PHYSICAL AND GENETIC MAPPING OF THE ASGR-CARRIER CHROMOSOME IN *PENNISETUM SQUAMULATUM*

by

SIRJAN KUMAR SAPKOTA

(Under the Direction of Joann A. Conner)

ABSTRACT

Apospory is a form of gametophytic apomixis in which embryos develop from unreduced embryo sacs derived from nucellar cells of the ovule bypassing meiosis and fertilization of gametes. Apospory in the *Pennisetum and Cenchrus* species is controlled by a physically large, hemizygous, heterochromatic chromosomal block called the apospory-specific genomic region (ASGR). Two major studies were conducted to identify possible synteny of the ASGR-carrier chromosome in *Pennisetum squamulatum* to related and sequenced reference genomes and to use Illumina duplex specific nuclease (DSN) generated sequences to produce additional genetic markers for the ASGR. In the first study, macro-collinearity of the ASGR-carrier chromosome in *P. squamulatum*, outside the ASGR , to chromosome 2 of sorghum and foxtail millet was identified based on physical mapping of eight bacterial artificial chromosome (BAC) clones carrying ASGR-carrier chromosome markers as compared to the location of the orthologous gene on chromosome 2 of sorghum and foxtail millet. In our second study, we were able to genetically map more than 85% of sequence characterized amplified region (SCAR) markers developed from duplex specific nuclease normalized Illumina sequences, to the ASGR in *P. squamulatum*.

INDEX WORDS: apomixis, ASGR, ASGR-carrier chromosome, duplex specific nuclease, SCAR, *Pennisetum*, fluorescence *in situ* hybridization, collinearity

PHYSICAL AND GENETIC MAPPING OF THE ASGR-CARRIER CHROMOSOME IN *PENNISETUM SQUAMULATUM*

by

SIRJAN KUMAR SAPKOTA

B.S., Tribhuvan University, Nepal, 2010

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2014

PHYSICAL AND GENETIC MAPPING OF THE ASGR-CARRIER CHROMOSOME IN *PENNISETUM SQUAMULATUM*

by

SIRJAN KUMAR SAPKOTA

Major Professor:	Joann A. Conner
Committee:	Cecilia E. McGregor
	Brian M. Schwartz

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
December 2014

DEDICATION

This thesis is dedicated to my parents and grandparents for the wonderful childhood I had and the support I got throughout those days, and many more to come.

ACKNOWLEDGEMENTS

I would like to thank Dr. Joann Conner for her support and constant advisement. I believe the skills and techniques you taught me will take me far in my career and many thanks to you for that. My appreciation goes to Dr. Peggy Ozias-Akins for her advice and help for my projects. I would also like to acknowledge the advices and guidance of my committee members, Dr. Cecilia McGregor and Dr. Brian Schwartz. I cannot overstate my appreciation for the lab members of the Ozias-Akins lab; Juliet (Dr. Ye Chu), Miss Evelyn, Josh, Greg, Keun, Thiago, Stephenie, Marian, Maricel, thanks a lot for creating such a nice and supportive environment where I could learn and grow more into science. Kathleen, I cannot thank you enough for your support and help throughout. I acknowledge all the funding sources which made my education and degree acquirement possible. Last but not the least, dad, mom and sisty you guys are who keeps me going now and for years to come. I missed you guys all these years.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

Sexual Reproduction and Apomixis in Flowering Plants 1

Significance of Apomixis 2

Developmental Mechanisms and Genetic Control of Apomixis 3

Apomixis in *Pennisetum* and *Cenchrus* 15

Comparative Genomics and Collinearity in Grass Family 21

Next-Generation Sequencing Technology and Duplex Specific Nuclease 23

References 25

2 PHYSICAL MAPPING OF THE ASGR-CARRIER CHROMOSOME REVEALS COLLINEARITY TO CHROMOSOME 2 OF SORGHUM AND FOXTAIL MILLET OUTSIDE THE ASGR 42

Abstract 43

Introduction 44

Materials and Methods 47

Results 54

Discussion 59

References 66

3 GENETIC MAPPING OF SEQUENCE CHARACTERIZED AMPLIFIED REGION MARKERS IN *PENNISETUM SQUAMULATUM* AND *CENCHRUS CILIARIS* USING DUPLEX SPECIFIC NUCLEASE NORMALIZED ILLUMINA SEQUENCES 90

