lt;/sub&gt;</td>
<td>50</td>
<td>286</td>
<td>257-365</td>
</tr>
<tr>
<td>NJPpGA892</td>
<td>HM136726</td>
<td>F: TGGCGCTAGGTCTACCTTTT R: GGAATCCCTACGTCTCTGTA</td>
<td>(GA)&lt;sub&gt;13&lt;/sub&gt;</td>
<td>20</td>
<td>250</td>
<td>224-291</td>
</tr>
<tr>
<td>NJPpGA897</td>
<td>HM136727</td>
<td>F: ACTCCAAAGCACCAGAGTGGAG R: CTCCTGCCGCTCTTGGG</td>
<td>(GA)&lt;sub&gt;30&lt;/sub&gt;GG(GA)&lt;sub&gt;3&lt;/sub&gt;</td>
<td>40</td>
<td>300</td>
<td>267-362</td>
</tr>
<tr>
<td>NJPpGA900</td>
<td>HM136728</td>
<td>F: ATCCGATTTGGCACGCAGTTA R: GCTAGGAGGCTTAAAGCTAGG</td>
<td>(CT)&lt;sub&gt;22&lt;/sub&gt;</td>
<td>42</td>
<td>299</td>
<td>152-374</td>
</tr>
<tr>
<td>NJPpGA914</td>
<td>HM136729</td>
<td>F: CCCCCATTCCACTCAAGTTT R: GATATGGACACCAACCATGC</td>
<td>(GA)&lt;sub&gt;19&lt;/sub&gt;</td>
<td>60</td>
<td>278</td>
<td>274-353</td>
</tr>
<tr>
<td>NJPpGA931</td>
<td>HM136730</td>
<td>F: TCAGGAGGCTACGTACCTTACG R: TCCAAACGCTACCTATCCAG</td>
<td>(GA)&lt;sub&gt;25&lt;/sub&gt;</td>
<td>23</td>
<td>299</td>
<td>289-342</td>
</tr>
<tr>
<td>NJPpGA937</td>
<td>HM136731</td>
<td>F: TACGCTTTCGGTGGAGAGTTGG R: TCCAACGGCTACCTTACG</td>
<td>(GA)&lt;sub&gt;22&lt;/sub&gt;</td>
<td>20</td>
<td>298</td>
<td>293-337</td>
</tr>
<tr>
<td>Marker ID</td>
<td>GenBank Accession no</td>
<td>Primers (5'-3')</td>
<td>Repeat motif</td>
<td>$N_s$</td>
<td>Original cloned allele (bp)</td>
<td>Allele size range (bp)</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------</td>
<td>----------------</td>
<td>-------------</td>
<td>------</td>
<td>---------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>NJPpGA947</td>
<td>HM136732</td>
<td>F: TGATCAGTTGGCTTTGTGCA</td>
<td>(CT)$_4$GT(CT)$_3$</td>
<td>22</td>
<td>245</td>
<td>183-290</td>
</tr>
<tr>
<td>NJPpGA954</td>
<td>HM136733</td>
<td>F: ATCTTCAGATCTGTTGGATGC</td>
<td>(GT)$_7$A(GA)$_6$</td>
<td>12</td>
<td>289</td>
<td>286-419</td>
</tr>
<tr>
<td>NJPpGA957</td>
<td>HM136734</td>
<td>F: TCCACGGGTATTTTTCCGTTA</td>
<td>(CT)$_17$</td>
<td>25</td>
<td>286</td>
<td>284-3550</td>
</tr>
<tr>
<td>NJPpGA963</td>
<td>HM136734</td>
<td>F: TACGCTGACACACAAATTTGC</td>
<td>(CT)$_15$</td>
<td>28</td>
<td>291</td>
<td>289-350</td>
</tr>
<tr>
<td>NJPpGA964</td>
<td>HM136736</td>
<td>F: ATATTGAGCGGTTGCTCT</td>
<td>(GA)$_4$CA(GA)$_3$</td>
<td>28</td>
<td>291</td>
<td>264-318</td>
</tr>
<tr>
<td>NJPpGA986</td>
<td>HM136736</td>
<td>F: GACATCAAGACCTTTGACGAA</td>
<td>(GA)$_17$</td>
<td>25</td>
<td>298</td>
<td>292-346</td>
</tr>
<tr>
<td>NJPpGA993</td>
<td>HM136736</td>
<td>F: AITAAGCAGTGGGCAATGAG</td>
<td>(CT)$_25$</td>
<td>29</td>
<td>292</td>
<td>274-338</td>
</tr>
<tr>
<td>NJPpGA1023</td>
<td>HM136739</td>
<td>F: CATGAGAGACGACACAGGA</td>
<td>(CT)$_19$</td>
<td>35</td>
<td>296</td>
<td>285-351</td>
</tr>
<tr>
<td>NJPpGA1033</td>
<td>HM136740</td>
<td>F: GGGGATACAAACTCGACAT</td>
<td>(GA)$_27$</td>
<td>23</td>
<td>293</td>
<td>270-316</td>
</tr>
<tr>
<td>NJPpGA1054</td>
<td>HM136741</td>
<td>F: ATACATGGGCGATGTG</td>
<td>(CT)$_23$</td>
<td>13</td>
<td>286</td>
<td>243-305</td>
</tr>
<tr>
<td>NJPpGA1071</td>
<td>HM136742</td>
<td>F: CACATTGGTTGGATCAT</td>
<td>(GA)$_15$GG(GA)$_23$</td>
<td>61</td>
<td>334</td>
<td>240-357</td>
</tr>
<tr>
<td>NJPpGA1092</td>
<td>HM136743</td>
<td>F: CATGACGTTAAGAAGGGTG</td>
<td>(GA)$_10$TA(GA)$_28$</td>
<td>38</td>
<td>299</td>
<td>241-317</td>
</tr>
<tr>
<td>NJPpGA1095</td>
<td>HM136744</td>
<td>F: AACTTTCAGATCAGATCATTATCC</td>
<td>(CT)$_2$GTCTGT(CT)$_4$C(CT)$_22$</td>
<td>54</td>
<td>298</td>
<td>238-337</td>
</tr>
<tr>
<td>NJPpGA1100</td>
<td>HM136745</td>
<td>F: CTAATTTGGTGAGATTTGCT</td>
<td>(CT)$_14$</td>
<td>47</td>
<td>250</td>
<td>249-336</td>
</tr>
<tr>
<td>NJPpGA1101</td>
<td>HM136746</td>
<td>F: GAGACCIAAAATACTGTCTC</td>
<td>(CT)$_18$</td>
<td>35</td>
<td>298</td>
<td>285-342</td>
</tr>
<tr>
<td>Marker ID</td>
<td>GenBank Accession no.</td>
<td>Primers (5’-3’)</td>
<td>Repeat motif</td>
<td>$N_a$</td>
<td>Original cloned allele (bp)</td>
<td>Allele size range (bp)</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-------</td>
<td>---------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>NJpGA110</td>
<td>HM136747</td>
<td>F: TTCTCTCTCTCCATCTTTG</td>
<td>(CT)$_{20}$</td>
<td>38</td>
<td>300</td>
<td>247-329</td>
</tr>
<tr>
<td>NJpGA117</td>
<td>HM136748</td>
<td>F: TGAGGATGGTCTGCTGCTGAG</td>
<td>(GA)$_{26}$</td>
<td>39</td>
<td>291</td>
<td>240-365</td>
</tr>
<tr>
<td>NJpGA1112</td>
<td>HM136749</td>
<td>F: AACGTCATGCTTCCGACCACATA</td>
<td>(GA)$<em>{8}$AA(GA)$</em>{14}$</td>
<td>55</td>
<td>281</td>
<td>236-328</td>
</tr>
<tr>
<td>NJpGA112</td>
<td>HM136750</td>
<td>F: CTTGTATTGGAGAGTTGTGG</td>
<td>(GA)$_{25}$</td>
<td>25</td>
<td>299</td>
<td>253-360</td>
</tr>
<tr>
<td>NJpGA118</td>
<td>HM136751</td>
<td>F: TCCACGGCTACCTATTTCA</td>
<td>(GA)$_{10}$</td>
<td>40</td>
<td>285</td>
<td>257-344</td>
</tr>
<tr>
<td>NJpGA148</td>
<td>HM136752</td>
<td>F: CAAAGGGAATGTAAGAGG</td>
<td>(GA)$_{14}$</td>
<td>40</td>
<td>250</td>
<td>234-289</td>
</tr>
<tr>
<td>NJpGA152</td>
<td>HM136753</td>
<td>F: AACGTCATGCTTCCGACAAT</td>
<td>(GA)$<em>{8}$AA(GA)$</em>{13}$</td>
<td>58</td>
<td>281</td>
<td>216-328</td>
</tr>
<tr>
<td>NJpGA153</td>
<td>HM136754</td>
<td>F: TTAAGACCCCTACCAATGCAA</td>
<td>(CT)$_{14}$</td>
<td>41</td>
<td>291</td>
<td>276-360</td>
</tr>
<tr>
<td>NJpGA9307</td>
<td>HM136755</td>
<td>F: AGACAAAAAGACACACAAAGA</td>
<td>(GA)$_{20}$</td>
<td>67</td>
<td>283</td>
<td>258-394</td>
</tr>
<tr>
<td>NJpGA9314</td>
<td>HM136756</td>
<td>F: GGACAACGTAGCTCTTTCAGT</td>
<td>(GA)$_{10}$</td>
<td>48</td>
<td>290</td>
<td>273-410</td>
</tr>
<tr>
<td>NJpGA9317</td>
<td>HM136757</td>
<td>F: GTGACACGTAGCTGTTCAGT</td>
<td>(GA)$_{13}$</td>
<td>53</td>
<td>294</td>
<td>240-357</td>
</tr>
<tr>
<td>NJpGA9324</td>
<td>HM136758</td>
<td>F: TCATGATGTAGTGACCAGC</td>
<td>(CT)$<em>{10}$CG(CT)$</em>{6}$</td>
<td>33</td>
<td>278</td>
<td>245-329</td>
</tr>
<tr>
<td>NJpGA9326</td>
<td>HM136759</td>
<td>F: TAGGCGTCCAGTTAAGAAAA</td>
<td>(GA)$_{13}$</td>
<td>38</td>
<td>269</td>
<td>257-341</td>
</tr>
<tr>
<td>NJpGA9335</td>
<td>HM136760</td>
<td>F: ATTTTTCGTTGGCGGTTT</td>
<td>(CT)$_{15}$</td>
<td>31</td>
<td>250</td>
<td>237-298</td>
</tr>
<tr>
<td>NJpGA9361</td>
<td>HM136761</td>
<td>F: TAGCAATGCTCCAAAGTAGR</td>
<td>(CA)$_{11}$</td>
<td>40</td>
<td>250</td>
<td>235-309</td>
</tr>
</tbody>
</table>
Table 3. cont.

<table>
<thead>
<tr>
<th>Marker ID</th>
<th>GenBank Accession no.</th>
<th>Primers (5’-3’)</th>
<th>Repeat motif</th>
<th>$N_a$</th>
<th>Original cloned allele range (bp)</th>
<th>Allele size range (bp)$^1$</th>
<th>Range of PIC values for $N_a$ alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJPpGT8</td>
<td>HM136762 F: AAAAAATTAAGCGGATGCTATTGCTAG R: AATCGCAGGATTTCAGTATG</td>
<td>(GT)$_{12}$</td>
<td>31</td>
<td>223</td>
<td>206-321</td>
<td>0.01 – 0.48</td>
<td></td>
</tr>
<tr>
<td>NJPpGT12</td>
<td>HM136763 F: GAGTTGGTCGCTGGTGTAG R: CGCTATTCTCTTTAGTGGTATGC</td>
<td>(CA)$<em>4$CG(CA)$</em>{11}$CG(CA)$_4$</td>
<td>11</td>
<td>226</td>
<td>227-320</td>
<td>0.01 – 0.50</td>
<td></td>
</tr>
<tr>
<td>NJPpGT102</td>
<td>HM136764 F: GGTCCTGGGATGCTGTAG T: CGTGTAACATCCGGCTTAC</td>
<td>CAGA(CA)$_{13}$</td>
<td>21</td>
<td>178</td>
<td>164-246</td>
<td>0.01 – 0.46</td>
<td></td>
</tr>
<tr>
<td>NJPpGT110</td>
<td>HM136765 F: GTCCTCGTGGTTCTCTCGT</td>
<td>(GT)$_{11}$</td>
<td>24</td>
<td>221</td>
<td>197-254</td>
<td>0.01 – 0.50</td>
<td></td>
</tr>
<tr>
<td>NJPpGT113</td>
<td>HM136766 F: CGGTGGAGAGGATGATGAGTATG R: CACCCCGCTCTTAAAGGCTATCG</td>
<td>(GA)$<em>{10}$(GT)$</em>{10}$TT(GT)$_6$</td>
<td>40</td>
<td>190</td>
<td>151-239</td>
<td>0.01 – 0.49</td>
<td></td>
</tr>
<tr>
<td>NJPpGT123</td>
<td>HM136767 F: GATTGGGGCTTTCAACTGATTC R: CACCCCGCTCTTAAAGGCTATCG</td>
<td>(CA)$_{13}$</td>
<td>32</td>
<td>142</td>
<td>128-221</td>
<td>0.01 – 0.50</td>
<td></td>
</tr>
<tr>
<td>NJPpGT134</td>
<td>HM136768 F: CCAAAAAGATTACTCAGCGTAC R: TGTGAAGAAACACCAAGTGGAA</td>
<td>(GT)$_{33}$</td>
<td>70</td>
<td>138</td>
<td>95-254</td>
<td>0.01 – 0.50</td>
<td></td>
</tr>
<tr>
<td>NJPpGT135</td>
<td>HM136769 F: GCCGCTCTCTTGCTGATT R: CGGGTAAGGTCTCTGTGTTG</td>
<td>(GT)$_{29}$</td>
<td>57</td>
<td>174</td>
<td>132-241</td>
<td>0.01 – 0.50</td>
<td></td>
</tr>
<tr>
<td>NJPpGT136</td>
<td>HM136770 F: TTCTTTGTCCCTTACTTCTG R: ATGACGGAGGCTTTACG</td>
<td>(GT)$_2$GC(GT)$_7$</td>
<td>17</td>
<td>226</td>
<td>237-253</td>
<td>0.01 – 0.46</td>
<td></td>
</tr>
<tr>
<td>NJPpGT9314</td>
<td>HM136771 F: CTCAGTGACGGAACAGCTTAA R: GCTGCTAAAGGGCACCCTCCTAACC</td>
<td>(CA)$_4$CG(CA)$_5$</td>
<td>45</td>
<td>295</td>
<td>244-313</td>
<td>0.01 – 0.48</td>
<td></td>
</tr>
<tr>
<td>NJPpGT9318</td>
<td>HM136772 F: CGAGTAGGAAATCACTCCAACA R: ACCACCATGCAGGACCTCTG</td>
<td>(CA)$_{15}$</td>
<td>18</td>
<td>284</td>
<td>248-302</td>
<td>0.01 – 0.47</td>
<td></td>
</tr>
<tr>
<td>NJPpGT9434</td>
<td>HM136773 F: CCCCTATGGCCAGACGT</td>
<td>(GT)$_{15}$</td>
<td>43</td>
<td>249</td>
<td>182-311</td>
<td>0.01 – 0.50</td>
<td></td>
</tr>
<tr>
<td>NJPpGT9446</td>
<td>HM136774 F: TGTTCCTGTGACTGAATCTCACTCTG R: GCACACCATGCTCATCTTGA</td>
<td>(GT)$_{10}$</td>
<td>37</td>
<td>248</td>
<td>254-308</td>
<td>0.01 – 0.48</td>
<td></td>
</tr>
</tbody>
</table>

$N_a$: number of alleles

$^1$, includes M13(-21) 18 bp sequence (5’ – TGTAAAACGACGGCCAGT – 3’) on the forward primer
investigation. The PIC value ranges for the majority of the SSR markers approached the maximum PIC value for dominant markers, indicating that each SSR marker produced alleles that were highly polymorphic (Table 3). PIC values and expected heterozygosity (H_E) of the 3373 individual alleles are available upon request. The high level of polymorphism observed for the described microsatellite markers support their application in genetic studies of Kentucky bluegrass. The potential utility of these markers includes assessment of genetic diversity, creation of genetic linkage maps, assessment of levels of apomixis in cultivars and experimental varieties, and identification of aberrant progeny in apomictic Kentucky bluegrass breeding programs. Additionally, the large number of microsatellite markers described in this report falls within a range (75 – 100 SSRs) that other researchers have suggested to be appropriate for PVP analyses in other crops (Heckenberger et al., 2005b; Smith et al., 2009).
Literature Cited


Heckenberger, M., M. Bohn, D. Klein, and A. E. Melchinger. 2005b. Identification of essentially derived varieties obtained from biparental crosses of homozygous lines: II. Morphological distances and heterosis in comparison with simple sequence repeat and amplified fragment length polymorphism data in maize. Crop Sci. 45:1132-1140.


Shortell, R. R., W. A. Meyer, and S. A. Bonos. 2009. Classification and inheritance of
morphological and agronomic characteristics in Kentucky bluegrass (*Poa

Genetic diversity among US sunflower inbreds and hybrids: assessing probability
of ancestry and potential for use in plant variety protection. Crop Sci. 49:1295-
1303.

(*Pisum sativum* L.) by molecular, biochemical and morphological markers. J. of

Soreng, R. J. and F. R. Barrie. 1999. Proposal to conserve the name *Poa pratensis*
(Grassinae) with a conserved type. Taxon. 48:157-159.

Transferability of genomic and EST-microsatellites from *Festuca arundinacea*
Schreb. to *Lolium pericicum* Boiss. and Hohen. ex Boiss. Intl. J. of Bot. 4:476-
480.