Abstract 91

Introduction 92

Materials and Methods 96

Results 103

Discussion 108

References 115

4 SUMMARY 152

LIST OF TABLES

Page

Table 2.1: Primers designed for isolation of the ASGR-carrier chromosome BACs 72

Table 2.2: BlastN analysis of the ASGR-carrier chromosome transcripts with corresponding BAC information 73

Table 3.1: Primers designed for mapping 133

Table 3.2: Number of 100 bp Illumina sequence reads generated from the DSN treated libraries 136

Table 3.3: Statistics from the Velvet assemblies of BC_8 line 58 apomictic DSN sequences using different K-mer lengths 137

Table 3.4: Contigs selected for SCAR primer design 138

Table 3.5: Number of contigs showing ≥ 99% sequence identity against ASGR-BAC and ASGR-recombinant BAC clone sequences 142

Table 3.6: Summary of mapping results 144

Table 3.7: Contigs chosen and location of SCAR primers for possible gene expression analysis 146

LIST OF FIGURES

Page

Figure 2.1: Example of PCR confirmation of BAC clones containing the ASGR-carrier chromosome SCAR primers 77

Figure 2.2: DNA fragments generated from a *Hae*III digestion of BAC clones with strong hybridization signals with the 160 bp *Kpn*I centromeric repeat sequences 78

Figure 2.3: Assessment of repetitiveness of the ASGR-carrier chromosome BACs by Southern blot 79

Figure 2.4: Southern blot of the ASGR-carrier chromosome BACs with apomictic BC_8 line 58 DNA 80

Figure 2.5: Physical mapping of the ASGR-carrier chromosome BAC clones p142D19, p236E19, p258L05 and p057M05 on metaphase chromosome spreads of apomictic BC_8 line 58 root tips 81

Figure 2.6: Physical mapping of the ASGR-carrier chromosome using BACs p036L06 and p220A02 on metaphase chromosome spreads of apomictic BC_8 line 58 root tips 83

Figure 2.7: Physical localization of the ASGR-carrier chromosome BACs p285J18 and p181G17 on metaphase chromosome spreads from apomictic BC_8 line 58 root tips 85

Figure 2.8: Graph showing enrichment of apomictic but not sexual signal on sorghum chromosome 2 during CGH analysis 87

Figure 2.9: Collinear alignments between the ASGR-carrier chromosome and chromosome 2 of sorghum and foxtail millet 88

Figure 3.1: Count vs. abundance of 31 K-mers for BC_8 line 58 apomictic control and DSN treated sequences 147

Figure 3.2: Count vs. abundance of 31 K-mers for BC_8 line 58 sexual control and DSN treated sequences 148

Figure 3.3: Examples for pre-screening of SCAR primers for mapping to the ASGR and the ASGR-carrier chromosome 149

Figure 3.4: Example of mapping ASGR-linked SCAR primers to a larger R5-F_1 population....150

Figure 3.5: Mapping of SCAR primers from contig *A7770001.min7_55_Ctg232696* in F_1 population 151

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Sexual Reproduction and Apomixis in Flowering Plants

Sexual reproduction is the most widespread form of reproduction in flowering plants. In sexual reproduction, the life cycle of the plant alternates between a diploid sporophytic phase and a haploid gametophytic phase (REISER and FISCHER 1993; GROSSNIKLAUS and SCHNEITZ 1998). During sexual reproduction, the diploid sporophyte develops into a haploid gametophyte that produces haploid spores through meiosis. The male and female spores are produced by microsporogenesis and megasporogenesis, respectively. Megasporogenesis results into a tetrad of megaspores from meiosis of the megaspore mother cell. One of the megaspores survives and becomes functional while the other three megaspores degenerate. In microsporogenesis all four microspores resulting from meiosis of the microspore mother cell survive and are functional. The gametophytes are developed from the mitotic cell proliferation and differentiation of the respective spores. Male gametes are produced in the male gametophyte, a product of the anther. Female gametophytes or sexual embryo sacs, a product of the ovule, are the site for development of the egg cells or female gametes. A polygonum-type embryo sac, the most common type of female gametophyte, comprises one egg cell, two synergid cells, one binucleate central cell and three antipodal cells (REISER and FISCHER 1993). In sexual reproduction, the egg cell fuses with the sperm cell to produce a zygote that develops into the embryo, while endosperm is produced by fertilization of the central cell by other sperm cell.

Although sexual reproduction is ubiquitous in flowering plants, apomixis has been observed in more than 400 genera of almost 40 families with numerous apomictic species in families from Poaceae, Asteraceae and Rosaceae (CARMAN 1997; BICKNELL and KOLTUNOW 2004; OZIAS-AKINS and VAN DIJK 2007). Apomixis is defined as an asexual form of reproduction through seed (NOGLER 1984). The term apomixis was introduced by Winkler (1908) to mean "Substitution of sexual reproduction by an asexual multiplication process without nucleus and cell fusion" (BICKNELL and KOLTUNOW 2004). In apomictic reproduction, unlike sexual reproduction, the embryo develops from unreduced cells of the ovule without meiotic division and fertilization of gametes. The endosperm in apomictic seeds can, however, be generated either autonomously (autogamy) or through fertilization of polar nuclei (pseudogamy) (HAND and KOLTUNOW 2014).

Significance of Apomixis

Since plants derived from apomictic seeds are genetically identical to the maternal plant, apomixis is seen as a potential mechanism to fix hybrid vigor in crop species and maintain specific genotypes through successive seed generations in agriculture breeding (HANNA and BASHAW 1987; HANNA 1995; SAVIDAN 2000). Preliminary analyses on potential benefits of introducing apomixis in major crops have suggested substantial overall welfare gains associated with apomictic crops (SAVIDAN *et al.* 2001; KAUSHAL *et al.* 2004; FRISVOLD *et al.* 2005). Although introduction of apomixis to major crops would provide a cost-effective alternative to various breeding practices used to produce hybrid seed, apomixis has not been observed in major grain crops and attempts to integrate the trait by interspecific hybridization have had limited success (DUJARDIN and HANNA 1989b; SAVIDAN 2000; OZIAS-AKINS *et al.* 2003; HAND and KOLTUNOW 2014). The introgression of apomixis into sexual species has been hindered by the

limited number of apomictic species with significant relatedness to sexual crops and ploidy barriers during apomictic crop breeding (DUJARDIN and HANNA 1984; DUJARDIN and HANNA 1989a; HÖRANDL and TEMSCH 2009). A potential alternative to development of apomictic crops may be the introduction of one or more well characterized apomictic genes into sexual plants through genetic engineering. Recent exploration of the molecular mechanisms underlying apomixis in several species might eventually enable crop breeding to produce apomictic crops (KOLTUNOW *et al.* 1995; GROSSNIKLAUS 2001; OZIAS-AKINS 2006; HEQIANG *et al.* 2009).

Developmental Mechanisms and Genetic Control of Apomixis

Types of Apomixis

Although the actual genetic mechanism of apomixis is still undetermined, the study of apomixis in several species across different genera has expanded our knowledge of apomixis. Apomictic mechanisms are historically subdivided into two categories and classified as either sporophytic or gametophytic based on whether the embryo develops from a gametophyte or directly from unreduced somatic (sporophytic) cells within the ovule (NOGLER 1984; KOLTUNOW 1993; HAND and KOLTUNOW 2014).

Sporophytic apomixis occurs late in ovule development when unreduced somatic ovule cells surrounding the embryo sac differentiate, acquire an embryogenic cell fate and start developing as adventitious embryos. Typically adventitious embryos form from nucellar cells while integumentary originated adventitious embryos are much less common (KOLTUNOW 1993; OZIAS-AKINS 2006). Polyembryony is frequently observed in sporophytic apomicts because sporophytic apomixis coexists with the angiosperm sexual pathway within the same ovule (KOLTUNOW *et al.* 1995). Endosperm resulting from the sexual pathway can provide nutrients for either the survival of the adventitious embryo or both the sexual and apomictic derived embryos

(OZIAS-AKINS 2006; DE MEEÛS *et al.* 2007). When fertilization of the central cell is not successful, adventitious embryos can obtain nutrients for their development from the degrading nucellar and integument cells (KOLTUNOW 1993). Sporophytic apomixis appears to be genetically complex and quantitatively inherited (GARCÍA *et al.* 1999). A haplotype-specific physical map covering a locus strongly associated with polyembryony in *Citrus* was constructed through screening of a *Citrus* BAC library (NAKANO *et al.* 2012). Genetic mapping and association analysis have characterized the candidate genomic region for polyembryony to be approximately 300 Kb containing 70 predicted open reading frames. Single nucleotide polymorphism (SNP) detected in the polyembryony genomic region showed strong association with embryo type in *Citrus* cultivars, indicating that a common polyembryony genomic region shared among widely diverse *Citrus* cultivars and species may play major role in determination of either the polyembryony or monoembryony in *Citrus* species (NAKANO *et al.* 2012).