60:351-360.

sequence repeat (SSR) markers to complement distinctness, uniformity and
106:1091-1101.

illustrations based on lettuce and barley. Euphytica. 137:129-137.

The establishment of ‘essential derivation’ among rose varieties, using AFLP.

CHAPTER 3

Classification of Kentucky bluegrass (*Poa pratensis* L.) cultivars and accessions I: Characterization based on microsatellite markers

Abstract

Kentucky bluegrass (*Poa pratensis* L.) is an important facultative apomictic temperate perennial grass species utilized for both forage and cultivated turf. Through apomixis, this species is able to propagate diverse and odd ploidy levels, resulting in many genetically distinct phenotypes. A wide range of diverse cultivars and accessions of Kentucky bluegrass have been previously characterized based on pedigree, common turf performance, and morphological characteristics to create a Kentucky bluegrass cultivar classification system. The objectives of the current study were to genotype 265 Kentucky bluegrass cultivars, experimental selections, collections, and hybrids using microsatellite markers, and compare the results with the original classification system. Contrary to previous research using RAPD markers, SSR markers showed a strong correlation between genetic relatedness as assessed by molecular markers and the original Kentucky bluegrass classification system, and also provided justification for a revision/update of the classification system. Traditional classification types that are supported by the current SSR analysis include BVMG, Compact, Compact-America, Mid-Atlantic, Julia, Shamrock, and Midnight types. Newly proposed classification types include Eurasian, Sydsport related, Limousine related, and P-105 related. All cultivars, experimental selections, collections, and hybrids were uniquely identified with the current set of SSR markers. Genetic relationships of individuals as assessed by SSR
markers closely matched known pedigrees. The current set of SSR markers can be used to rapidly genotype and assign new cultivars/accessions to Kentucky bluegrass classification types and assess genetic relatedness among individuals, and should be considered for use in a Kentucky bluegrass Plant Variety Protection program.

**Introduction**

The bluegrasses, also commonly referred to as meadowgrasses, are one of the most economically important genera of the *Poaceae* (Soreng and Barrie, 1999; Huff, 2010). Kentucky bluegrass (*Poa pratensis* L.) is the botanical-type species for the genus *Poa* (Soreng and Barrie, 1999), and is recognized as one of the most widely utilized temperate perennial grass species for both forage and amenity turf in the northern United States and Canada (Huff, 2003; Huff, 2010). According to Funk (2000), Kentucky bluegrass is lauded as the ‘king of grasses’ and the ‘final triumph of nature’, and is acclaimed as the premiere lawn grass, dominating or contributing to the turf of 40 million lawns throughout Canada and all but the warmest regions of the United States. It is also widely used on golf courses, athletic fields, and in permanent pastures throughout the cool-humid and transitional climates of the world (Beard, 1973). Kentucky bluegrass is attractive, hardy, aggressive, and persistent, and is adapted to a wide range of soils and climates (Bashaw and Funk, 1987; Funk, 2000). Of all the cool-season turfgrasses, Kentucky bluegrass has the most extensive rhizomes which contribute significantly to its ability to produce an excellent sod and survive stressful environments/conditions (Meyer and Funk, 1989). Its overall aesthetic appeal as a lawn grass sets the standard against
which most other turfgrasses are compared due to its combination of softness, medium-to-fine leaf texture, high shoot density, dark green color, and persistence (Huff, 2003). Additionally, as a highly palatable and nutritious forage, Kentucky bluegrass is used for grazing on over 15 million hectares of pastures throughout the northeastern and north-central U.S. and substantial areas of Canada and Europe (Bashaw and Funk, 1987; Funk, 2000).

Kentucky bluegrass is a facultative apomictic species, which has a highly variable chromosome number, creating a series of polyploidy and aneuploidy ranging from $2n = 28$ – $154$ (Muntzing, 1933; Akerberg, 1939; Tinney, 1940; Nielson, 1946; Grazi et al., 1961; Love and Love, 1975; Meyer and Funk, 1989; Huff, 2003). While this complex polyploidy may sometimes present a challenge to breeding efforts in this species, the ability of Kentucky bluegrass to propagate diverse and odd ploidy levels through apomixis results in many genetically distinct phenotypes within the species (Huff, 2010). Highly diverse cultivars have been developed from directly selecting promising apomictic plants from natural populations (Pepin and Funk, 1971; Bashaw and Funk, 1987; Bonos et al., 2000) and from the improvement of intraspecific hybridization techniques (Pepin and Funk, 1971).

A classification system, based on varying combinations of pedigree information, common turf performance characteristics, and morphological traits, was previously developed to characterize the large number of Kentucky bluegrass cultivars and accessions (Bara et al., 1993; Murphy et al., 1997; Bonos et al., 2000; Shortell et al., 2009). This classification system was developed to provide an overview of the similarities and differences between cultivars, and as a guide to help turf managers
develop blends of cultivars that are genetically diverse yet still uniform in morphological and performance characteristics (Shortell et al., 2009). A detailed discussion of the original classification scheme, the description of the classification types, as well as subsequent updates, is available in previous publications (Bara et al., 1993; Murphy et al., 1997; Bonos et al., 2000; Shortell et al., 2004; Shortell et al., 2009). A brief summary and description of the most recent version of the classification system (adapted from Shortell et al., 2009) is shown in Table 4.

The Kentucky bluegrass classification system has been useful to both researchers and turf managers, however, the use of morphological traits and cultivar performance characteristics to distinguish varieties has numerous drawbacks including the maintenance of increasingly large reference collections for comparative analyses, a limited number of descriptors available to distinguish varieties, time-consuming field based measurement of large numbers of samples and replicates, and the potential for the expression of morphological traits to be influenced by the environment (Lombard et al., 2000; Roldan-Ruiz et al., 2001; Giancola et al., 2002; Kwon et al., 2005; Ibanez et al., 2009). Additionally, common morphological or performance characteristics may not equate to genetic relatedness. Due to the above drawbacks of using morphological characters to distinguish varieties, numerous researchers have proposed using molecular markers for variety discrimination, genetic diversity studies, and Plant Variety Protection (PVP) (Roldan-Ruiz et al., 2001; Giancola et al., 2002; Heckenberger et al., 2002; Cooke and Reeves, 2003; Heckenberger et al., 2003; Tommasini et al., 2003; van Eeuwijk and Law, 2004; Vosman et al., 2004; Heckenberger et al., 2005; Kwon et al., 2005; Borchert
Table 4. Kentucky bluegrass classification types, brief description of types, and example cultivars.

<table>
<thead>
<tr>
<th>Classification type</th>
<th>Attributes of cultivars in classification type</th>
<th>Example cultivars</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Density (Aggressive)</td>
<td>aggressive lateral growth resulting in a high density turf; cultivars may dominate a turf stand when used in blends with other bluegrass types or other turf species; high thatch production under certain environmental and management conditions</td>
<td>Princeton 104, A-34, Limousine, Touchdown</td>
</tr>
<tr>
<td>BVMG</td>
<td>medium low growing turf with medium density and medium-wide leaves; moderate to good turf quality; high seed yield potential; high susceptibility to certain races of stripe smut (<em>Ustilago striiformis</em>)</td>
<td>Baron, Victa, Merit, Gnome</td>
</tr>
<tr>
<td>Common (Midwest Ecotype)</td>
<td>erect growth habit; narrow leaf blades; early flowering; high seed yields; high susceptibility to leaf spot (<em>Drechslera poae</em>); utilized for pastures, conservation purposes and lower maintenance utility turf</td>
<td>Arboretum, Kenblue, South Dakota Certified, Ginger</td>
</tr>
<tr>
<td>Compact</td>
<td>low compact growth habit forming highly attractive turf; good to excellent resistance to leaf spot; slight purple winter color; later spring greenup</td>
<td>Wildwood, Hallmark, Goldstar, Blacksburg, Blackstone</td>
</tr>
<tr>
<td>Compact – America</td>
<td>cultivars with similar growth and performance characteristics to the cultivar America; low compact growth forming highly attractive turf; moderate winter dormancy with some purpling; moderate recovery from summer stress; good resistance to dollar spot (<em>Sclerotinia homoeocarpa</em>), leaf spot, and stripe smut diseases</td>
<td>America, Bedazzled, Brilliant, Delight, Mallard, Arrow, Boutique Showcase, Kingfisher, Barnique, Sonoma, Bordeaux, Unique</td>
</tr>
<tr>
<td>Classification type</td>
<td>Attributes of cultivars in classification type</td>
<td>Example cultivars</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Compact – Midnight</td>
<td>cultivars with similar growth and performance characteristics to the cultivar Midnight; low compact growth forming highly attractive turf; long winter dormancy/late spring greenup; dark green color; susceptibility to powdery mildew (<em>Erysiphe graminis</em>)</td>
<td>Midnight, Midnight II, Arcadia, Quantum Leap, Impact, Beyond, Total Eclipse, Odyssey, NuGlade, Perfection, Tsunami, Awesome</td>
</tr>
<tr>
<td>CELA</td>
<td>cultivars with similar growth and performance characteristics to the cultivars Challenger, Eclipse, Liberty, and Adelphi; early spring greenup; moderate to good stripe smut resistance; good resistance to leaf spot; variable winter color and dormancy</td>
<td>Challenger, Eclipse, Liberty, Adelphi</td>
</tr>
<tr>
<td>Julia</td>
<td>cultivars with similar growth and performance characteristics to the cultivar Julia (a European variety); medium to dark green color; medium leaf texture; good winter color and spring greenup; good wear tolerance; poor resistance to dollar spot</td>
<td>Julia, Ulysses, Pick 453, H92-558, Rampart</td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td>vigorous turf of medium-high density with a deep and extensive rhizome system; tolerance or good recovery from billbug damage; excellent summer stress tolerance; moderate to good leaf spot resistance</td>
<td>Cabernet, Starburst, Appalacian, PST-161, Riverside Park, Valsburg Park</td>
</tr>
<tr>
<td>Shamrock</td>
<td>cultivars with similar growth and performance characteristics to the cultivar Shamrock; moderate resistance to leaf spot and billbugs (<em>Sphenophorus parvulus</em>); high seed yield production</td>
<td>Shamrock, Lakeshore, Moonshine, Mongoose, Champlain</td>
</tr>
</tbody>
</table>
Table 4 cont. Kentucky bluegrass classification types, brief description of types, and example cultivars.

<table>
<thead>
<tr>
<th>Classification type</th>
<th>Attributes of cultivars in classification type</th>
<th>Example cultivars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texas x Kentucky bluegrass hybrids</td>
<td>Intra-specific hybrids between Texas bluegrass and Kentucky bluegrass</td>
<td>Thermal Blue, Longhorn</td>
</tr>
<tr>
<td>Other</td>
<td>A large group of cultivars falling under the category of ‘other turf types’, wide range of characteristics,</td>
<td>Lily, Limerick, Bodacious, Allure,</td>
</tr>
<tr>
<td></td>
<td>requiring further study to accurately classify cultivars or assign to different groups</td>
<td>PST-York Harbor 4, Jewel,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Barzan, Baronie, Rita</td>
</tr>
</tbody>
</table>
et al., 2008; Gunjaca et al., 2008; Smykal et al., 2008; Bonow et al., 2009; Ibanez et al., 2009; Smith et al., 2009). In contrast to morphological characters, molecular markers offer a number of advantages including a nearly unlimited number of characters, high degree of polymorphism, ease of scoring, and they are unaffected by the environment (Lombard et al., 2001; Tommasini et al., 2003; Smykal et al., 2008).

Randomly amplified polymorphic DNA (RAPD) markers have also previously been used to characterize Kentucky bluegrass germplasm (Huff, 2001; Johnson et al., 2002; Curley and Jung, 2004). It would be reasonable to expect a correlation between genetic relatedness as assessed by molecular markers and the previously described Kentucky bluegrass classification system, as the classification of cultivars is based, in part, on pedigree information and morphological traits that are commonly used in PVP analysis. However, the most recent work utilizing RAPD markers by Curley and Jung (2004) reported that a correlation existed only for the BVMG, Compact-‘Midnight’, and possibly common types, with no other obvious patterns emerging that correlated the grouping of cultivars using molecular marker data to the existing Kentucky bluegrass classification scheme. There are a number of likely possibilities that could explain this discrepancy: 1) Performance characteristics and morphological traits are not accounting for all of the genetic variability at the DNA level captured in the RAPD analysis; 2) the RAPD analysis is not accurately reflecting the genetic relatedness of plant material within a classification group; or 3) the seedlings chosen for study in the RAPD analysis did not accurately reflect the apomictic character of Kentucky bluegrass cultivars.

In the current study we used microsatellite, or simple sequence repeat (SSR), markers to study the genetic relationships of 265 Kentucky bluegrass cultivars and
accessions. The specific objectives of this study were: 1) to assess the amount of genetic divergence between commercial Kentucky bluegrass cultivars, experimental selections, plant collections, and hybrids between Kentucky bluegrass and Texas bluegrass (*Poa arachnifera* Torr.), and 2) to assess the correlation between the original cultivar classification system, the previously reported RAPD marker based grouping, and the cultivar grouping using recently described Kentucky bluegrass SSRs (Honig et al., 2010).

**Materials and Methods**

**Plant material**

Two hundred sixty five Kentucky bluegrass cultivars, experimental selections, collections and hybrids were evaluated in this study, and are listed in Table 5. The entries listed in this table are organized according to groupings based on previous versions of the Kentucky bluegrass classification system, with particular emphasis given to the most recent report by Shortell et al. (2009). Single seedlings or single tillers of all entries were transplanted into 48-cell flats (90 cm x 45 cm) and allowed to establish in the greenhouse. The plants were visually screened for apomixis in the greenhouse, and all off-types (aberrant progeny) were discarded. Plants were then established in various spaced-plant nurseries at the Rutgers University Plant Biology and Pathology Research and Extension Farm at Adelphia, NJ, in April 2003 through April 2006 on a well drained Freehold sandy loam (fine-loamy, mixed, mesic, Typic Hapludult). Plants were maintained in the field for a minimum of two years for further apomixis screening. A single transplant of each entry, representing the repeating apomictic mother clone, was
Table 5. Kentucky bluegrass cultivars, experimental selections, collections, and hybrids genotyped by 80 microsatellite markers in the current study†.

<table>
<thead>
<tr>
<th>Classification type</th>
<th>Cultivar or selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Density</td>
<td>HV140‡, Limousine, Julius, NorthStar, Bariris</td>
</tr>
<tr>
<td>BVMG</td>
<td>Baron, Envicta, Goldrush, Abbey, Raven, Marquis</td>
</tr>
<tr>
<td>Common</td>
<td>Kenblue, GO9LM9‡</td>
</tr>
<tr>
<td>Compact</td>
<td>Wildwood, Hallmark, PST-B5-125‡, Bluetastic, PST-B4-246‡, Alpine, Ascot, Ba 82-288‡, Skye, Diva, Goldstar, P-105, Chicago II, Moonshadow, Moonlight, Blackstone, Hampton</td>
</tr>
<tr>
<td>Compact-America</td>
<td>Bedazzled, Royale, Glenmont, PST-604‡, Brilliant, PST-222‡, Mallard, Apollo, Delight, Arrow, Dynamo, Casablanca, Valor, PST-H6-150‡, Boutique, Showcase, Kingfisher, Sonoma, Bordeaux, SR 2284, Langara, PST-B3-170‡, Baroness, Barnique, BAR Pp 0566‡, BAR Pp 0573‡, Unique</td>
</tr>
<tr>
<td>CELA</td>
<td>Jefferson</td>
</tr>
<tr>
<td>Julia</td>
<td>Avalanche, Pick 453‡, Rampart, H92-558‡, Julia</td>
</tr>
<tr>
<td>Mid.-Atlantic</td>
<td>Eagleton, PST-161‡, Cabernet, Appalacian, Starburst</td>
</tr>
</tbody>
</table>
Shamrock Shamrock, Lakeshore, Moonshine, Katie, Brooklawn, Champagne,
Durham, Mongoose, A98-1028‡, Champlain

Texas x Kentucky A03TB775‡, A03TB386‡, A03TB676‡, A03TB256‡, BH006001‡,
bluegrass hybrids BH006002‡, BH006003‡

Other (and unclass-
ified) BlueSapphire, Rita, PST-H5-35‡, Blacksburg II, MonteCarlo,
Voyager II, BlueRidge, Mercury, H92-203‡, A97-1409‡, Royce,
A98-365‡, A96-739‡, A99-3245‡, A99-3122‡, A99-3127‡, Bd99-
2103‡, A03-141‡, Shiraz, A96-363‡, SIA96386‡, Ulysses, A93-
201‡, Misty, DLF769036‡, 99AN53‡, Bartitia, Ba83113‡,
Ba84140‡, Baronie, A03-55‡, Bluemax, PST-YorkHarbor4§,
Baronette, SRX2114‡, A98-407‡, DLF769037‡, A97-1294‡, A00-
1100‡, A99-3116‡, A99-2678‡, A03-66‡, A00-99‡, A99-523‡, A03-
132‡, A97-1303‡, A99-2427‡, A00-430‡, A05-322‡, H02-99xH98-
767‡, A05-332‡, A03-335‡, A05-335‡, A05-314‡, A05-336‡,
HV238‡, Washington, Avid, A95-410§, A00-1254‡, A03-6‡, A03-
77‡, A03-37‡, H03-180xA99-2874‡, A97-1328‡, Limerick, A98-
3320PilicaWarkaPol§, A98-3367RekorvoPol§, H02-
603UlricehannSwe§, H04-390AristoFin§, A04-1268PorkalaFin§,
A04-1427JurbarkasLith§, A04-387PasulprisloRom§, Cheetah,
A98-3384AnimalPKSwe§, A98-3366RekorvoPol§, A94MH94‡,
H02-608UlricehannSwe§, H07-697NorrkopingSwe§, A04-
1357PargasFin§, A04-1474SiauliaiLith§, A04-1569MoletLith§,
A04-1547SpitrenaiLith§, A04-1470SiauliaiLith§, A04-
Kentucky bluegrass entries are grouped according to historical classification typing (pedigree, performance characteristics, and morphological measurements).