In gametophytic apomixis, the embryo sac is mitotically formed from an unreduced cell in the ovule and therefore bypasses meiosis. This process of mitotic embryo sac development is termed apomeiosis and is the first step of the apomictic process. In the second step of the process, the egg cell of the unreduced embryo develops into an embryo without fertilization and is termed parthenogenesis. Endosperm development in gametophytic apomixis can occur either by fertilization of the central cell by the sperm cell of the pollen or autonomously by fertilization-independent means. Autonomous endosperm formation, if occurring, is considered the third step of apomixis (GRIMANELLI *et al.* 2001; OZIAS-AKINS 2006; HAND and KOLTUNOW 2014). Gametophytic apomixis can be further subdivided into diplospory and apospory (GRIMANELLI *et al.* 2001; OZIAS-AKINS 2006; HAND and KOLTUNOW 2014).

In diplospory, the megaspore mother cell develops into an unreduced female gametophyte by variable developmental processes prior to diplosporous embryo sac initiation. The megaspore mother cell may either be inhibited from starting meiosis or initiate meiosis but arrest during meiosis I thereby keeping the somatic cell chromosome number. Mitotic divisions form an unreduced gametophyte that is not visually distinct from a sexually derived embryo sac. One cell of the diplosporous embryo sac is specified to perform the function of an egg cell, although it is chromosomally unreduced compared with meiotically derived egg cells (GRIMANELLI *et al.* 2001; HAND and KOLTUNOW 2014). It is still unclear when and how the megaspore mother cell becomes diplosporous or the mechanisms of inhibition of meiosis or arrest of meiosis that are crucial for understanding diplospory.

In contrast to diplospory, apospory involves development of unreduced embryo sacs directly from mitosis, originating not from megaspore mother cell, but from nucellar cell(s) adjacent to megaspore mother cell (KOLTUNOW 1993; HAND and KOLTUNOW 2014). These nucellar cells are referred to as aposporous initials and they usually possess large nuclei and dense cytoplasm. The aposporous initial cell undergoes mitosis to form the unreduced embryo sac. Meiosis of the megaspore mother cell can either be completed or be arrested during the process, but in both cases the products of sexual reproduction usually degenerate in plants that almost completely reproduce by apomixis (obligate apomicts). In facultative apomixis, sexual and apomictic embryo sacs are both produced side by side in the apomictic plant. The number of aposporous embryo sacs formed in an ovule differs between aposporous species and even within genotypes of a species (NOGLER 1984). In most aposporous species the unreduced egg cell is rarely fertilized, but in *Hypericum perforatum* and *Arabis holboellii* fertilization of unreduced egg cells have been reported (MATZK *et al.* 2001; NAUMOVA *et al.* 2001). Endosperm

development in apospory usually requires pollination and fertilization of the central cell (pseudogamy).

Endosperm Development

During sexual reproduction, 2C:3C and 4C:6C DNA content of embryo:endosperm is observed during G1 and G2 phases of the cell cycle, respectively. Seed viability in individuals that reproduce by sexual means is affected by the maternal to paternal genome contributions to endosperm, while apomicts are either more tolerant of variation in the maternal:paternal genome ratio or have adaptations to restore such ratios (GRIMANELLI *et al.* 1997; OZIAS-AKINS 2006). *Pennisetum* and other panicoid apomicts display this kind of adaptation. The four nucleate aposporous embryo sacs have a uninucleate central cell that is fertilized to create 2n+n endosperm which reconstitutes a 2:1 maternal to paternal genome ratio (OZIAS-AKINS *et al.* 2003; OZIAS-AKINS 2006). A change in this pattern was observed, a drop in the frequency of the uninucleate central cell in favor of binucleate central cells, when apomixis was introgressed into sexual *P. glaucum* from an apomictic relative. This change was found to correlate with a reduction in seed set (MORGAN *et al.* 1998).