Denotes experimental selection.

Denotes collection (Pol = Poland; Swe = Sweden; Fin = Finland; Lith = Lithuania; Rom = Romania; Uz = Uzbekistan; Fra = France; Sp = Spain; Ger = Germany; Ch = China).
established in the greenhouse and maintained for DNA extraction in the laboratory. An additional single entry of a wild ecotype of annual bluegrass (*Poa annua* L.), collected from the Rutgers University Plant Biology and Pathology Research and Extension Farm at Adelphia, NJ, was included in this study as an outgroup for phylogenetic analysis.

**Microsatellite development and genotyping**

The development of 88 Kentucky bluegrass microsatellite markers, characteristics of the SSRs, and the primer sequences for those markers was previously described by Honig et al., (2010). Microsatellite markers NJPpGA6, NJPpGA931, NJPpGA937, NJPpGA993, NJPpGA1117, NJPpGT8, NJPpGT12, and NJPpGT9318 were not used in the current analysis because they produced more than 5% missing data (7%, 16%, 16%, 6%, 13%, 15%, 14%, and 10%, respectively) in the 266 entries. DNA was isolated from the 266 entries using Sigma GenElute Plant Genomic DNA Miniprep Kit (St. Louis, MO) according to the manufacturer’s instructions. For genotyping, amplifications of microsatellite loci with fluorescently labeled PCR primers were performed in 13 μl reactions using 50 ng of Kentucky bluegrass DNA, 1X Ramp-Taq PCR buffer (Denville Scientific, Metuchen, NJ), 2 mM MgCl₂, 0.25 mM each dNTP (Denville Scientific), 0.5 U Ramp-Taq DNA polymerase (Denville Scientific), 0.5 pmol forward primer with 18 bp M13(-21) addition at the 5’-end (Schuelke, 2000), 1 pmol reverse primer, and 1 pmol fluorescent dye-labeled 18bp M13(-21) primer (FAM, NED, PET or VIC). Thermal-cycling parameters were 94 °C for 5 min., followed by 30 cycles of 94 °C for 30 s, 55 °C for 45 s, 72 °C for 45 s, followed by 20 cycles of 94 °C for 30 s, 53 °C for 45 s, 72°C for 45 s, ending with a final extension of 72 °C for 10 min. PCR reaction products were
analyzed on an ABI 3130xl Genetic Analyzer and sized using Genemapper 3.7 (Applied Biosystems, Foster City, CA) and LIZ 500(-250) size standard.

Although microsatellites can generate codominant data, problems can arise during the identification of polyploid genotypes because it is difficult to determine the number of copies of an allele in heterozygotes (Saltonstall, 2003; Markwith et al., 2006; Liao et al., 2008). Thus, banding patterns observed at particular loci are referred to as ‘allele phenotypes’ (Becher et al., 2000). The individual alleles of the SSRs utilized in the current study of 265 polyploid Kentucky bluegrasses and one annual bluegrass ecotype were treated as dominant markers, where the banding phenotypes were scored as band absence (0) or presence (1) (binary data matrix).

Data analysis

Genetic distance and UPGMA clustering

Only strong and reliably amplified SSR marker alleles were scored and assembled into a binary data matrix. Pairwise comparisons of the proportion of shared alleles between individual genotypes (plants) were determined by the Jaccard (1908) coefficient, where $J = \frac{\text{matching fragments}}{\text{matching} + \text{polymorphic fragments}}$. The resulting genetic similarity coefficient was then transformed into a distance matrix ($d = 1 - \text{Jaccard similarity}$), averaged over 1000 random draws from the interval of the minimum and maximum Jaccard distance. Cluster analysis was generated from the distance matrix by the unweighted pair group method using the arithmetic averages (UPGMA) algorithm. All calculations were performed by the computer program FAMD 1.23 (Schluter and
Harris, 2006). The UPGMA tree was visualized using the computer program Dendroscope 2.6.1 (Huson et al., 2007).

**Model-based clustering**

Model based clustering analysis employing a Bayesian algorithm was applied to infer the genetic structure and to define the number of clusters (Kentucky bluegrass types/populations) in the dataset using the computer program STRUCTURE 2.3.3 (Pritchard et al., 2000; Falush et al., 2003). In this analysis, genotyped individuals are allocated to a predetermined number of populations/clusters ($K$), where ($K$) is chosen in advance and can be varied across different runs. Plants can have membership in several clusters, with membership coefficient equaling 1.0 across clusters. The method employs a Markov chain Monte Carlo algorithm to estimate the allele frequencies in each of the ($K$) populations/clusters and, for each individual, the proportion of its genome derived from each population/cluster ($q_k$). We assumed that all loci were independent and in linkage equilibrium. A no-admixture ancestry model was used and allele frequencies were correlated, with a burnin length of 10000 iterations followed by 100000 run iterations at each ($K$). For other settings, program defaults were used, and no prior information was used to define the clusters ($K$). ($K$) values were tested from 2 through 25, with three replicate runs at each ($K$). The wild ecotype of *P. annua* was excluded from this analysis.
**Principle coordinate analysis**

Principle coordinate analysis (PCO), which provides estimates of genetic similarity between individuals, has been used as an alternative method to represent inter-individual and inter-group relationships. The PCO multivariate method was chosen as a complement to the UPGMA cluster analysis method, because cluster analysis is more accurate for closely related individuals, whereas PCO is more accurate in assessing distances among major groups (Hauser and Crovello, 1982; Sun et al., 2001; Taran et al., 2004). The previously described SSR allele binary data matrix was used to calculate a pairwise genetic distance matrix between individuals using the computer program GenAlEx 6 (Peakall and Smouse, 2006). The pairwise genetic distance was used in a PCO analysis in GenAlEx 6 to validate and further define naturally occurring genetic clusters. The first three eigenvalues of the PCO analysis were visualized using the computer program SigmaPlot 11.0 (Systat Software Inc., San Jose, CA, USA).

**Kentucky bluegrass classification type revision**

Based on a combination of known pedigrees, the results of the UPGMA cluster analysis, the results of the model based clustering analysis, and the PCO analysis, we assigned all 266 individuals of the current study into 12 revised Kentucky bluegrass classification types and one outgroup (*P. annua*). The 265 Kentucky bluegrass cultivars, experimental selections, collections, and hybrids are listed in Table 6 according to this newly revised classification scheme. This newly revised classification scheme is, therefore, a representation of genetic relatedness based on known pedigrees and the current microsatellite data, and differs from previous classification systems that were
Table 6. Revised Kentucky bluegrass classification system based on a combination of pedigree, UPGMA cluster analysis, model-based clustering analysis, and principle coordinate analysis (PCO) of microsatellite data.

<table>
<thead>
<tr>
<th>Classification type</th>
<th>Cultivar or selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>BVMG</td>
<td>Baron, Enviicta, Goldrush, Misty, BH006000, Abbey, Raven, Marquis, DLF769036, 99AN53, Bartitia</td>
</tr>
<tr>
<td>Compact</td>
<td>Alpine, Ba83-113, Ba84-140, Ba 82-288, BlueSapphire, Moonshadow, Baronie, A03-55, A03TB775, BH006003, Hallmark, PST-B5-125, PST-YorkHarbor4, Blacksburg II PST-B4-246, Ascot, Baronette, SRX2114, PST-H5-35, PST-B3-170, Blackstone, Rita, Hampton, Wildwood,</td>
</tr>
</tbody>
</table>
P-105 related  P-105, HV238†, Avalanche, Jefferson, Champlain, Washington, 
Avid, A95-410†
Mid.-Atlantic  A03TB386†, PST-161†, Cabernet, Appalacian, Starburst, 
A00-1254, A03-6†, A03-77†, H03-180xA99-2874†, A97-1328†, 
A03-37†
EurAsian  Limerick, HV140†, A98-3320PilicaWarkaPol‡, A98-
3367RekorvoPol‡, H02-603UltrehamnSwe‡, H04-390AristoFin‡, 
A04-1268PorkalaFin‡, A04-1427JurbarkasLith‡, A04-
387PasulprisloRom‡, Cheetah, A98-3384AnimalPKSwe‡, A98-
3366RekorvoPol‡, A94MH94‡, H02-608UltrehamnSwe‡, H07-
697NorrkopingSwe‡, A04-1357PargasFin‡, A04-1474SiauliaiLith‡, 
A04-1569MoletLith‡, A04-1547SpitrenaiLith‡, A04-
1470SiauliaiLith‡, A04-1484KursenaiLith‡, H04-535KorpoFin‡, 
Bodacious, Boomerang, PpH7929†, PpH7907†, SRX27921†, 
Chelsea, Bluenight, DLF769032†, GO9LM9†, NorthStar, 
A03TB676†, U2998Uz‡, A98-3369TeologPol‡, A04-
1384AristoFin‡, H04-376KaisterFin‡, A04-1354PetsmoFin‡, A04-
1272PorkalaFin‡, A05-847ColdePortFra‡, A04-1342KustaviFin‡, 
A04-1347LepainenFin‡, Eagleton, Jewel, DLF769034†, A98-183†, 
Kenblue, A96-742†, A98-3297OgrodzieniecPol‡, H01-
847SomcutaMareRom‡, H99-1722Uz‡, A98-3322ZyrardowPol‡, 
A983323ZychlinPol‡, H94-305‡, H01-804ComarnicRom‡, A04-
1415KrudonisLith‡, A04-1423BuktaLith‡, A04-383PojorataRom‡,
A04-1271PorkalaFin\(^\ddagger\), A04-394PetroojaniRom\(^\ddagger\), H01-894HongyanCh\(^\ddagger\), Barzan

Julia

H01-912FrankfurtGer\(^\ddagger\), A04-1283HankoFin\(^\ddagger\), A05-894VillaboneSp\(^\ddagger\), Ulysses, Pick 453\(^\dagger\), Rampart, CVB20631\(^\dagger\), H92-558\(^\dagger\), Julia, Baritone, Bariris

Shamrock

A97-857\(^\dagger\), Shamrock, Lakeshore, Moonshine, Brooklawn, Champagne, Mongoose, A98-1028\(^\dagger\), A98-948Aries, A01-701\(^\dagger\), A93-201\(^\dagger\), A03TB256\(^\dagger\), Bd0384\(^\dagger\), A97-1799\(^\dagger\), A05-313\(^\dagger\)

Sydsport related

SRXQG245\(^\dagger\), A04-1315FinstromFin\(^\ddagger\), H04-389SodderRuderFin\(^\ddagger\), Lily, Coventry, PST-10879\(^\dagger\), B543\(^\dagger\), B545\(^\dagger\), B5144\(^\dagger\), Fairfax, Chateau, Allure, Serene

Limousine related

Limousine, Julius, SRX26351\(^\dagger\), B4128A\(^\dagger\), A00-4083PoriFin\(^\ddagger\), PpH6366\(^\dagger\), Freedom II

Compact-Midnight

Skye, Chicago II, Midnight, Midnight II, Quantum Leap, Arcadia, Impact, Total Eclipse, Odyssey, NuGlade, Perfection, Tsunami, Ginney, Courtyard, Alexa, FrontPage, Blue Velvet, Everest, Awesome, Excursion, Freedom III, EverGlade, NuDestiny, Barrister, Beyond, Rugby II, Award, Liberator, Bluestone

Other

Wellington, NA-K992\(^\dagger\)

\(^\dagger\)Denotes experimental selection

\(^\ddagger\)Denotes collection (Pol = Poland; Swe = Sweden; Fin = Finland; Lith = Lithuania; Rom = Romania; Uz = Uzbekistan; Fra = France; Sp = Spain; Ger = Germany; Ch = China)
based on pedigree, morphological characters, and cultivar performance data. The PCO analysis was color coded according to the revised classification types, listed in Table 6, for easier visualization of the PCO data.

**Analysis of molecular variation (AMOVA)**

The 12 revised classification types, described in Table 6, were treated as populations in an AMOVA analysis, performed in GenAlEx 6, to examine the distribution of variation among and within populations (classification types), and to assess the interpopulation pairwise genetic distance ($\Phi_{ST}$). Statistical significance was tested by random permutation, with the number of permutations set to 999. Tests of statistical significance of the AMOVA were used to validate the proposed 12 revised classification types. The wild ecotype of *P. annua* was excluded from this analysis.

**Results**

**Microsatellite markers**

The 80 SSR markers used in the current study produced 3353 alleles in the 266 entries. The number of alleles for individual SSRs ranged from 5 to 85, with an average of 41.9 alleles per SSR. Specific details about individual SSR markers can be found in Honig et al. (2010).

**UPGMA Clustering**

The results of the UPGMA clustering analysis are presented in Fig. 1. Cultivars and accessions are color coded according to a combination of known pedigrees and the
Fig. 1. Phylogeny of Kentucky bluegrass cultivars, experimental selections, collections, and hybrids. Diagram is a UPGMA cladogram based on the average Jaccard distance of SSR marker alleles between individuals, color coded according to known pedigrees. Black color indicates unknown pedigree or individual that may be clustering in a clade that differs from known pedigree.
current SRR classification type assignment. The UPGMA clustering analysis clearly groups the Kentucky bluegrass entries into several distinct classification types. Very strong support exists for the delineation of the traditional Kentucky bluegrass types Shamrock, Compact-Midnight, BVMG, and Compact-America. The majority of European, as well as a few Asian, plant collections are tightly clustered in the lower left quadrant of Fig. 1. This is strong justification for a newly defined type referred to as the Eurasian type. The traditional Julia type, originally comprised of European cultivars, is closely grouped within the main clustering of the newly proposed Eurasian type. The cultivars ‘Allure’, ‘Fairfax’, and ‘Serene’, as well as the experimental selections B545, and B543 are all tightly clustered in the upper left quadrant of Fig. 1. The pedigrees of these entries can all trace, in part, back to the cultivar ‘Sydsport’ (Crystal A. Rose-Fricke – personal communication), providing strong justification for a newly defined type referred to as Sydsport-related. Additional support is likely for grouping the traditional Mid-Atlantic type, and new groups defined as Limousine-related and P105-related. Cultivars that had previously been referred to as Compact type cultivars are split into two smaller subclusters that encompass the Compact-America, P-105-related, and Mid-Atlantic types.

The main clustering of Eurasian collections and cultivars are basal, being most closely related to the outgroup, *P. annua*. The remainder of the groups and cultivars, which were predominantly developed from US breeding programs, appear to be derived from the more basal Eurasian cultivars. Within the large grouping of predominantly US germplasm, Shamrock, Sydsport-related, Limousine-related, and Compact-Midnight type cultivars are more closely related and are aligned with the BVMG type cultivars.
Another large subgroup within the US germplasm consists of the Compact, Compact-America, P-105-related, and Mid-Atlantic types.

All cultivars/accessions were uniquely identified with the current panel of 80 Kentucky bluegrass SSR markers. The majority of the individual cultivars/accessions were separated by a Jaccard’s genetic distance greater than 0.10 (Fig. 1). Exceptions included a few cultivars/accessions in the Compact-America type and several cultivars in the Compact-Midnight type.

**Model-based clustering**

The results of the model-based clustering (STRUCTURE 2.3.3) analysis are presented in Fig. 2. Color coding is separated based on the groupings produced from the STRUCTURE output. Groupings of cultivars from the model-based clustering analysis most closely resembled the groupings delineated in the UPGMA cladogram at \( K = 18 \). At all other values of \( K \) (2-17 and 19-25), groups were composed of individuals belonging to clusters that differed substantially from known pedigrees and the UPGMA cladogram (data not shown), making it difficult to identify the underlying classification criterion. At \( K = 18 \), clusters identified by STRUCTURE were very closely matched to clusters in the UPGMA cladogram for Shamrock, Sydsport-related, Limosine-related, Compact-Midnight, BVMG, Compact-America, P-105-related, Mid-Atlantic, and Julia cultivar types. Compact cultivars were predominantly found in two groups in the STRUCTURE analysis: one group composed of a mixture of Compact cultivars and cultivars previously defined as belonging to the Other type (Table 3); and the other composed of a mixture of Compact cultivars and Eurasian cultivars/accessions. The
Fig. 2. Estimated population (classification type) structure for Kentucky bluegrass cultivars, experimental selections, collections, and hybrids. Each individual is represented by a vertical colored line, which can be partitioned into \( K \) segments that represent the individual’s estimated membership fractions in \( K \) clusters. Groupings of cultivars from this model-based clustering analysis most closely resembled the groupings delineated in the UPGMA cladogram at \( K = 18 \).
Fig. 2. Cont.
Eurasian types were divided into eight clusters, including a Julia cluster. The eight Eurasian clusters from the STRUCTURE analysis are a departure from the clustering represented in the UPGMA cladogram, however, taken all together the eight Eurasian clusters closely match the main group of Eurasian cultivars/accessions in Fig. 1. An additional cluster identified in the STRUCTURE analysis that was not present in the UPGMA cladogram was composed of the cultivars ‘Chicago II’ (Brede, 2004a), ‘Limerick’, ‘Chelsea’, and ‘Skye’.

**Principle coordinate analysis**

For the PCO analysis (Fig. 3) PCO1, PCO2, and PCO3 explained 40.6%, 21.2%, and 12.7% of the total variation, respectively. The results of the PCO analysis were generally consistent with those obtained through the UPGMA cluster analysis and model-based cluster analysis. Distinct clusters of cultivar groupings that are similar to both the UPGMA cluster analysis and model-based clustering are easily identifiable in the PCO graph. The Eurasian, Julia, Compact-America, Limousine related, Compact, P-105 related, and Mid-Atlantic types form a larger cluster in the center-left of the diagram. The Compact-Midnight type forms a very distinct cluster to the right of the main grouping, while the BVMG, Shamrock, and Sydsport related types form distinct clusters above the main grouping.