Genetic Control and Inheritance of Apomixis

The presence of varied apomictic mechanisms and phylogenetic analyses of apomicts in different genera and families imply that apomixis has evolved independently multiple times (CARMAN 1997; VAN DIJK and VIJVERBERG 2005). With the exception of *Tripsacum*, where diplospory is determined by a single genetic locus (KINDIGER *et al.* 1996), the majority of diplosporous species such as *Erigeron* and *Taraxacum* are found to be controlled by two independent loci for apomeiosis and parthenogenesis (VAN DIJK *et al.* 1999; NOYES and RIESEBERG 2000). Apospory is transmitted as a dominant trait and controlled by a single genetic

locus in *Pennisetum*, *Brachiaria*, *Panicum*, and *Paspalum* species (PESSINO *et al.* 1997; OZIAS-AKINS *et al.* 1998; MARTÍNEZ *et al.* 1999; EBINA *et al.* 2005). In aposporous *Poa pratensis* and *Cenchrus ciliaris*, recombination has been observed showing separation in loci controlling apomeiosis and parthenogenesis (ALBERTINI *et al.* 2001; CONNER *et al.* 2013). Genetic studies in *Hieracium*, capable of fertilization-independent endosperm formation, revealed that this trait also segregates independently of the other two apomictic components (OGAWA *et al.* 2013).

The genetic analysis of apomicts via genetic crossing experiments with sexual plants show that apomixis is inherited as a dominant trait with simple inheritance involving a few Mendelian genes (DUJARDIN and HANNA 1989b; LEBLANC *et al.* 1995; OZIAS-AKINS and VAN DIJK 2007; BARCACCIA and ALBERTINI 2013; HAND and KOLTUNOW 2014). In contrast, molecular and cytogenetic analyses of the chromosomal region(s) of the various apomixis locus of several species have revealed a rather complex structure (GRIMANELLI *et al.* 1998; MARTINEZ *et al.* 2003; OZIAS-AKINS *et al.* 2003; AKIYAMA *et al.* 2004). The complex structure of the apomictic loci is likely based on the need to link multiple genes which could involve mechanisms such as reduced recombination at the locus with DNA rearrangements or supernumerary chromatin structures (OZIAS-AKINS *et al.* 2003; PUPILLI and BARCACCIA 2012). In *Poa pratensis*, based on observation of discrete classes of expressivity, a model with five major genes required to control unlinked events of apospory initiation and parthenogenesis was proposed (MATZK *et al.* 2005). The apomixis controlling locus (ACL) in *Paspalum* showed strong suppression of recombination and large-scale rearrangements due to transposable elements, based on comparison to a syntenic region in rice (CALDERINI *et al.* 2006). Chromosomes carrying the ASGR in *Pennisetum* and the *LOSS OF APOMEIOSIS (LOA)* locus in *Hieracium* are associated with extensive repetitive sequence and transposon-rich regions

(AKIYAMA *et al.* 2004; OKADA *et al.* 2011). In *Hieracium*, a recent study showed that a progeny arising from a cross between sexual *H. pilosella* and apomictic *H. praealtum*, that are recombinant for *LOA*-linked markers, still remains apomictic despite having lost the extensive repetitive sequence structure associated with the locus (KOTANI *et al.* 2014). The structure of the apomixis locus in *Pennisetum* and *Cenchrus* will be discussed later in this chapter.

Sexual reproduction has been shown to be the default pathway in apomictic *Hieracium praealtum* by characterization of gamma deletion mutants (KOLTUNOW *et al.* 2011). A series of deletion mutants lacking the apomeiosis locus (called *LOSS OF APOMEIOSIS* or *LOA*) and the parthenogenesis locus (called *LOSS OF PARTHENOGENESIS* or *LOP*) were developed (CATANACH *et al.* 2006). Endosperm formation in *H. praealtum* is fertilization independent, segregates independent of apomeiosis and parthenogenesis, and is controlled by a separate locus for autonomous endospermy. Deletion of either *LOA* or *LOP* alone results in the return of the sexual pathway for that component. Plants with *LOP* deletions resulted in development of unreduced embryo sacs with unreduced egg cells which did not form embryos when pollination was prevented, but formed hybrid seeds with increased ploidy when pollinated with a tetraploid tester plant. With *LOA* mutants, the megaspore mother cell progressed through meiosis with the egg and central cell forming a reduced embryo and endosperm, respectively in the absence of fertilization. Deletion of both the *LOA* and *LOP* loci resulted in complete reversion to sexual development of seed (CATANACH *et al.* 2006; KOLTUNOW *et al.* 2011). This suggests that in *Hieracium*, the fate of cells with gametic potential during apomixis may be redirected rather than being a detached and completely independent pathway (HAND and KOLTUNOW 2014).

Polyploidy and Apomixis

Most apomicts are polyploid in nature. The presence of diploid gametophytic apomicts suggests that polyploidy is not absolutely required for the expression of apomixis (BICKNELL 1997; KOJIMA and NAGATO 1997; KANTAMA *et al.* 2007). Nevertheless, the frequency of unreduced embryo sacs and seed formation in diploid apomicts was much lower compared to the polyploid apomicts, suggesting that the expression of apomixis may have been enhanced by higher ploidy level in apomicts (BICKNELL and KOLTUNOW 2004; OZIAS-AKINS 2006). It has been reported that simple chromosome doubling of sexual diploid *Paspalum* produced apomictic autotetraploids (QUARIN *et al.* 2001), suggesting that changes brought on by genome duplication might be involved in apomixis induction. Studies have shown that changing ploidy level of plants changes the expression of reproductive regulatory genes (GUO *et al.* 1996; COMAI *et al.* 2000; EDGER and PIRES 2009). However, it is clear that polyploidy alone is not sufficient for apomixis based on the fact that more than half of the angiosperm species are polyploid but most of them reproduce sexually (GRIMANELLI *et al.* 2001). Apomixis has been argued to be transmitted preferentially through diploid gametes rather than haploid gametes (GRIMANELLI *et al.* 1998; TAS and VAN DIJK 1999). This establishment suggests the increased genome size of apomicts might be the consequence of apomixis rather than the cause for apomixis.

Candidate Apomixis Genes

Several efforts have been made to identify genes that have a role in apomixis. Differential gene expression analysis between apomixis and closely related sexual genotypes has been used for identification of genes differentially regulated in apomictic and sexual pathways. The advent of next generation sequencing has been crucial in identification of new clues to the differential

transcriptional pathways involved in apomictic and sexual species (HAND and KOLTUNOW 2014). Transcriptome comparisons between apomictic and closely related sexual individuals in reproductive tissues at different developmental stages have revealed several mechanisms that are central to the apomictic process like protein degradation, transcription, stress response, and cell-to-cell signaling (ALBERTINI *et al.* 2005; LASPINA *et al.* 2008; SHARBEL *et al.* 2010; SILVEIRA *et al.* 2012; OKADA *et al.* 2013).

Transcriptome profiles of microdissected live ovules at four developmental stages between a diploid sexual and diploid apomictic *Boechera* were compared by sequencing >2 million SuperSAGE tags (SHARBEL *et al.* 2010). Most heterochonic tags and stage-specific tags were significantly downregulated during early stages of apomictic ovule development, while most late stage-specific tags were upregulated in the apomictic ovules showing that the apomixis-specific gene expression is characterized by overrepresentation of transcription factor activity (SHARBEL *et al.* 2010). While not identifying any individual candidate gene(s), the authors hypothesize that apomeiosis is correlated with a global downregulation of gene expression during megaspore mother cell formation (SHARBEL *et al.* 2010; HAND and KOLTUNOW 2014). *APOLLO* (for apomixis-linked locus) gene, an Aspartate Glutamate Aspartate Aspartate histidine exonuclease whose transcripts are down-regulated in sexual ovules entering meiosis while being upregulated in apomeiotic ovules at the same stage of development in apomict, has been identified in *Boechera*. It was suggested that a single deregulated allele (apoallele) could have a role in inducing the cascade of events leading to asexual female gamete formation in an apomictic *Boechera* plant (CORRAL *et al.* 2013).

Laser capture microdissection was used to collect specific cell types from different stages of apospory in aposporous *Hieracium praealtum* for comparative transcriptome analyses.