**Kentucky bluegrass classification type revision**

The combination of pedigree information, UPGMA cluster analysis, model based clustering analysis, and PCO analysis all support the proposed revised Kentucky
Principle Coordinate Analysis of Kentucky bluegrass cultivars, experimental selections, collections, and hybrids measured by 80 microsatellite markers

Fig. 3. Principle coordinate analysis (PCO) of Kentucky bluegrass cultivars, experimental selections, collections, and hybrids based on microsatellite allele data. PCO1, PCO2, and PCO3 accounted for 40.6%, 21.2%, and 12.7% of the variation, respectively.
bluegrass classification system outlined in Table 6. With just a few exceptions, the
UPGMA clustering analysis and model based clustering analysis at \( (K) = 18 \) produce
similar grouping results for the Shamrock, Sydsport-related, Limousine-related,
Compact-Midnight, BVMG, Compact-America, P-105-related, Mid-Atlantic, and Julia
classification types. Additionally, the eight clusters (including the Julia type) composed
of predominantly Eurasian collections revealed in the model based clustering, closely
match the main grouping of Eurasian cultivars in the UPGMA analysis. The PCO
multivariate method produced very similar grouping patterns relative to both the
UPGMA and model based clustering analyses. Taken together, the results of these three
analyses indicate that the revised 12 classification types described in Table 6 are a
reasonable representation of the genetic relationships among Kentucky bluegrass
cultivars.

**Analysis of molecular variance (AMOVA)**

The results of the AMOVA are presented in Table 7. The AMOVA was
carried out treating the newly revised 12 classification types (Table 6) as populations.
The AMOVA was, in part, meant to serve as a validation step for the newly revised
classification scheme.

The majority of the SSR variation (70%) observed among the 265 individuals
from the 12 populations was accounted for by within population variance, although a
significant portion (30%) was attributable to differences between populations (Table 7).
Within population variation was spread relatively uniformly among the various
Table 7. Analysis of molecular variance (AMOVA) of 265 Kentucky bluegrass individuals from 12 different populations (classification types) based on 80 microsatellite markers.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of squares</th>
<th>Variance components†</th>
<th>P-value‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among population</td>
<td>11</td>
<td>19383.26</td>
<td>77.91 (30%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within populations</td>
<td>253</td>
<td>46945.03</td>
<td>185.55 (70%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of squares within populations</th>
<th>Variance components within populations†</th>
<th>P-value‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>BVMG</td>
<td>10</td>
<td>1625.09</td>
<td>162.50 (7.88%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Compact</td>
<td>23</td>
<td>6006.08</td>
<td>261.13 (12.66%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Compact-America</td>
<td>70</td>
<td>10699.87</td>
<td>152.85 (7.41%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>P-105 related</td>
<td>7</td>
<td>1265.00</td>
<td>180.71 (8.76%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td>10</td>
<td>1699.18</td>
<td>169.92 (8.24%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>EurAsian</td>
<td>61</td>
<td>17144.34</td>
<td>281.05 (13.63%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Julia</td>
<td>10</td>
<td>1672.36</td>
<td>167.23 (8.11%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>203.00</td>
<td>203.00 (9.84%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Shamrock</td>
<td>14</td>
<td>2423.80</td>
<td>173.27 (8.40%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sydsport related</td>
<td>12</td>
<td>1618.85</td>
<td>134.90 (6.54%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Limousine related</td>
<td>6</td>
<td>656.29</td>
<td>109.38 (5.30%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Comp.-Midnight</td>
<td>29</td>
<td>1931.17</td>
<td>66.60 (3.23%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

†Value in parenthesis indicates the percentage of the total.
‡Probability computed by nonparametric procedures from 999 data permutations.
classification types (Table 7), however, the Compact-Midnight type had the lowest within population variation (Table 7). Pair-wise $\Phi_{ST}$ values derived from the AMOVA highlighted a large number of differences between classification types, when individual pairs of classification types were compared (Table 8). Two pair-wise $\Phi_{ST}$ comparisons, between the newly defined Other type and the Eurasian type, as well as the newly defined Other type and the Compact type, were not significant ($P > 0.05$), but all other pair-wise $\Phi_{ST}$ comparisons were significant at $P < 0.05$ (Table 8). This indicates that the vast majority of classification types are differentiated from other classification types, the $\Phi_{ST}$ distance values are not random, and the AMOVA supports the revised classification scheme presented in Table 6.

Discussion

Validation and revision of Kentucky bluegrass classification types

The UPGMA depiction (Fig. 1) of the genetic relatedness of Kentucky bluegrass cultivars, experimental selections, collections, and hybrids based on SSR data represents a cogent argument for the grouping of Kentucky bluegrass entries into several distinct classification types. With only a few exceptions, model based clustering (Fig. 2) and PCO analysis (Fig. 3) corroborate the grouping scenario depicted in Fig. 1. These three separate data analyses are in strong agreement with known pedigree information, and therefore, support a revision of previous Kentucky bluegrass classification systems resulting in a newly proposed classification scheme outlined in Table 6. The results of the AMOVA (Tables 7 and 8) conclusively validate the revised classification system,
Table 8. Pair-wise $\Phi_{st}$ values calculated by AMOVA illustrating differences between populations (classification types) of Kentucky bluegrass ($\Phi_{st}$ values are given below the diagonal, and the associated $P$ values are given above the diagonal).

<table>
<thead>
<tr>
<th></th>
<th>BVMG</th>
<th>Compact</th>
<th>Comp.-America</th>
<th>P-105 related</th>
<th>Mid.-At.</th>
<th>EurAsian</th>
<th>Julia</th>
<th>Other</th>
<th>Shamrock</th>
<th>Sydsport related</th>
<th>Limousine related</th>
<th>Comp.-Midnight</th>
</tr>
</thead>
<tbody>
<tr>
<td>BVMG</td>
<td>—</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.016</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Compact</td>
<td>0.226</td>
<td>—</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.079</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Comp.-America</td>
<td>0.416</td>
<td>0.188</td>
<td>—</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.025</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>P-105 related</td>
<td>0.325</td>
<td>0.128</td>
<td>0.333</td>
<td>—</td>
<td>0.001</td>
<td>0.001</td>
<td>0.021</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td>0.373</td>
<td>0.215</td>
<td>0.368</td>
<td>0.306</td>
<td>—</td>
<td>0.001</td>
<td>0.017</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>EurAsian</td>
<td>0.144</td>
<td>0.082</td>
<td>0.228</td>
<td>0.110</td>
<td>0.157</td>
<td>—</td>
<td>0.078</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Julia</td>
<td>0.372</td>
<td>0.213</td>
<td>0.407</td>
<td>0.301</td>
<td>0.354</td>
<td>0.146</td>
<td>—</td>
<td>0.011</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Other</td>
<td>0.362</td>
<td>0.102</td>
<td>0.366</td>
<td>0.281</td>
<td>0.343</td>
<td>0.060</td>
<td>0.360</td>
<td>—</td>
<td>0.002</td>
<td>0.010</td>
<td>0.024</td>
<td>0.005</td>
</tr>
<tr>
<td>Shamrock</td>
<td>0.326</td>
<td>0.193</td>
<td>0.342</td>
<td>0.286</td>
<td>0.354</td>
<td>0.145</td>
<td>0.364</td>
<td>0.312</td>
<td>—</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Sydsport related</td>
<td>0.323</td>
<td>0.243</td>
<td>0.394</td>
<td>0.378</td>
<td>0.417</td>
<td>0.171</td>
<td>0.411</td>
<td>0.432</td>
<td>0.291</td>
<td>—</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Limousine related</td>
<td>0.428</td>
<td>0.237</td>
<td>0.421</td>
<td>0.379</td>
<td>0.416</td>
<td>0.166</td>
<td>0.440</td>
<td>0.488</td>
<td>0.386</td>
<td>0.464</td>
<td>—</td>
<td>0.001</td>
</tr>
<tr>
<td>Comp.-Midnight</td>
<td>0.542</td>
<td>0.402</td>
<td>0.526</td>
<td>0.568</td>
<td>0.586</td>
<td>0.277</td>
<td>0.596</td>
<td>0.677</td>
<td>0.532</td>
<td>0.602</td>
<td>0.645</td>
<td>—</td>
</tr>
</tbody>
</table>
indicating that the proposed types are composed of distinct populations or genetically related classification groups.

A significant improvement to the previous classification system is the near elimination of almost all entries in the current study from the Other type (Table 5 versus Table 6), placing the majority of these cultivars and accessions into previously existing classification types or newly proposed types based on SSR relatedness. The entries ‘Wellington’ and NA-K992 remain classified as Other types because the UPGMA analysis shows these entries to be closely aligned with the outgroup and possibly grouping with the Eurasian entries, yet according to pedigree information Wellington was a naturalized ecotypic selection from Oregon (Brentano, 2004) and NA-K992 was an aberrant selection out of the Compact type cultivar ‘Wildwood’ (Thomas E. Brentano – personal communication). Interestingly, the AMOVA results seem to partially support this pedigree information in that one of the only two non-significant pair-wise ΦST comparisons occurs between the revised Other type (NA-K992 and Wellington) and the revised Compact type (possibly attributed to the relationship of NA-K992 to Wildwood) (Table 8). Additional molecular marker, agronomic or performance data will likely be needed to place these entries into existing classification types.

The UPGMA cluster analysis, model based clustering, and PCO analysis provide strong support for the traditional classification types of Shamrock, Compact-Midnight, BVMG, Compact-America, and Julia, as well as new classification types referred to as Sydsport related and Eurasian (with the Julia type now being a subtype within the Eurasian type). Additional support is likely for new types referred to as Limousine related and P-105 related, as well as the traditional Mid-Atlantic type, however, these
classification types have limited numbers of entries in the current study. SSR genotyping of additional cultivars and accessions in the latter three types would help to support or refute the validity of these classification groups.

Unlike most of the classification types in the current study, the Compact type was split among multiple clades in the UPGMA analysis and multiple (K) groupings in the model based clustering analysis. This is interesting in that the current SSR data indicates that the Compact type may not represent a distinct genetic group. The original description of Compact type cultivars included cultivars that exhibited “low compact growth habit, forming a highly attractive turf” (Bara et al., 1993). While pedigree information was considered when assigning cultivars to this grouping, a number of genetically divergent cultivars may have originally been included in the Compact group based simply on a common low, compact growth habit. SSR genotyping of additional cultivars and accessions in this classification type would help clarify whether or not this group needs to be split into two or more revised types.

Contrary to the RAPD marker data presented by Curly and Jung (2004), the SSR data in the current study showed a strong correlation between genetic relatedness as assessed by molecular markers and the previously described Kentucky bluegrass cultivar classification system based on pedigree information, morphological traits, and cultivar performance characteristics (with due consideration given to the current proposed revisions). We believe this correlation is to be expected, as the original classification system was based in part on known pedigrees of Kentucky bluegrass cultivars and accessions. The discrepancy between the RAPD marker data and the SSR marker data is likely the result of general inherent limitations associated with RAPD markers. Although
initially attractive as a marker system when first described (simple utility, low cost, and high return), RAPD markers have since been found to have significant disadvantages including: 1) low stringency PCR conditions (low annealing temperature and short-length primer), which can increase the likelihood of nonspecific binding and primer mismatch (Williams et al., 1990; Tyler et al., 1997; Perez et al., 1998; Mueller and Wolfenbarger, 1999; Zhang et al., 2006) resulting in the inability of RAPD marker assays to distinguish between true and artifactual PCR products (Williams et al., 1990; Ellsworth et al., 1993; Lamboy, 1994a; Lamboy, 1994b; Tyler et al., 1997; Perez et al., 1998; Mueller and Wolfenbarger, 1999); 2) high intra- and inter-laboratory variation even under controlled conditions (Penner et al., 1993; Jones et al., 1997; Perez et al., 1998); and 3) limitations related to sequence homology where bands of apparently identical molecular weight in different individuals are assumed to be the same homologous DNA fragment, when in fact PCR bands of the same size may represent a different genetic location (locus) (Devos and Gale, 1992; Clark and Lanigan, 1993; van de Zande and Bijlsma, 1995; Karp et al., 1996; Karp et al., 1997; Gupta et al., 1999; Kuras et al., 2004; Semagn et al., 2006). Additionally, once amplified fragment length polymorphisms (AFLPs) (Vos et al., 1995), SSR markers, and other marker systems were developed, numerous authors have reported that RAPD markers are not as informative or efficient for genetic diversity studies in direct comparison to other marker systems (Mueller and Wolfenbarger, 1999; Yue et al., 2002; Agrama and Tuinstra, 2003; Belaj et al., 2003; Garcia et al., 2004; Kjolner et al., 2004; Kuras et al., 2004; Sobotka et al., 2004; Gallego et al., 2005; Jeung et al., 2005; Yu et al., 2005; Behera et al., 2008). Lastly, a possible contributing factor associated with the lack of correlation between the RAPD marker work of Curley and Jung (2004) and
the traditional Kentucky bluegrass classification system could have been related to the plant material sampling techniques. Curley and Jung (2004) sampled three random seeds of each entry in their study, which may have included ‘off-types’ that were not representative of the cultivar, while the current study using SSR markers included two levels of apomixis screening (at the seedling stage and the mature plant stage) to ensure that the samples chosen for DNA analysis represented the repeating apomictic maternal plant.

Genetic relationships among classification types

The main grouping of Eurasian accessions (lower left quadrant of Fig. 1), being most closely related to the outgroup, *P. annua*, is basal to the remainder of the Kentucky bluegrass cultivars, experimental selections, and hybrids in the current study, many of which originated from US breeding programs. This grouping scenario fits with observations that Kentucky bluegrass is native to the old world, being distributed naturally throughout the temperate and cooler regions of Europe and Asia (Bashaw and Funk, 1987), and with the proposed center of origin of the genus Poa, which based on morphological, cytological, and species diversity is considered to be Eurasia (Huff, 2003). Additionally, these Eurasian collections are genetically diverse, as indicated by the deep nodes in the UPGMA analysis (Fig. 1) and the highest within population variance in the AMOVA (Table 7), and therefore, will be important sources of new genetic variation if incorporated into US breeding programs.

The next most basal large delineation consists of the Shamrock, Sydsport related, Limousine related, Compact-Midnight, and BVMG types (Fig. 1). This large lineage
includes a limited number of Eurasian collections aligned with the Sydsport related, BVMG, and Limousine related types. The Eurasian collections aligned with these types likely follows pedigree information, as ‘Limousine’ originated in Germany (Alderson and Sharp, 1994), Sydsport originated in Sweden (Alderson and Sharp, 1994), and ‘Baron’ (Hurley and Ghijsen, 1980) (BVMG type) originated in Holland. It will be interesting to determine if the Eurasian collections aligned with the BVMG type are susceptible to the fungal pathogen stripe smut (Ustilago striiformis (West.) Niessl.), as this has historically been a defining characteristic of this classification group. The cultivar ‘Sydsport’ was one of the earlier improved Kentucky bluegrass cultivars utilized in the US, likely accounting for the placement of the new Sydsport related type in this quadrant of the UPGMA diagram which includes both US and European germplasm (Fig. 1). The cultivar ‘Shamrock’ (Bailey et al., 1995) (type name for the Shamrock type), was a single plant progeny from A80-336, a derivative of ‘Glade’ (Jacklin et al., 1977), pollinated in a polycross that included the cultivar Sydsport, which likely explains the proximity of the Shamrock and Sydsport related types in the UPGMA analysis (Fig. 1). The cultivar ‘Midnight’ (Meyer et al., 1984) originated from the progeny of F64-603, a selection made from an old lawn in Washington, D.C., crossed with the cultivar Glade (Alderson and Sharp, 1994). The fact that both Shamrock types and Midnight types trace back to Glade, or derivatives of Glade, is a likely explanation for the placement of both the Compact-Midnight and Shamrock types in this major lineage (Fig. 1). Additionally, many of the newer Compact-Midnight type cultivars are progeny from crosses between Midnight and Limousine (e.g. Brede, 2001; 2003; 2004b; 2004c; 2006), which explains the proximity of these two types in Fig. 1. A final interesting observation associated with this large
grouping of classification types is the placement of the Compact-Midnight type within this major lineage, as opposed to the lineage that includes the Compact and Compact-America types (Fig. 1). Earlier classification systems, based on morphology and performance characteristics, implied that the Compact-Midnight type was related to the Compact and Compact-America classification types based on the inclusion of the name “Compact-”. The current SSR data disputes this claim based on the location of the Compact-Midnight type in the UPGMA analysis (Fig. 1), as well as a strikingly unique placement in the PCO analysis (Fig. 3). Additionally, pair-wise $\Phi_{ST}$ distance values calculated by AMOVA (Table 8) indicate that the Compact-Midnight type is one of the most distinct classification groups relative to all other revised groups in the current study. Taken together, these findings suggest that the Compact-Midnight type should from this point forward be named the Midnight type, as the “Compact-” designation for this specific type seems to refer only to growth habit and not genetic relationships among groups.

The final large lineage is comprised of the Compact, Compact-America, P-105 related, and Mid-Atlantic types. This lineage is almost exclusively comprised of cultivars and accessions from US breeding programs, with the possible exceptions of ‘Baronie’ and ‘Baronette’ (unknown pedigrees). Within this lineage, it will be interesting to continue to refine the classification of the revised Compact type, as this type appears to be divided into two genetically distinct groups: one group composed of the cultivars ‘Alpine’, ‘Ascot’, ‘Bluesapphire’, ‘Blackstone’ (Rose-Fricker et al., 2002), and ‘Hampton’; and the other composed of cultivars including Wildwood, ‘Hallmark’, ‘Rita’, and ‘Blacksburg II’ (Fig. 1). Unfortunately, pedigree information does not adequately
explain this division of Compact types for the entries in the current study, so it will be necessary to continue to genotype additional entries in this classification type to properly divide this type into two or more revised groups. The Compact-America type appears to be more closely related to the split group of Compact type cultivars that includes Wildwood, Hallmark, Rita, and Blacksburg II (Fig. 1). Genotyping of additional entries in the Compact type may provide a definitive pedigree connection to the Compact-America type, although a possible explanation for the current grouping association exists through the cultivar ‘Moonshadow’ (Bonos et al., 2005). In the current analysis, Moonshadow is the basal entry of the Compact-America type (Fig. 1), and was a single plant selected from the progeny of an aberrant derivative of H86-974, pollinated by C-74. H86-974 was an apomictic F1 hybrid selected from the progeny of the cross A-25 x ‘Blacksburg’. Blacksburg was an early Compact type cultivar, and C-74, or closely related cultivars such as ‘Unique’ (Rose-Fricker et al., 1999), have been widely used in crosses of cultivars in the Compact-America classification type. Blacksburg was not included in the current study, however, Blacksburg II, a repeating aberrant from Blacksburg, is in the split Compact group most closely related to the Compact-America type (Fig. 1). Genotyping of the cultivar Blacksburg would help confirm this association. The cultivar ‘Cabernet’ (Bonos et al., 2004) originated from the polycross of the experimental selection RSP pollinated by plants of ‘P-105’ (Hurley et al., 2000), Rita, and C-74, which likely explains at least one the associations between the Mid-Atlantic type and the P-105 related, Compact-America, or Compact types in the UPGMA analysis (Fig. 1). Genotyping of additional P-105 related and Mid-Atlantic types will likely further resolve broad genetic relationships within this major lineage.
Genetic relationships among cultivars/accessions

In addition to the broad revision of the Kentucky bluegrass classification system, there are a number of interesting observations or changes at the cultivar/accession level. Some brief examples include:

1. All cultivars, experimental selections, collections, and hybrids were uniquely identified by the current set of SSR markers. Most entries were separated by a Jaccard’s distance greater than 0.10 (Fig. 1). Notable exceptions included a few entries in the Compact-America type and several entries in the Midnight type (Fig. 1). Additionally, the AMOVA results show that the Midnight type has the lowest within population variance (Table 7), indicating that the Midnight type cultivars as a group have lower genetic variability in comparison to the remaining revised classification types. This similarity in many of the individual Midnight type cultivars confirms their suspected close relationship based on pedigree, morphological characteristics, and performance characteristics (Shortell et al., 2009).

2. The cultivars ‘Skye’ and ‘Chicago II’ were previously assigned to the Compact type based on morphological data and performance characteristics (Table 5). The revised classification system assigns these cultivars to the Midnight type (Table 6). This change is supported by the UPGMA analysis showing these two varieties to be the basal varieties of the Midnight type (Fig. 1), and the PCO analysis which places these two cultivars between the main
concentration of Midnight cultivars and the remainder of the entries in the study (Fig. 3). The UPGMA and PCO analysis are consistent with the pedigree information for these varieties, both of which are hybrids with the cultivar ‘Midnight’.

3. The cultivars ‘Diva’, ‘Moonlight’, and ‘Bluetastic’ were previously assigned to the Compact type based on morphological data and performance characteristics (Table 5). The revised classification system assigns these cultivars to the Compact-America type (Table 6). This change is supported by the UPGMA analysis (Fig. 1), model based clustering analysis (Fig. 2), and PCO analysis (Fig. 3), all of which group these three cultivars with other Compact-America cultivars. This is consistent with pedigree information, as all three cultivars can be traced back to the Compact-America type through the cultivar ‘America’ (Funk et al., 1982) or C-74.

4. The cultivars ‘Cheetah’, ‘Boomerang’ and ‘Bodacious’, as well as the experimental selections SRX27921, PpH7907, and PpH7929 all trace back to the European cultivar ‘Cynthia’ (Kenneth Hignight – personal communication). The revised classification system places these entries in the Eurasian type (Table 6) based on the UPGMA, model based clustering, and PCO analyses. All six of these entries form a small clade, located adjacent to the Julia sub-type, within the Eurasian type (Fig. 1).

5. The cultivar ‘Kenblue’ and the experimental selection GO9LM9 have been previously classified in the Common type. The revised classification system places these entries in the Eurasian type (Table 6) based on UPGMA, model
based clustering, and PCO analyses. The original description of the Common type (Bara et al., 1993) stated that cultivars in this group were commonly selected from naturalized ecotypes surviving within old pastures located in the Midwestern United States. A plausible explanation for these entries grouping in the Eurasian type is that these cultivars may be descendants from remnant populations of pasture grasses originally seeded by European settlers in the Midwestern US. SSR genotyping of additional cultivars and accessions from the Common type will be needed to support this hypothesis.

The current SSR data set, and corresponding Jaccard’s distance, is a logical starting point for the development of a boundary for distinctness, uniformity, and stability (DUS) testing within a Kentucky bluegrass Plant Variety Protection (PVP) program. The key issue in a DUS testing program is the establishment of distinctness. A variety may be considered to be distinct if the characteristics of the variety are consistent, clear, and distinguishable from other varieties (UPOV, 2002). Distinctness refers to the minimum distance that should exist between two plant varieties so as to provide a framework of protection around an existing variety already granted intellectual property rights. Similar concepts for DUS testing using genetic distance/similarity based on molecular markers have been proposed in a wide variety of crop species including sugar beet (Beta vulgaris L.) (De Riek et al., 2001), perennial ryegrass (Lolium perenne L.) (Roldan-Ruiz et al., 2001), rape (Brassica napus L.) (Tommasini et al., 2003), aromatic rice (Oryza sativa L.) (Singh et al., 2004), pepper (Capsicum annuum L.) (Kwon et al., 2005), maize (Zea mays L.) (Gunjaca et al., 2008), and durham wheat (Triticum durum Desf.) (Noli et al., 2008),
to name a few. Ultimately, the general procedures to establish minimum distance must be implemented on a crop by crop basis, and are dependent on factors such as the genetic variability of the species, mode of reproduction, and type of molecular marker system used to generate the data (Ibanez et al., 2009). Additionally, DUS testing using molecular markers must be adopted by treaty organizations that govern plant breeders rights (e.g. UPOV Working Group on Biochemical and Molecular Techniques) (van Eeuwijk and Baril, 2001). The current research could serve as a starting point for the review of the use of SSR markers for Kentucky bluegrass PVP.

**Kentucky bluegrass classification types not included in the current classification revision**

There are four previous classification types that are not represented in the current classification revision based on SSR markers (Table 5 versus Table 6): CELA type, Common type, High Density (Aggressive) type, and Texas x Kentucky bluegrass hybrids. The Common type has already been discussed (see previous section). High Density (Aggressive) type cultivars were underrepresented in the current study. Bonos et al. (2000) reported that this type may only be related by the common growth characteristic of high shoot density, indicating that this type may not be a genetically related group. This observation may be supported by the fact that some cultivars that were previously a part of this type (Limousine and ‘Julius’) are now classified as Limousine related in the current revision (Table 6; Fig. 1; Fig. 2). SSR genotyping of additional cultivars that were previously considered to be High Density (Aggressive) type cultivars will be needed to determine whether or not this group should be retained in the classification system.
The CELA type, represented by the type cultivars ‘Challenger’ (Meyer et al., 1987), ‘Eclipse’ (Funk et al., 1981), ‘Liberty’ (Brilman et al., 1989), and ‘Adelphi’ (Funk et al., 1973), was also underrepresented in the current study. The situation with this group is different than the High Density (Aggressive) type in that the CELA type may be a genetically related classification type. The cultivars Adelphi and Eclipse share a common pedigree (Alderson and Sharp, 1994), as do Challenger and Liberty (Alderson and Sharp, 1994). It may, however, be problematic to determine the genetic relationships of this type using SSR markers because many of the cultivars in this group currently have very limited commercial production and/or have infrequently been used in the development of new cultivars. The final type not included in the current revised classification system is the Texas x Kentucky bluegrass hybrids. In the current study, the entries of this type do not form a distinct genetic group, but rather, are dispersed among multiple classification types (Table 5 versus Table 6; Fig.1; Fig. 2). These entries may be grouping according to the pedigree associated with the Kentucky bluegrass parent(s), however, the modified backcross procedure used to breed many of these entries is complex (multiple Kentucky bluegrass recurrent parents in different classification types), making pedigree analysis difficult. The current findings that this group is genetically diverse according to SSR marker genotypes is not surprising in light of the fact that performance characteristics of Texas x Kentucky bluegrass hybrids have been shown to be highly variable (Suplick-Ploense, 2002; Abraham et al., 2004; Reinert and Read, 2008; Richardson et al., 2009).
Conclusions

Contrary to the RAPD marker work of Curley and Jung (2004), SSR markers in the current study showed a strong correlation between genetic relatedness as assessed by molecular markers and the previously described Kentucky bluegrass cultivar classification system based on pedigree information, morphological traits, and cultivar performance characteristics. Additionally, SSR marker analysis in the current study provided justification for a revision/update of the Kentucky bluegrass cultivar classification system. This revision includes the near elimination of almost all entries in the current study from the Other type. Classification types that are supported by the current SSR analysis include BVMG, Compact, Compact-America, P-105 related, Mid-Atlantic, Eurasian, Julia, Shamrock, Sydsport related, Limousine related, and Midnight types. All cultivars, experimental selections, collections, and hybrids were uniquely identified with the current set of SSR markers. Genetic relationships of individuals as assessed by the current set of SSR markers closely match known pedigrees. The current set of SSR markers can be used to rapidly genotype and assign new cultivars/accessions to Kentucky bluegrass classification types and assess genetic relatedness among individuals. Additionally, the current SSR data set and corresponding Jaccard’s distance, should be considered as a starting point for the review of the use of SSR markers for Kentucky bluegrass PVP.
Literature Cited


Soreng, R. J. and F. R. Barrie. 1999. Proposal to conserve the name Poa pratensis (Gramineae) with a conserved type. Taxon. 48:157-159.


CHAPTER 4

Classification of Kentucky bluegrass (Poa pratensis L.) cultivars and accessions II: A comparison of morphological and molecular methods

Abstract

Estimation of genetic diversity and variety identification is important in plant breeding programs for understanding genetic relationships among germplasm resources and for variety identification in the process of commercialization and protection of new cultivars. Plant variety protection (PVP) testing typically involves growing candidate and previously registered varieties together, and testing for statistical differences between crop-specific morphological characteristics called descriptors. Although PVP testing using morphological characters has been sufficient to distinguish plant varieties, the use of plant comparative morphology has numerous limitations. A wide range of diverse cultivars and accessions of Kentucky bluegrass (Poa pratensis L.) have been previously characterized based on pedigree, common turf performance, morphological characteristics, and molecular markers to create a Kentucky bluegrass cultivar classification system. The objectives of the current study were to assess the genetic relationships of 172 Kentucky bluegrass cultivars and experimental selections using both microsatellite markers and morphological descriptors, and to determine which data set more accurately reflects known historical Kentucky bluegrass pedigree information. Genetic relationships of individuals as assessed by SSR markers closely matched known pedigrees, while morphological data showed very poor relationships with known pedigrees of Kentucky bluegrass cultivars and experimental selections. Results of this
study indicate that the current set of SSR markers could be used as a complement to PVP testing procedures for Kentucky bluegrass, as SSR markers more accurately reflect known genetic relationships of a large number of Kentucky bluegrass cultivars and experimental selections than do morphological descriptors.

**Introduction**

Estimation of genetic diversity and variety identification are of vital importance in plant breeding programs. Plant breeders can use the knowledge of genetic relationships among germplasm resources and cultivars as a parental selection tool to improve performance and quality, introduce novel traits, and to maintain genetic diversity. Additionally, variety identification is a critical step in the process of commercialization of new varieties and the application of plant variety protection (PVP). The development of PVP programs is required for countries that are members of the World Trade Organization (WTO) and have signed the Trade-Related Aspects of the Intellectual Property Rights (TRIPS) agreement (Kwon et al., 2005). PVP systems are designed to comply with the Convention of the International Union for the Protection of New Plant Varieties (UPOV, 1991), and grant Plant Breeders Rights (PBR) based on testing of varieties for Distinctness, Uniformity, and Stability (DUS). DUS testing typically involves growing candidate and previously registered varieties together, and testing for statistical differences between crop-specific morphological characteristics called descriptors (Roldan-Ruiz et al., 2001; Kwon et al., 2005). Although DUS testing using morphological characters has been sufficient to distinguish plant varieties, the use of plant comparative morphology has numerous limitations including the maintenance of
increasingly large reference collections for comparative analysis, a limited number of descriptors available to distinguish varieties, time-consuming field based measurement of large numbers of samples and replicates, and the potential for the expression of morphological traits to be influenced by the environment (Lombard et al., 2000; Roldan-Ruiz et al., 2001; Giancola et al., 2002; Kwon et al., 2005; Ibanez et al., 2009). Additionally, and perhaps most importantly, assessment of morphological characteristics may not equate to genetic relatedness, as genetic relatedness between varieties is ultimately determined by similarities/differences in DNA sequences (Roldan-Ruiz et al., 2001). This raises the question of whether or not DNA markers should be used as a complement or an alternative to morphological descriptors for Distinctness testing in PVP programs. While numerous researchers have presented arguments in favor of the use of DNA markers in PVP programs (Roldan-Ruiz et al., 2001; Giancola et al., 2002; Heckenberger et al., 2002; Cooke and Reeves, 2003; Heckenberger et al., 2003; Tommasini et al., 2003; van Eeuwijk and Law, 2004; Vosman et al., 2004; Heckenberger et al., 2005; Kwon et al., 2005; Borchert et al., 2008; Gunjaca et al., 2008; Smykal et al., 2008; Bonow et al., 2009; Ibanez et al., 2009; Smith et al., 2009), relatively few papers (Giancola et al., 2002; Taran et al., 2004; Gunjaca et al., 2008; Noli et al., 2008) have assessed whether morphological descriptor or molecular marker analyses better reflect known pedigree information in the species under investigation. The objectives of the current paper were: 1) to assess the genetic relationships of Kentucky bluegrass cultivars and experimental selections using microsatellite markers; 2) to assess the genetic relationships of Kentucky bluegrass cultivars and experimental selections using morphological descriptors commonly used in Kentucky bluegrass PVP analysis; and 3) to
determine which data set (molecular marker or morphological descriptor) more accurately reflects known historical Kentucky bluegrass pedigree information.

**Materials and Methods**

**Plant Material**

One hundred seventy-two Kentucky bluegrass cultivars and experimental selections were evaluated in this study. These cultivars and experimental selections included the complete set of unique entries evaluated for morphological characteristics by Shortell et al. (2009) and a matching subset of entries evaluated using microsatellite (SSR) markers by Honig et al. (2010a). Morphological measurements used in the current analysis were collected by Shortell et al. (2009). Single seedlings of each entry were transplanted into 48-cell flats (90 cm x 45 cm) and allowed to establish in the greenhouse. Plants were screened for apomixis in the greenhouse and all off-types (aberrant progeny) were discarded. The plants were then established in a spaced-plant nursery at the Rutgers University Plant Biology and Pathology Research and Extension Farm in Adelphia, NJ, in April 2003 on a well drained Freehold sandy loam (fine-loamy, mixed, mesic, Typic Hapludult). The experiment was established in a randomized complete block design with three replications. Four plants of each entry were planted per replicate for a total of 12 plants per entry. The field received 98 kg N ha$^{-1}$/year using a 3.5-4.4-8.3 fertilizer (elemental N, P$_2$O$_5$, and K$_2$O, respectively) applied as two split applications, half in the spring and half in the fall. The field was mowed at a height of 20 cm twice per year, once after seed maturity and once before the onset of winter. Labeled pre- (DCPA [dimethyl 2,3,5,6-tetrachlorobenzene-1,4dicarboxylate] and dithiopyr [3,5-pyridinedicarbothioic...
acid, 2-(difluoromethyl)-4-(2-methylpropyl)-6-(trifluoromethyl)-S,S-dimethyl ester]) and postemergence (dicamba [3,6-dichloro-2-methoxy benzoic acid] and halosulfuron [Methyl 5-[(4,6-dimethoxy-2pyrimidinyl)amino)carbonylaminosulfonyl]-3-chloro-1-methyl-1H-pyrazole-4-carboxylate]) herbicides were used to control broadleaf weeds, sedges, and Poa annua (L). The field did not receive supplemental irrigation, and no disease outbreaks were observed.

Plants utilized for the microsatellite analysis were collected as previously described by Honig et al. (2010a). Single seedlings or single tillers of all entries were transplanted into 48-cell flats (90 cm x 45 cm) and allowed to establish in the greenhouse. The plants were visually screened for apomixis in the greenhouse, and all off-types (aberrant progeny) were discarded. Plants were then established in various spaced-plant nurseries at the Rutgers University Plant Biology and Pathology Research and Extension Farm at Adelphia, NJ, in April 2003 through April 2006 on a well drained Freehold sandy loam (fine-loamy, mixed, mesic, Typic Hapludult). Plants were maintained in the field for a minimum of two years for further apomixis screening. A single clone of each entry, representing the repeating apomictic mother clone, was collected from the field, established in the greenhouse and maintained for DNA extraction in the laboratory.

**Morphological Data Collection**

The morphological characteristics were measured on spaced plants in 2004 and 2005 following the experimental methods required for PVP applications (USDA, 2005). Plant height, panicle length, flag leaf height, width, and length were taken approximately two weeks after anthesis, when all plant panicles were fully expanded (late-May to mid-
July in both years). Rhizome spread was measured after the first mowing. Plant height was measured from the soil surface to the average height of the majority of the panicles. Rhizome spread was measured in two directions perpendicular to each other, and represented the longest extending rhizomes in those two directions. Rhizome spread and plant height were taken on each of the four individual plants per replicate and averaged to obtain one value per replicate. Panicle length, as well as flag leaf height, width and length were calculated as the average of three separate flowering culms from each of the four plants per replicate, and averaged to obtain one value per replicate. Panicle length was measured from the bottom node to the tip of the longest panicle. Flag leaf height was measured from the soil surface to the collar of the flag leaf. Flag leaf blade width was measured at the widest point on the flag leaf. Flag leaf blade length was measured from the flag leaf collar to the tip of the blade. Due to the large number of entries and high temperatures of late spring/early summer plants were sampled in order of maturity, which served to minimize the effects of plant desiccation caused by seed maturation.