Transcriptome and protein motif analyses showed that the vast majority of sequences were homologous to *Arabidopsis* genes which were expressed in *Arabidopsis* during early embryo development. Gene expression indicative of meiosis were notably absent in enlarging aposporous initials cells. The transcriptomes from both aposporous initial and early aposporous embryo sac were similar and shared expressed genes involved in gametophyte development suggesting enlarging aposporous initial cells transition to embryo sac development prior to entering meiosis (OKADA et al. 2013).

In *Brachiaria brizantha*, expression analysis of expressed sequence tag (EST) from ovaries of sexual and apomictic individuals showed the presence of nine expressed sequences preferentially detected in ovaries of the apomicts as confirmed by RT-qPCR. Two expressed sequences, *BbrizSti1* and *BbrizHelic*, showing putative involvement in early steps of Panicum-type embryo sac differentiation were identified to have similarity to a stress induced protein and a helicase, respectively (SILVEIRA *et al.* 2012). A MADS-box gene from *B. brizantha*, named *BbrizAGL6*, was localized to the megaspore mother cell of ovaries from sexual and apomictic plants and additionally to the nucellus of apomictic plants where the aposporous initials differentiate (GUIMARÃES *et al.* 2013).

In *Poa pratensis*, a cDNA-AFLP (amplified fragment length polymorphism) strategy was applied for identification and characterization of two candidate genes: *SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE* (*PpSERK1*) and *APOSTART* (ALBERTINI *et al.* 2004; ALBERTINI *et al.* 2005). *PpSERK* is a tyrosine kinase proposed as the switch that allows aposporous initial cells to form and develop into an embryo sac. In *P. pratensis*, *PpSERK* was found to be expressed in cells neighboring the megaspore mother cell (ALBERTINI *et al.* 2005). The ortholog of *APOSTART* in *Arabidopsis* is expressed in mature female embryo sacs, and the

phenotype of mutants for this gene suggests the role of this gene in embryo and seed development (BARCACCIA and ALBERTINI 2013).

A differential display analysis was carried out on immature inflorescences of sexual and apomictic tetraploid genotypes of *Paspalum notatum* to identify genes associated with apospory initiation. From 94 differentially expressed sequences, a total of 65 candidate unigenes were isolated from assembly analysis with 45 unigenes functionally characterized. The differentially expressed sequences appeared to be involved in several cellular processes like signal transduction cascade, cell-cycle control, protein turnover, transcription regulation, transposon activity and endoplasmic reticulum-mediated biosynthesis (LASPINA *et al.* 2008). In a recent study in *P. notatum*, expression of *PnSERK2*, a member of the *SERK* family, seems to be correlated with the onset of apomixis. *PnSERK2* was identified to be expressed in nucellar cells of the apomictic genotype at meiosis, but in sexual genotype it was expressed only in the megaspore mother cell (PODIO *et al.* 2014).

The *BABY BOOM (BBM)* gene, a member of the AP2/ERF multigene family, was identified as a marker for embryo development in *Brassica napus* microspore-derived embryo cultures. Overexpression of *BnBBM* in *Arabidopsis* induced spontaneous somatic embryo formation from vegetative tissue (BOUTILIER *et al.* 2002). Sequence analysis of partially sequenced bacterial artificial chromosome (BAC) clones from the apospory-specific genomic region (ASGR) in *P. squamulatum* and *C. ciliaris* identified 40 putative protein coding regions. Putative protein coding regions sharing sequence similarity to the rice *BABY BOOM (BBM*) gene were identified (CONNER *et al.* 2008). Comparative analysis of the expression of the *ASGR-BBM-like* gene in *C. ciliaris* through semi-quantitative RT-PCR conducted with aposporous ovaries from different stages indicated that the transcription of the *CcASGR-BBM-like* genes was

initiated before pollination and up-regulated upon pollination with a significant increase one day after pollination. However, another copy of the *BBM-like* gene (*CcN-ASGR-BBM-like*), similar to the *ASGR-BBM-like* gene in sequence but genetically unlinked to the ASGR, showed no expression until one day after pollination in aposporous *C. ciliaris* with transcription levels lower than that of *CcASGR-BBM-like* gene at one day after pollination in the apomictic plant. The differential expression patterns of *CcASGR-BBM-like* and *CcN-ASGR-BBM-like* genes in aposporous ovaries suggests that the *ASGR-BBM-like* gene could play a role in initiation and/or maintenance of parthenogenesis of aposporous embryo in *C. ciliaris* (ZENG 2009). Additionally, the A8 plant, a *C. ciliaris* recombinant plant that is capable of producing unreduced aposporous embryo sacs but has lost the ability to undergo parthenogenesis, has lost the UGT197 marker that carries the *ASGR-BBM-like* gene (CONNER *et al.* 2013).