**Microsatellite development and genotyping**

The development of 88 Kentucky bluegrass microsatellite markers, characteristics of the SSRs, and the primer sequences for those markers was previously described by Honig et al., (2010b). Microsatellite markers NJPpGA6, NJPpGA931, NJPpGA937, NJPpGA993, NJPpGA1117, NJPpGT8, NJPpGT12, and NJPpGT9318 were not used in the current analysis because they produced more than 5% missing data in the 172 entries. DNA was isolated from the 172 entries using Sigma GenElute Plant Genomic DNA Miniprep Kit (St. Louis, MO) according to the manufacturer’s instructions. For
genotyping, amplifications of microsatellite loci with fluorescently labeled PCR primers were performed in 13 μl reactions using 50 ng of Kentucky bluegrass DNA, 1X Ramp-Taq PCR buffer (Denville Scientific, Metuchen, NJ), 2 mM MgCl₂, 0.25 mM each dNTP (Denville Scientific), 0.5 U Ramp-Taq DNA polymerase (Denville Scientific), 0.5 pmol forward primer with 18 bp M13(-21) addition at the 5’-end (Schuelke, 2000), 1 pmol reverse primer, and 1 pmol fluorescent dye-labeled 18bp M13(-21) primer (FAM, NED, PET or VIC). Thermal-cycling parameters were 94 °C for 5 min., followed by 30 cycles of 94 °C for 30 s, 55 °C for 45 s, 72 °C for 45 s, followed by 20 cycles of 94 °C for 30 s, 53 °C for 45 s, 72 °C for 45 s, ending with a final extension of 72 °C for 10 min. PCR reaction products were analyzed on an ABI 3130xl Genetic Analyzer and sized using Genemapper 3.7 (Applied Biosystems, Foster City, CA) and LIZ 500(-250) size standard.

Although microsatellites can generate codominant data, problems can arise during the identification of polyploid genotypes because it is difficult to determine the number of copies of an allele in heterozygotes (Saltonstall, 2003; Markwith et al., 2006; Liao et al., 2008). Thus, banding patterns observed at particular loci are referred to as ‘allele phenotypes’ (Becher et al., 2000). The individual alleles of the SSRs utilized in the current study of 172 polyploid Kentucky bluegrasses were treated as dominant markers, where the banding phenotypes were scored as band absence (0) or presence (1) (binary data matrix).

**Morphological Data analysis**

To determine differences in morphological characteristics between Kentucky bluegrass cultivars, morphological data were subjected to analysis of variance (ANOVA)
using the generalized linear model (GLM) procedure of SAS version 9.1 (SAS Institute, Cary, NC). Means were separated using Fisher’s protected least significant difference at the 0.05 $P$ level. Due to space considerations data on individual cultivar responses were not included in this article. All morphological data were previously reported by Shortell et al. (2006). The relationships between cultivars and experimental selections based on all morphological characteristics were tested using the multivariate principal component analysis procedure PROC PRINCOMP in SAS Version 9.1 (SAS Institute, Cary, NC). Principal components were then used as input variables for a cluster analysis using the unweighted pair group method with arithmetic mean (UPGMA) algorithm to generate a dendrogram using the PROC CLUSTER procedure in SAS Version 9.1 (SAS Institute, Cary, NC) (Cortese et al., 2010). The cultivar ‘Barzan’ was chosen as the root of the UPGMA tree. Of the cultivars utilized in the current analysis, Barzan was found to be most closely related to the outgroup species, *P. annua*, in the analysis conducted by Honig et al. (2010a).

**Microsatellite Data analysis**

Only strong and reliably amplified SSR marker alleles were scored and assembled into a binary data matrix. Pairwise comparisons of the proportion of shared alleles between individual genotypes (plants) were determined by the Jaccard (1908) coefficient, where $J = \frac{\text{matching fragments}}{\text{matching} + \text{polymorphic fragments}}$. The resulting genetic similarity coefficient was then transformed into a distance matrix ($d = 1 – J$), averaged over 1000 random draws from the interval of the minimum and maximum Jaccard distance. Cluster analysis was generated from the distance matrix by
the unweighted pair group method with arithmetic mean (UPGMA) algorithm. The cultivar Barzan was chosen as the root of the UPGMA tree. All calculations were performed by the computer program FAMD 1.23 (Schluter and Harris, 2006). The UPGMA tree was visualized using the computer program Dendroscope 2.6.1 (Huson et al., 2007).

Results

The ANOVA of the morphological data is presented in Table 9. A significant year x cultivar interaction was observed for all morphological measurements (Table 9). Therefore, the cluster analysis of the morphological data is presented separately by year. The UPGMA cluster analysis based on SSR data is shown in Fig. 4. The UPGMA cluster analyses based on the morphological data are shown in Fig. 5 and Fig. 6 for 2004 and 2005, respectively. Immediately apparent when looking at all three diagrams is the lack of grouping by Kentucky bluegrass classification type for the morphological data in both years (Fig. 5; Fig. 6), compared to the strong grouping by Kentucky bluegrass classification type for the SSR data (Fig. 4). The UPGMA cluster analysis based on SSR data for the 172 cultivars and experimental selections in the current study is similar to the UPGMA cluster analysis based on SSR data for the full set of 265 Kentucky bluegrass cultivars, experimental selections, collections, and hybrids in Honig et al. (2010a): In both studies SSR data clearly defined the classification types Eurasian, Julia sub-type within Eurasian, BVMG, Shamrock, Sydsport related, Limousine related, Midnight,
Table 9. Analysis of Variance (ANOVA) of cultivar and year effects on various morphological measurements of Kentucky bluegrass cultivars and experimental selections measured on spaced plants in 2004 and 2005 at the Rutgers University Plant Biology and Pathology Research and Extension Farm in Adelphia, NJ.

<table>
<thead>
<tr>
<th>ANOVA Source</th>
<th>Plant Height</th>
<th>Panicle Length</th>
<th>Flag Leaf Height</th>
<th>Flag Leaf Width</th>
<th>Flag Leaf Length</th>
<th>Plant Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultivar</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Year</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Cultivar x Year</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>*</td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>

*,*** Significant at the 0.05 and 0.001 probability levels, respectively.
Fig. 4. An UPGMA cladogram of 172 Kentucky bluegrass cultivars using 80 SSR primer pairs
Fig. 4 cont. An UPGMA cladogram of 172 Kentucky bluegrass cultivars using 80 SSR primer pairs
Fig. 5. An UPGMA cluster analysis of six morphological characteristics of 172 Kentucky bluegrass cultivars and experimental selections in 2004. Distances between clusters are expressed in $R^2$ values, where high $R^2$ values indicate greater similarity.
Fig. 5 cont. An UPGMA cluster analysis of six morphological characteristics of 172 Kentucky bluegrass cultivars and experimental selections in 2004. Distances between clusters are expressed in $R^2$ values, where high $R^2$ values indicate greater similarity.
Fig. 6. An UPGMA cluster analysis of six morphological characteristics of 172 Kentucky bluegrass cultivars and experimental selections in 2005. Distances between clusters are expressed in $R^2$ values, where high $R^2$ values indicate greater similarity.
Fig. 6 cont. An UPGMA cluster analysis of six morphological characteristics of 172 Kentucky bluegrass cultivars and experimental selections in 2005. Distances between clusters are expressed in $R^2$ values, where high $R^2$ values indicate greater similarity.
Compact, Compact-America, Mid-Atlantic, and P-105 related. The topology of the SSR UPGMA diagram in the current study is also similar to Honig et al. (2010a). The Eurasian and Julia types are basal to the remainder of the Kentucky bluegrass cultivars and experimental selections (Fig. 4), and the remaining entries are split into two major lineages: one consisting of the Shamrock, Sydsport related, Limousine related, Midnight, and BVMG types (Fig. 4), and the other comprised of the Compact, Compact-America, P-105 related, and Mid-Atlantic types (Fig. 4). Within the latter two lineages, there are some minor topological changes in the ordering of classification types in the current study compared to the SSR UPGMA clustering in Honig et al. (2010a), however, it is clear that the classification grouping is preserved and the major relationships are similar. The minor topological differences referred to above could be due to the removal of 93 entries from the current SSR analysis. In the current study, the clear grouping patterns for the SSR data, by Kentucky bluegrass classification type, are consistent with known pedigrees of Kentucky bluegrass cultivars and experimental selections.

In contrast to the strong grouping of SSR data by Kentucky bluegrass classification type (Fig. 4), the UPGMA diagrams based on morphological data (Fig. 5; Fig. 6) show very little grouping according to classification type. Exceptions may include some Midnight type cultivars in both 2004 and 2005 and some Compact-America type cultivars in 2005 (Fig. 5; Fig. 6); however, it is clear that even these groupings are loosely assembled and incomplete relative to the SSR data (Fig. 4). Overall, the UPGMA diagrams based on morphological data show very poor relationships with known pedigrees of Kentucky bluegrass cultivars and experimental selections.
Discussion

The goal of PVP is to award a certificate of protection to the owner of a variety after an examination determines that the variety is new, distinct from other varieties, and genetically uniform and stable through successive generations (USDA, 2005). Giancola et al. (2002) stated that the best system to differentiate and identify varieties must be practical, precise, reliable, and robust, and must offer high discrimination power, lack of interaction with the environment, capacity to generate equivalent results among laboratories, consistency of calculated genetic distances with pedigree data, good genomic coverage, public availability, susceptibility to automation and adequate cost/profit relationship. According to our results, the use of morphological data in Kentucky bluegrass fails two of the criteria outlined above: 1) lack of interaction with the environment (statistically significant genotype x year interaction); and 2) consistency of calculated genetic distances with pedigree data. Although not clearly stated in existing PVP laws, it would seem that common sense should dictate that descriptors/characters used to distinguish a new variety from previously existing varieties should accurately reflect the genetic relationships of the varieties being tested. The US Plant Variety Protection Office (PVPO) provides guidelines to breeder’s for writing a Statement of Distinctness which includes the language: “1) State the most similar previously existing variety, varieties, or identifiable group of varieties; and 2) State the character or characters that clearly distinguish the applicant’s variety from the varieties stated in step 1” (PVPO, 2010). The first statement clearly indicates the importance of understanding the relationship of the new variety to previously existing varieties. This is important for several reasons. First, disagreement found between morphological and molecular data
sets means that results of similarities or distinction tests could be different depending on the type of trait used for testing (Bernet et al., 2003). Second, as pointed out by Cooke and Reeves (2003), debates over the use of morphological or molecular characters in PVP has had the interesting effect of making people question the purpose of DUS testing. In other words, if the purpose of DUS testing is simply only to find some way of declaring two varieties to be different, then the current use of morphological characters/descriptors is appropriate (Cooke and Reeves, 2003). If, on the other hand, the purpose of DUS testing is to reward the outcome of a scientifically-based plant breeding program by establishing a ‘zone of protection’ around a variety, based on relatedness and taking associations between varieties into account, then the testing system should be based on metrics that reflect those associations (Cooke and Reeves, 2003; Law et al., 1999). Of the two data sets investigated in the current analysis, only the SSR data accurately depicts the genetic relationships among a large group of Kentucky bluegrass cultivars and experimental selections.

The goal of the current analysis is not necessarily intended to advocate for discontinuing the use of morphological measurements and the adoption of molecular markers as the sole method for DUS testing in Kentucky bluegrass. Rather, we believe the most important finding of this work is that conclusions reached in pair-wise comparisons of large numbers of cultivars using morphological characters and molecular markers are not the same. Results reported here indicate that morphological characters do little to characterize Kentucky bluegrass cultivars based on genetic relatedness, however, it is recognized that morphological measurements have been used successfully in DUS testing of Kentucky bluegrass cultivars since the adoption of PVP programs.
Additionally, morphological measurements will always be useful for characterizing varieties because the data can provide information as to maturity, color, density, growth habit, and other measurements important for successful breeding programs. Moving forward in the short term, the results presented here indicate that it would be logical to use the current set of SSR markers as a complement to the current DUS testing procedures for Kentucky bluegrass. As an example, SSR analysis of a new cultivar could accurately assess the most similar previously existing variety, varieties, or identifiable group of varieties as indicated in the guidelines developed by the US PVPO. Final DUS determination could still rely on traditional morphological or agronomic descriptors, with the advantage being that only two or a few cultivars have to be compared to reach a conclusion concerning distinctness. Additional advantages include the use of SSR markers as an additional/alternative determinate of distinctness/similarity when morphological descriptors are inconclusive. Longer term objectives could include developing a consensus among breeders and governing bodies (US PVPO and UPOV) to adopt a core set of Kentucky bluegrass cultivars and SSR markers with an agreed upon genetic distance measure that could serve as a basis for DUS testing of new cultivars in the future.
Literature Cited


Heckenberger, M., J. R. van de Voort, A. E. Melchinger, J. Peleman, and M. Bohn. 2003. Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties: II. Genetic and


CHAPTER 5

PCR marker-based genetic linkage map in allotetraploid creeping bentgrass

(Agrostis stolonifera L.)

Introduction

*Agrostis*, or bentgrass, is a large genus in the Poaceae, consisting of approximately 150 to 200 species worldwide (Hitchcock, 1951; Harvey, 2007). Five species are utilized as turfgrasses, including creeping bentgrass (*A. stolonifera* L.), colonial bentgrass (*A. capillaris* L.), velvet bentgrass (*A. canina*), dryland bentgrass (*A. castellana* Boiss. and Reut.), and redtop bentgrass (*A. gigantea* Roth.) (Warnke, 2003). These species are perennial, outcrossing grasses used for lawns, athletic fields, and golf courses. Creeping bentgrass is the most widely utilized cool-season turf species for intensively managed sports playing surfaces such as bowling greens and golf course putting greens, tees, and fairways (Turgeon, 1996). As the most desirable cool-season turfgrass for intensively managed sports surfaces, creeping bentgrass is an economically important turfgrass species in the United States.

DNA markers are used for a number of purposes in genetic studies including genotypic profiling, population genetic analyses, genetic linkage and comparative mapping, positional cloning, detection of quantitative trait loci (QTLs), and marker-assisted selection (MAS). Molecular marker-based genetic linkage maps were developed in a number of economically important crop species beginning in the late 1980’s and early 1990’s: early work was conducted in rice (*Oryza sativa* L.) (Causse et al. 1994; Kurata et al. 1994b; Harushima et al. 1998), hexaploid wheat (*Triticum aestivum* L. em.
Thell) (Chao et al. 1989; Liu and Tsunewaki 1991; Devos et al. 1992; Devos et al. 1993b; Xie et al. 1993; Nelson et al. 1995a,b,c), barley (*Hordeum* sp.) (Graner et al. 1991; Heun et al. 1991; Kleinhofs et al. 1993), oat (*Avena* sp.) (O’Donoughue et al. 1992; O’Donoughue et al. 1995; Rayapati et al. 1995; Van Deynze et al. 1995a), maize (*Zea mays* L.) (Burr and Burr 1991; Gardiner et al. 1993), sorghum (*Sorghum* sp.) (Chittenden et al. 1994), and sugarcane (*Saccharum* sp.) (da Silva et al. 1993). These early linkage maps, as well as linkage maps that immediately followed, were initially constructed with restriction fragment length polymorphisms (RFLPs) (O’Brien, 1993) or early PCR-based markers such as randomly amplified polymorphic DNA (RAPD) markers and amplified fragment length polymorphism (AFLP) markers. Although first generation maps have proved informative, particularly when utilizing RFLPs for comparative mapping within the Poaceae, these early marker systems have since been found to have numerous disadvantages. Hybridization based markers, such as RFLPs, require relatively large amounts of high quality DNA (Miao et al., 2005), the use of radioactivity for labeling, and the hybridization procedure itself is time-consuming and laborious making RFLPs impractical for use in applied agricultural research (Ramsey et al., 2000). RAPD markers use short random oligonucleotides as primers and low annealing temperature to amplify genomic DNA sequences. Due to the unpredictable behavior of the short primers and low annealing temperature, RAPDs tend to have low repeatability (Penner et al., 1993; Jones et al., 1997; Perez et al., 1998; Liu et al., 2005). AFLP markers are more reliable than RAPD markers, however, the AFLP technique requires extensive pre-PCR sample preparation, including restriction digestion followed by the ligation of adaptors and pre-amplification, making AFLP relatively more laborious (Liu et al., 2005).
The disadvantages associated with the above marker systems has led to considerable interest in the use of more advanced PCR-based markers for linkage mapping such as microsatellites or simple sequence repeats (SSRs). SSR markers have been proven to have tremendous utility in linkage mapping studies because they are: (1) PCR-based; (2) generally codominant; (3) multiallelic; (4) highly polymorphic; (5) highly reproducible; (6) randomly and uniformly distributed throughout eukaryotic genomes; and (7) are easily accessible to other research laboratories via published primer sequences (Yu et al., 2000; Miao et al., 2005).

Another more advanced technique is the use of completed genome sequences from model plant species and important agricultural crops to provide templates for the design of genomic tools in orphan species lacking sequence information (Feltus et al., 2006). One such approach is the use of intron polymorphisms as genetic markers. An early strategy for this approach was referred to as exon-primed intron-crossing (EPIC)-PCR (Palumbi, 1995), where PCR primers are designed in exonic sequences and prime into intervening intronic sequences. The advantage of EPIC-PCR is that exonic sequences are relatively more conserved; therefore, the primers designed in exons may have more extensive applications, such as cross-taxa amplification, than those designed in non-coding sequences (Wang et al., 2005). Primers can be designed from conserved exon sequences that flank introns in order to maximize intronic polymorphism discovery rates within a taxon while maintaining cross taxa applicability via DNA conservation in the priming sites (Lohithaswa et al., 2007). Both conserved intron spanning primers (CISP) (Feltus et al., 2006; Lohithaswa et al., 2007) and intron length polymorphism (ILP) (Wang et al., 2005) markers are considered to be EPIC-PCR markers, where each
are designed so that PCR primers are anchored in conserved exons and prime into non-conserved introns. Both types of markers were originally developed using conserved exonic sequence information, and have been previously aligned to the rice (*Oryza sativa*) genome (Wang et al., 2005; Feltus et al., 2006; Lohithaswa et al., 2007). Cross taxa amplification in a number of Poaceae species, as well as *Musa, Allium,* and *Gossypium* has been previously demonstrated (Wang et al., 2005; Feltus et al., 2006; Lohithaswa et al., 2007). These types of markers serve as comparative anchor markers to the rice genome, have high rates of cross taxa amplification due to PCR primers anchored in conserved exonic sequences, and have high rates of polymorphism due to priming into non-conserved intronic sequences (Wang et al., 2005; Feltus et al., 2006; Lohithaswa et al., 2007).

A final reason to shift from markers such as RFLPs, RAPDs, and AFLPs to SSRs and other more advanced PCR-based marker systems is the potential to use constructed linkage maps in traditional plant breeding programs (Gupta and Varshney, 2000). Modern plant breeding programs require rapid, reliable, low-cost molecular analysis of thousands of plants to identify QTLs and implement MAS programs (Somers et al., 2004). This is only possible with SSRs or other advanced PCR based maker systems that are amenable to repeatable, rapid and high throughput genotyping technologies.

The development of the first genetic linkage map in allotetraploid creeping bentgrass was reported by Chakraborty et al. (2005). The map was constructed using 169 RAPD, 180 AFLP, 39 heterologous cereal cDNA RFLP, and 36 homologous bentgrass cDNA RFLP markers. The linkage map consisted of 424 mapped loci covering 1110 cM in 14 linkage groups. While this map was very useful for identifying the genomic
structure of creeping bentgrass, the RAPD markers are likely to have limited
transferability among laboratories, while the RFLP and AFLP markers are not easily
amenable to high throughput genotyping platforms. As a result these markers may have
limited practical value in an applied creeping bentgrass breeding program. The
objectives of the current study were to develop a framework linkage map of creeping
bentgrass using AFLP, SSR, CISP (Feltus et al. 2006), and ILP (Wang et al., 2005)
markers. SSR, CISP and ILP markers should offer superior utility for construction of
linkage maps, identification of QTLs, and the implementation of MAS in a creeping
bentgrass breeding program because they are more amenable to rapid, reliable, low-cost
molecular analysis of large numbers of plants than marker systems previously used in
creeping bentgrass linkage map construction.

Materials and Methods

Plant material and mapping population

Two heterozygous parental creeping bentgrass clones, 7418-3 and L93-10, were
selected from an evaluation trial of 500 clones collected from golf course fairways and
putting greens in Arizona, Illinois, New Jersey, and New York, and from commercially
available cultivars (Bonos et al. 2004). 7418-3 was collected from Piping Rock Golf
Course in Long Island, NY in 1996 (Bonos, 2010). L93-10 was a single clone selected
from the cultivar L93. These two creeping bentgrass clones exhibit differences in leaf
color, shoot density, lateral growth rate and disease response to dollar spot (Sclerotinia
homoeocarpa F.T. Bennett) (Bonos, 2006). 7418-3 exhibits lighter green leaf color, lower
shoot density, longer internodes, wider leaves, and greater susceptibility to dollar spot
than L93-10.
During the spring of 2000, two pseudo F₂ (F₁) populations were created from a reciprocal cross of the two creeping bentgrass clones 7418-3 and L93-10, crossed in a greenhouse located at the New Jersey Agricultural Experiment Station, New Brunswick, NJ. Seeds were collected in late July from both parental clones. The first population, which utilized 7418-3 as the mother and L93-10 as the father, produced 95 individuals, and the second population, which utilized L93-10 as the mother and 7418-3 as the father, produced 86 individuals, for a total of 181 progeny from both crosses. Seeds from each population were initially sown separately into 15.2 cm round pots. After germination and growth in the greenhouse, individual seedlings were transferred to single 6.4 cm cells of 48 cell flats and assigned a mapping population number.

**SSR development and analysis**

Microsatellite development and genotyping was previously described by Kubik et al. (2010). Briefly, for SSR development total genomic DNA was extracted from MCB17, a single creeping bentgrass clone collected from Piping Rock Golf Course in Long Island, NY, using Qiagen DNAeasy Plant Mini Kit (Valencia, CA). DNA was sent to Genetic Identification Services Inc. (GIS, Chatsworth, CA) for construction of an SSR library enriched for GA, GT, GCT, and AAT SSRs. Methods for DNA library construction and enrichment were developed following Jones et al. (2002): Genomic DNA was partially digested with a cocktail of seven blunt-end restriction enzymes (RsaI, HaeIII, BsrBI, PvuII, Stul, Scal, EcoRV). Fragments in the size range of 300 to 750 bp were adapted with 20 bp oligonucleotides, which contained a HindIII site at the 5’ end and subjected to magnetic bead capture (CPG, Inc., Lincoln Park, NJ) using 5’-biotinylated GA₁₅, GT₁₅, GCT₁₂, and AAT₁₂ as capture molecules according to the
manufacturer’s instructions. Captured molecules were amplified using a primer complimentary to the adaptor, digested with HindIII to remove the adaptor sequences, and ligated into the HindIII site of pUC19. Recombinant plasmids were then electroporated into *Escherichia coli* DH5α. GIS delivered DNA libraries 50% enriched for GA and GT repeats and libraries 10% enriched for the tri-nucleotides GCT and AAT.

Several 100-μl aliquots of each SSR library were plated out onto LB agar plates containing ampicillin, IPTG and Bluo-Gal. The plates were incubated at 37 °C overnight. Several thousand individual white colonies were chosen from the GA and GT enriched libraries, and grown in 6 ml LB broth containing ampicillin. Several hundred colonies were chosen from the GCT and AAT enriched libraries. DNA was isolated from the cultures using Qiagen Miniprep spin kit. Samples were sequenced using an ABI 3130xl DNA sequencer (Applied Biosystems, Foster City, CA). Sequence data from the clones containing SSRs were analyzed for primer selection, and polymerase chain reaction (PCR) primers were designed to flank regions surrounding the SSR motif using Primer3 software (Rozen and Skaletsky, 2000). The forward primer in the pair was elongated at the 5’ end using the M13(-21) 18 bp sequence (5’ – TGTAAACGACGGCCAGT – 3’) for economic fluorescent labeling (Schuelke, 2000) and synthesized by Integrated DNA Technologies (Coralville, IA). Primers were initially tested on the two creeping bentgrass parental clones (7418-3 and L93-10) to reveal polymorphisms and assess amplification success.

For genotyping, 217 SSR primer pairs (Kubik et al., 2010) that revealed repeatable polymorphisms between the two parental clones were screened in all 181 progeny (95 progeny from 7418-3 x L93-10 and 86 progeny from L93-10 x 7418-3).
PCR was conducted in 13-ul reactions using 50 ng genomic DNA, 1X Ramp-Taq PCR buffer (Denville Scientific, Metuchen, NJ), 2 mM MgCl₂, 0.25 mM of each dNTP (Denville Scientific), 0.5 U Ramp-Taq DNA polymerase (Denville Scientific), 0.5 pmol forward primer with M13(-21) addition, 1 pmol reverse primer, and 1 pmol M13(-21) florescent dye-labeled primer (FAM, NED, PET or VIC). Thermal-cycling parameters were 94 °C for 5 min., followed by 30 cycles of 94 °C for 30 s, 55 °C for 45 s, 72 °C for 45 s, followed by 20 cycles 94 °C for 30 s, 53 °C for 45 s, 72 °C for 45 s, ending with a final extension of 72 °C for 10 min. PCR reaction products were analyzed on an ABI 3130xl Genetic Analyzer and sized using Genemapper 3.7 software (Applied Biosystems) and LIZ 500(-250) size standard (Applied Biosystems). An additional 58 EST-SSR anchor loci (Yu et al., 2004) were screened in the 181 progeny using the methods described above for the creeping bentgrass SSR markers.

**CISP and ILP analysis**

Two-hundred twenty CISP (Feltus et al., 2006; Lohithaswa et al., 2007) and 1897 ILP (Wang et al., 2005) primer pairs were screened in the 181 progeny using the methods described above for the creeping bentgrass SSR markers. This methodology allowed for the detection of insertions/deletions (INDELs), identified as size differences during capillary electrophoresis. CISP and ILP markers were utilized to assess marker conservation between creeping bentgrass and rice genomes.
**AFLP analysis**

For amplified fragment length polymorphism (AFLP) analysis, genomic DNA was extracted using a modified CTAB (cetyltrimethylammonium bromide) extraction method (Saghai-Maroof et al., 1984) from 100 mg lyophilized plant leaf tissue. DNA digestions, adaptor ligation, and preselective and selective amplifications were performed according to instructions provided with an AFLP analysis kit purchased from Invitrogen (Carlsbad, CA) and standard AFLP procedures (Vos et al., 1995). The selective amplifications with three or four selective bases per primer were performed by using 6-carboxy fluorescein (FAM) fluorescent primers labeled on the 5' nucleotide. The amplified fragments were detected with an ABI 3730xl instrument (Applied Biosystems). Each sample lane included the GeneScan 500-LIZ internal lane standard. The capillary electrophoresis procedures were performed by the Bovine Functional Genomics Laboratory (Beltsville, MD). Fluorescent fragments between 50 and 500 nucleotides were identified using GeneScan 3.1 software (Applied Biosystems). GeneScan trace files were analyzed for the presence or absence of AFLP products, in 1-nt intervals using the computer software Genographer (Benham et al., 1999).

**SSR, CISP, and ILP allele scoring**

Although SSR, CISP, and ILP markers can all generate co-dominant data, problems can arise during the identification of polyploid genotypes because it is difficult to determine the number of copies of an allele in heterozygotes (Saltonstall 2003; Markwith et al. 2006; Liao et al. 2008). Thus, banding patterns observed at particular loci are referred to as ‘allele phenotypes’ (Becher et al. 2000). The individual alleles of
the SSR, CISP, and ILP markers utilized in the current study of allotetraploid creeping bentgrass were treated as dominant markers, where the banding phenotypes were scored as band absence (0) or presence (1). Segregation of alleles was investigated by the presence of either maternal markers (fragments segregating from L93-10) or paternal markers (fragments segregating from 7418-3).

For allele coding of SSR, EST-SSR, CISP, and ILP markers in this publication we used the primer pair name from corresponding previous publications (Kubik et al., 2010; Yu et al., 2004; Feltus et al., 2006; Lohithaswa et al., 2007; Wang et al., 2005), followed by the allele size in creeping bentgrass in parentheses. As an example for the creeping bentgrass SSRs, allele GA1591 (236) refers to the SSR primer pair GA1591, with an allele size of 236 bp in creeping bentgrass (coding follows Kubik et al., 2010). EST-SSR anchor markers (alleles designated as “KSU…””) used the coding from Yu et al. (2004), followed by the allele size in creeping bentgrass in parentheses. CISP alleles (alleles designated as “SRSC…”, “PRSC…”, “BRSC…”, and “ORSC…””) are immediately followed by a series of numbers, and then followed by the allele size in creeping bentgrass in parenthesis. The number series following CISP alleles refers to the rice chromosome number (number immediately following the letter designations and before the underscore character) and a primer pair number from the original publication (number immediately following the underscore character) (coding follows Feltus et al., 2006; and Lohithaswa et al., 2007). As an example, the CISP allele SRSC2_011(227) refers to the CISP primer pair SRSC2, aligned to rice chromosome 2, primer pair number 011, with an allele size of 227 bp in creeping bentgrass. ILP alleles (alleles designated “Os…””) are also immediately followed by a series of numbers, and then followed by the
allele size in creeping bentgrass in parenthesis. The number series following ILP alleles refers to the rice chromosome number (number immediately following the “Os” letter designation and before the underscore character) and the rice alignment position on that chromosome (number immediately following the underscore character) (coding follows Wang et al. 2005). As an example, the ILP allele Os01g_40860(242) refers to the ILP primer pair Os01g, aligned to rice chromosome 1, aligned at starting position 40860 bp on rice chromosome 1, with an allele size of 242 bp in creeping bentgrass.

**Linkage Analysis and Map Construction**

All markers were coded as dominant markers according to the coding systems described in JoinMap 3.0 (van Ooijen and Voorrips, 2001) for a cross pollinated (CP) population. Markers present and heterozygous in the maternal parent were scored as lm x ll and markers present and heterozygous in the paternal parent were scored as nn x np. Initially, two parental framework maps (maternal and paternal) were constructed for each of the two populations using markers with alleles segregating in one or both parents. After constructing the LOD grouping trees, we selected linkage groups (LGs) with a LOD grouping threshold of 5.0 or greater. To estimate recombination and order of markers for each linkage group, the default parameters set in JoinMap were used. These include a pairwise recombination estimation using $r < 0.40$ and a LOD score $> 1.0$ (ripple value = 1, jump threshold = 5). Map distances were calculated using the Kosambi mapping function (Kosambi, 1944). Deviations of all markers from the expected segregation ratio of 1:1 were determined using $\chi^2$ contingency tables in JoinMap.
Two parental maps for each population were first constructed with markers segregating in a 1:1 ratio. Markers present in both parents of each population that segregated 3:1 were added to the maps individually to provide more markers for the identification of homologous linkage groups. Linkage groups were visualized using MAPCHART 2.1 (Voorrips, 2002).

**Examination of Chromosomal Pairing Behavior**

The comparison of the number of linked loci identified in coupling and repulsion phase has been used to determine meiotic pairing behavior (Sorrells, 1992; Wu et al., 1992) and distinguish between polysomic and disomic inheritance. Polysomy is characterized by low frequencies or absence of repulsion-phase linked markers, whereas in diploids or strict allopolyploids, the number of markers linked in coupling and repulsion phase is expected to be in a 1:1 ratio (Chakraborty, 2005). The ratio of coupling to repulsion phase linkages was used to assess the level of homoeologous chromosome pairing in tetraploid creeping bentgrass (Qu and Hancock, 2001) similar to Chakraborty et al. (2005). The SSR data set was split into male and female data sets to represent markers coming from each of the parents in the first mapping population (7418-3 x L93-10) (coupling-phase data set). The marker data was duplicated and then inverted (0 to 1 and 1 to 0) to create a repulsion-phase data set. The combined coupling- and repulsion-phase data sets for each parent were analyzed for linkages using MAPMAKER 3.0 EXP.
Combined population from reciprocal cross

After visualization of the parental maps from each of the two populations, it was determined that the individual linkage maps from the reciprocal crosses were in good agreement (data not shown). Any minor observed differences showed no systematic pattern. Recombination rates were generally not increased in one direction and reduced in the other, and therefore, the direction of the cross was not important. This provided justification for combining both populations to create a joined map for all fourteen linkage groups (Korzun et al., 1998; Korzun et al., 2001). Data from both populations was then pooled to create a single mapping population consisting of 181 individuals. Linkage analysis and map construction was then repeated for the pooled pseudo F₂ population as described above for the two original populations.

Identification of homologous and homoeologous groups in pooled population

Markers that detected multiple alleles and duplicated loci were used to define the homologous linkage groups, representing the same chromosome in each parent, and the homoeologous linkage groups, representing one of the two genomes present in creeping bentgrass (Chakraborty et al., 2005; Saha et al. 2005). Each pair of linkage groups was designated with the group numbers based on synteny similar to Chakraborty et al. (2005), followed by an arbitrary “.1(7418-3)”, “.1(L93-10)”, “.2(7418-3)” or “.2(L93-10)” to indicate that they represent the two different creeping bentgrass diploid genomes.
Results

AFLP Markers

A total of 153 AFLP markers were scored in the 181 progeny, and 71 of these markers were placed on the linkage map. Of the 153 total scored AFLP markers, 67 were scored as present in the 7418-3 parent, and 86 were scored as present in the L93-10 parent. Of the 71 total mapped AFLP markers, 45 were mapped in the 7418-3 parent map and 26 were mapped in the L93-10 parent. These markers tended to cluster together on specific linkage groups, and were not present on all linkage groups.

SSR Markers

A total of 200 SSR and EST-SSR primer pairs mapped 468 alleles in the 181 progeny. Of these alleles, 235 alleles were present in the 7418-3 parent and 233 alleles were present in the L93-10 parent. The number of polymorphic alleles per primer that mapped ranged from one to six with an average 2.3 alleles per primer. Two percent (4) of the 200 mapped SSR primers identified six alleles and two percent (4) identified five alleles. All of the primers that produced greater than four alleles identified separate alleles found in both parents and five of these primers identified alleles shared between both parents. More than four alleles is theoretically not possible in a tetraploid; however, the above percentages of five and six alleles is a small percentage of the total alleles and likely represents duplicated loci within a single creeping bentgrass genome. Fourteen percent (27) of the 200 SSR primer pairs identified four alleles. All of these primers identified separate alleles found in both parents and eight of these primers identified alleles shared between both parents. Nineteen percent (37) of the 200 SSR primer pairs
identified three alleles. The majority (excluding three) of these primers identified separate alleles found in both parents and 11 of these primers identified alleles shared between both parents. Thirty-three percent (66) of the SSR primer pairs identified two alleles: Twenty-eight of these identified separate alleles found in both parents; 27 of these identified two alleles in one parent or the other; 10 of these identified one parental allele and one shared allele; and 1 identified 2 shared alleles only. Thirty-one percent (62) identified one allele. Of these primers, 28 identified alleles found only in 7418-3 and 33 identified alleles found only in L93-10. One primer (GT 1345) identified 1 shared allele.

CISP and ILP Markers

Thirty-three CISP primer pairs mapped 62 alleles in the 181 progeny. Of these alleles, 35 alleles were present in the 7418-3 parent and 27 alleles were present in the L93-10 parent. Six percent (2) of the 33 CISP primers amplified four alleles: Of these, one primer pair identified separate alleles found in both parents and a shared allele and the other primer pair identified three alleles found in one parent and a shared allele. Twenty-one percent (7) of the CISP primer pairs amplified 3 alleles: Of these, two primers identified separate parental alleles and a shared allele; two primers identified parental alleles in one parent only and a shared allele; one primer identified separate alleles found in both parents; one primer identified three alleles in one parent; and one primer identified three shared alleles. Fifteen percent (5) of the mapped CISP primers identified two alleles: Of these, one of the primers identified separate alleles in both parents, while all the remaining primers identified two alleles in one parent only. Fifty-eight percent (19) of the CISP primers identified only one allele: of these, eight primers
identified alleles found only in 7418-3 and 11 identified alleles found only in L93-10. One primer (BRSC3_005V) identified one shared allele.

Fifty-nine ILP primer pairs mapped 105 alleles in the 181 progeny. Of these alleles, 64 alleles were present in the 7418-3 parent and 41 alleles were present in the L93-10 parent. One of the 59 ILP primers amplified five alleles and identified separate alleles in both parents. Seven percent (4) of the mapped ILP primers amplified four alleles: Of these, two primers identified separate alleles in both parents; one primer identified separate alleles in both parents and a shared allele; and one primer identified two alleles in one parent and two shared alleles. Eight percent (5) of the mapped ILP primers amplified three alleles: Of these, three primers identified alleles found in only one parent; one primer identified separate alleles found in both parents; and one primer identified separate alleles found in both parents and a shared allele. Nineteen percent (11) of the mapped ILP primers amplified two alleles: Of these, three primers identified separate alleles in both parents; six primers identified alleles found in only one parent; one primer identified an allele in only one parent and a shared allele; and one primer identified only shared alleles. Sixty-four percent (38) of the ILP primers amplified only one allele in the mapping population: Of these, 15 primers identified alleles found only in 7418-3 and 23 primers identified alleles found only in L93-10.

**Examination of chromosomal pairing behavior**

The two-point linkage coupling and repulsion-phase linkage analysis resulted in a 1:1 ratio of coupling- to repulsion- phase linkages (data not shown). This is a ratio that indicates disomic inheritance in a tetraploid species (Chakraborty et al., 2005; Qu and
Hancock, 2001). If tetrasomic inheritance through random bivalent pairing or multivalent formation (polysomic inheritance) were occurring, far fewer replulsion-phase linkages would be expected than were observed (Chakraborty et al., 2005). This result supports disomic inheritance in tetraploid creeping bentgrass.

**Segregation Distortion**

Of the total number of markers mapped, 33.8% and 38.1% from the 7418-3 and L93-10 parental maps, respectively, exhibited segregation distortion. The remainder of the mapped markers exhibited normal Mendelian segregation. All linkage groups possessed markers exhibiting segregation distortion, ranging from 10% in LG6.2(L93-10) to 71.4% in LG3.2(L93-10). Although some LGs had loci with distorted segregation throughout the LG, there did appear to be clustering of loci with distorted segregation on the ends of the LGs.

**Genetic Linkage Map and Comparative Relationships**

*A. stolonifera* is an allotetraploid with \(2n=4x=28\) (comprised of two subgenomes A2 and A3) based on studies of chromosome pairing and isozyme analysis (Jones 1956a, b, c; Warnke et al., 1998). As expected, 14 LGs were detected for each parental map, covering a total of 1424 cM and 1374 cM for the 7418-3 and the L93-10 parental maps, respectively (Fig. 7). A total of 385 markers were used to create the 7418-3 parent map, while 328 markers were used to create the L93-10 parent map. Two hundred and forty-one SSR, 45 AFLP, 35 CISP and 64 ILP alleles were used to create the 7418-3 map. Two hundred and thirty-four SSR, 26 AFLP, 27 CISP and 41 ILP alleles were used to
Fig. 7. Creeping bentgrass genetic linkage map. The seven homoeologous linkage groups are numbered from 1 to 7, with number assignments based on syntenic relationships to rice and original LG assignment described by Chakraborty et al. (2005). Markers exhibiting segregation distortion are marked with asterisks (*, **, *** indicate significance levels of 0.05, 0.01, and 0.001, respectively). The centiMorgan distance scale is located to the left of each individual LG. AFLP markers begin with two sets of three nucleotides corresponding to the selective amplification primers. CISP markers begin with BRSC, ORSC, PRSC, or SRSC. ILP markers begin with Os. EST-SSR markers begin with KSU. The remainder of the markers are creeping...
Fig. 7. (Contd.)

Linkage Group 2

2.1 (L93-10)

2.1 (7418-3)

2.2 (L93-10)

2.2 (7418-3)
Fig. 7. (Contd.) 

Linkage Group 3

3.1(L93-10)  3.1(7418-3)  3.2(L93-10)  3.2(7418-3)

331
Fig. 7. (Contd.)

Linkage Group 4

4.1(L93-10)

4.1(7418-3)

4.2(L93-10)

4.2(7418-3)
Fig. 7. (Contd.)

**Linkage Group 5**

### 5.1 (L93-10)

<table>
<thead>
<tr>
<th>0</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA179(210)</td>
<td>GT97(155)</td>
</tr>
</tbody>
</table>

### 5.1 (7418-3)

<table>
<thead>
<tr>
<th>0</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA179(210)</td>
<td>GT97(155)</td>
</tr>
</tbody>
</table>

### 5.2 (L93-10)

<table>
<thead>
<tr>
<th>0</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>KSU104(150)</td>
<td>ACG_CAG_416.92</td>
</tr>
</tbody>
</table>

---

333
6.1(L93-10)

6.2(L93-10)

6.1(7418-3)

6.2(7418-3)

Fig. 7. (Contd.)
create the L93-10 map. LG length varied from 62 cM to 169 cM, and marker distribution ranged from 8 to 46 markers per LG (Fig. 7).

Based on the presence of multiple alleles and duplicate loci (homoeoalleles) between LGs we have identified 14 pairs of homologous LGs and seven pairs of homoeologous LGs for each parent (Fig. 7). Such pairing is characteristic of species with allopolyploid genomes. Eleven of the homologous linkage groups were identified by the presence of at least five pairs of alleles from common loci, with a high of 12 pairs of alleles from common loci for homologs 6.2 (Fig. 7). Homologs 2.1 and 2.2 had one and two pairs of alleles from common loci, respectively (Fig. 7). Homologs 3.1 had zero pairs of alleles from common loci, but did have three duplicate loci that connected them to homoeologous groups 3.2 (Fig. 7). We considered a linkage group to be homoeologous to another linkage group based on the presence of at least two pairs of duplicate loci (Fig. 7). Homoeologous groups 1 and 7 had two pairs of duplicate loci; homoeologous groups 2, 3, 5, and 6 had three pairs of duplicate loci; and homoeologous group 4 had five pairs of duplicate loci (Fig. 7).

In an attempt to assign the LG numbers of all seven basic LGs in the current bentgrass map, we assessed comparative syntenous relationships with rice using CISP markers, ILP markers, and a BLAST alignment search of creeping bentgrass SSR sequences. In general, CISP and ILP markers from more than one rice chromosome were found on the same creeping bentgrass LG. This is not surprising in light of the fact that evidence suggests that both macro- and micro-scale rearrangements of grass colinearity and synteny can be common across taxa (Gaut, 2002; Sorrells et al., 2003; La Rota and
Sorrells, 2004; Gaut et al., 2007; Gui et al., 2010). Even with CISP and ILP markers from more than one rice chromosome on the same bentgrass LG, associations seemed to exist for bentgrass LGs 2, 4, 5, and 7, with somewhat weaker associations for bentgrass LGs 1, 3, and 6. According to LG assignment conducted by Chakraborty et al. (2005), bentgrass LGs 1, 2, 3, 4, 5, 6, and 7 should be represented by rice chromosomes 5, 7, 1, 3, both 8 and 9, 2, and 6, respectively. The following is a summary of the bentgrass LG / rice chromosome associations in the current study:

- Bentgrass LG5 should be represented by rice chromosomes 8 and 9 (Chakraborty et al., 2005). In the current analysis, three CISP / ILP markers from rice chromosome 8 and 15 CISP / ILP markers from rice chromosome 9 were found on the homoeologous bentgrass LG assigned as bentgrass LG5 (Fig. 7). No other bentgrass LGs had more than two CISP / ILP markers from rice chromosome 8.

- Bentgrass LG7 should be represented by rice chromosome 6 (Chakraborty et al., 2005). In the current analysis, six CISP / ILP markers from rice chromosome 6 were found on the homoeologous bentgrass LG assigned as bentgrass LG7 (Fig. 7). Additionally, bentgrass SSR markers CAG336 (found on LG7.2(7418-3)), CAG398 (found on LG7.2(7418-3)), and P378 (found on LG7.1(L93-10) and 7.1(7418-3)), all were found to have high identity (99%, 98%, and 72%, respectively) to sequences on rice chromosome 6 (Kubik et al., 2011).

- Bentgrass LG4 should be represented by rice chromosome 3 (Chakraborty et al., 2005). In the current analysis, eight CISP / ILP markers from rice chromosome 3 were found on the homoeologous bentgrass LG assigned as bentgrass LG4 (Fig.
Additionally, SSR GT824(141), found on LG4.2(L93-10), was found to have 71% identity to a sequence found on rice chromosome 3 (Kubik et al. 2011).

- Bentgrass LG2 should be represented by rice chromosome 7 (Chakroborty et al., 2005). In the current analysis, three ILP markers from rice chromosome 7 were found on the bentgrass LG assigned as bentgrass LG2.2 (Fig. 7). Additionally, GT 169(114), found on LG2.1(7418-3), was found to have 87% identity to a sequence found on rice chromosome 7 (Kubik et al., 2011).

- Although other bentgrass LGs (LG1, LG3, and LG6) had some associations with ILP and/or CISP markers from the respective rice chromosomes proposed by Chakroborty et al. (2005), less than three ILP or CISP markers from rice were found to be associated with these bentgrass LGs. Therefore, LG assignments for bentgrass LGs 1, 3, and 6 should be assumed to be arbitrary in the current study.

**Discussion**

We have constructed the first PCR marker-based genetic linkage map of creeping bentgrass, containing a substantial number of SSR loci as well as AFLP, CISP, and ILP loci. The latter two marker types are important because they can be used to assess syntenous relationships between orphan crops lacking significant molecular genetic resources, such as creeping bentgrass, and other model cereal crops such as rice. Additional utility of a PCR marker-based linkage map includes the potential to use the constructed linkage map in traditional plant breeding programs to rapidly and reliably analyze large numbers of plants for the identification of QTLs and the implementation of MAS programs, at a reasonable cost.
In the current study, we have found a significant number of alternate / multiple alleles and duplicate loci (homoeoalleles) between LGs. This has allowed us to identify 14 pairs of homologous LGs and seven pairs of homoeologous LGs for each creeping bentgrass parent. This is the expected number of chromosomes for creeping bentgrass (2n=4x=28), and suggests that the markers used in the current study are amplifying in the two different ancestral genomes that make up this species. This pattern of disomic inheritance provides further evidence of the allotetraploid nature of creeping bentgrass, thus supporting previous findings using chromosomal pairing behavior by Jones (1956b), isozyme analysis by Warnke et al. (1998), and genetic linkage map construction by Chakraborty et al. (2005). Additional support for disomic inheritance is suggested by a 1:1 ratio of coupling- and repulsion- phase marker linkages.

High levels of segregation distortion were observed in the cross between 7418-3 and L93-10. Segregation distortion is a common phenomenon that has been shown to affect diverse types of plants (Liang et al., 2006; Lunde et al., 2006). Mechanisms behind segregation distortion are not completely understood, but could include faulty chromosome pairing (Xu et al., 1995), aberrant segregation ratios due to differential transmission in either the male (Mangelsdorf and Jones, 1926; Liedl and Anderson, 1993) or female germline (Yanagihara et al., 1995), or as a result of post-zygotic selection prior to phenotypic evaluation (Gadish and Zamir 1986; Xu et al., 1997). High levels of segregation distortion have been reported in other turfgrasses, including ryegrass (Warnke et al., 2004) and tall fescue (Xu et al., 1995; Saha et al., 2005). The distribution of skewed markers in both parental maps indicates that both male and female parents were involved in segregation distortion.
Our results indicated some syntenic relationships between bentgrass and rice based on CISPs, ILPs, and SSR sequence identity to rice; however, the number of relevant markers is likely too low to draw definitive conclusions. Additional CISP, ILP, or anchor markers from other sources will provide more detail about the colinearity, synteny, and chromosomal rearrangements / duplications between creeping bentgrass and the other major cereal crops that currently have substantial molecular genetic resources. Additional benefits of continuing this work include comparative mapping / analysis with other forage, turf, and cereal grasses with the goal being the identification of putative orthologous loci or QTL with major effects on important traits such as disease resistance, drought tolerance, salinity tolerance, etc… Creeping bentgrass can be clonally propagated, and all parental and pseudo F2 lines have been maintained clonally. This fact, coupled with the PCR based map presented here provides for the possibility of easy integration of additional molecular markers from multiple sources, in any research laboratory interested in creeping bentgrass genetics / genomics.

Future use of this map includes the detection of QTL for resistance / susceptibility to the fungal disease dollar spot (Sclerotinia homoeocarpa F.T. Bennett) and the detection of QTL for drought tolerance (on-going work). Additional uses include leveraging genetic resources from well studied model species, such as rice and other cereal crops, for identification / locating genes or QTL for traits of interest in creeping bentgrass. Identification of QTL will then facilitate MAS in tradional creeping bentgrass breeding programs.
Literature Cited


annual x perennial ryegrass population. Theoretical and applied Genetics. 109:294-304.


BIBLIOGRAPHY


outcrossing grass species perennial ryegrass (Lolium perenne) by quantitative trait loci analysis and comparative genetics. New Phytologist. 178:559-571.


Funk, C. R. and W. A. Meyer. 2001. 70 years of turfgrass improvement a the New Jersey Agricultural Experiment Station: The Garden State's Rutgers University has long been in the forefront of turfgrass development. USGA Green Section Record. 39:19-23.


Heckenberger, M., M. Bohn, D. Klein, and A. E. Melchinger. 2005b. Identification of essentially derived varieties obtained from biparental crosses of homozygous lines: II. Morphological distances and heterosis in comparison with simple sequence repeat and amplified fragment length polymorphism data in maize. Crop Science. 45:1132-1140.


public reference set of SSR markers in Lolium perenne L. Molecular Ecology Notes. 5:951-957.


Lagercrantz, U. 1998. Comparative mapping between Arabidopsis thaliana and Brassica nigra indicates that brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics. 150:1217-1228.


Marcel, T.C., R. K. Varshney, M. Barbieri, H. Jafary, M. J. D. de Kock, A. Graner and R. E. Niks. 2007. A high-density consensus map of barley to compare the


(Puccinia coronata f. sp. loli) resistance in perennial ryegrass (Lolium perenne) using AFLP markers and a bulked segregant approach. Euphytica. 143:135-144.


Sokolov, B. P. 1990. Primer extension technique for the detection of single nucleotide polymorphisms in genomic DNA. Nucleic Acids Research. 18:3671-


Soreng, R. J. 1985. Poa in New Mexico, with a key to middle and southern Rocky Mountain species (Poaceae). Great Basin Naturalist. 45:395-422.

Soreng, R. J. and F. R. Barrie.  1999.  Proposal to conserve the name Poa pratensis (Gramineae) with a conserved type.  Taxon.  48:157-159.


UPOV. 2000. Progress report of the technical committee, the technical working parties and the working group on biochemical and molecular techniques, and DNA-profiling in particular. UPOV Document C/34/10 July 24, 2000.


USGA. 1943. Identity of creeping bent strains planted on experimental greens. Timely Turf Topics. December. p. 3.


types of molecular markers and their use in animal genetics. Genetics Selection
Evolution. 34:275-305.

Vigouroux, Y., J. S. Jaqueth, Y. Matsuoka, O. S. Smith, W. D. Beavis, J. S. Smith, and J.
Doebley. 2002. Rate and pattern of mutation at microsatellite loci in maize.
Molecular Biology and Evolution. 19:1251-1260.


maps and QTLs. Journal of Heredity. 93:77-78.

Vos, P., R. Hogers, M. Reijans, T. van de Lee, M. Hornes, A. Friters, J. Pot, J. Peleman,
M. Kupier, and M. Zabeau. 1995. AFLP: a new technique for DNA

Standardization and application of microsatellite markers for variety identification

The establishment of ‘essential derivation’ among rose varieties, using AFLP.

Wang, D. G., J.-B. Fan, C.-J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N.
Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou,
E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. Shen, D. Kilburn, J.
Rioux, C. Nusbaum, S. Rozen, T. J. Hudson, R. Lipshutz, M. Chee, and E. S.
Lander. 1998. Large-scale identification, mapping, and genotyping of single-

A. Pederson. 2005. Transfer of simple sequence repeat (SSR) markers from
major cereal crops to minor grass species for germplasm characterization and
evaluation. Plant Genetic Resources. 3:45-57.


CURRICULUM VITAE

Joshua Andrew Honig

Education


Professional Experience

Assistant superintendent

January 1996 – December 2001  Rutgers, Department of Plant Science, New 
Graduate Assistant

**Publications**