In *Hypericum perforatum*, a BAC clone from the *HYPERICUM APOSPORY* (*HAPPY*) locus was found to contain an ubiquitin-mediated E3 ligase gene. This *ARIADNE* 7-like E3 ligase (*HpARI*) has been postulated as a apomictic candidate gene from the *HAPPY* locus, as one of the four *HpARI* alleles within the tetraploid apomict is truncated as compared to the sexual plants (SCHALLAU *et al.* 2010). E3 ligases have a role in embryo sac development. In maize, MATH-BTB protein MAB interacts with an E3 ubiquitin ligase component (Cullin 3a) and is required for mitotic spindle function and nuclear fate in developing embryo sacs (JURANIC *et al.* 2012). Similarly in Arabidopsis, anaphase promoting complex/cyclosome, a multisubunit E3 ligase, plays a critical role in cell cycle control during female gametogenesis and embryogenesis (WANG *et al.* 2013). Alteration in expression or function of the *HpARI* E3 ligase might affect embryo sac development in *Hypericum*, leading to apomixis (HAND and KOLTUNOW 2014).

Possible Role of Deregulated/Mutated Sexual Gene and Apomixis

A comparative analysis of gene expression patterns of three members of the *FERTILIZATION-INDEPENDENT SEED (FIS)* class genes *MEDEA (MEA/FIS1), FIS2,* and *FERTILIZATION-INDEPENDENT ENDOSPERM (FIE/FIS3)* in *Hieracium* showed that the spatial and temporal expression patterns in apomictic *Hieracium* were comparable to those observed in embryo sacs, embryos, and endosperm in sexual *Hieracium* (TUCKER *et al.* 2003). Mutation studies in sexual species like *Arabidopsis* and maize have shown that mutations in genes with roles in meiosis or in epigenetic regulation during ovule-cell fate specification and embryo sac development can lead to apomeiotic phenotype in the progeny (BARRELL and GROSSNIKLAUS 2005; RAVI *et al.* 2008; GARCIA-AGUILAR *et al.* 2010; RAVI and CHAN 2010). A triple mutant in *Arabidopsis,* called *MiMe (Mitosis instead of Meiosis)* combines mutants in the *omission of second division 1* (*Osd1*), *Atspo11-1,* and *Atrec8* genes and produces viable diploid gametes that are genetically identical to their mother (D'ERFURTH *et al.* 2009). Auxin and cytokinin signaling can also affect the specification of the functional megaspore in *Arabidopsis.* Cytokinin mutants showed disrupted embryo sac development (CHENG *et al.* 2013). Similarly, the auxin efflux carrier gene *PIN-FORMED 1 (PIN1)* is necessary for embryo sac development in *Arabidopsis* (CECCATO *et al.* 2013). The timing, position and frequency of aposporous initial cell formation in *Hieracium* was affected by inhibition of polar auxin transport (TUCKER *et al.* 2012). *ARGONAUTE* genes contain small RNA that cleave mRNA and lead to RNA dependent DNA methylation (MEISTER 2013). Mutation in *ARGONAUTE* genes, *ago9* and *AGO104,* can lead to the development of apomictic-like phenotypes in *Arabidopsis* and maize, respectively (OLMEDO-MONFIL *et al.* 2010; SINGH *et al.* 2011). *AGO* genes are also known to have a role in germ-cell identity in rice. For example, *MEIOSIS ARRESTED AT LEPTOTENE1* (*MEL1*) is an

ARGONAUTE gene expressed specifically in gamete precursor cells within ovules of rice and regulates the cell division of premeiotic germ cells and meiosis in these cells (NONOMURA *et al.* 2007).

Apomixis in *Pennisetum* and *Cenchrus*

Pennisetum squamulatum is a polyploid wild relative of pearl millet and was used to introgress apospory into sexual pearl millet. The introduction of apomixis from *P. squamulatum* began with the production of allopolyploids F_1 hybrids from crosses between induced tetraploid *P. glaucum* (2n=4x=28) × *P. squamulatum* (2n=56) and *P. glaucum* (2n=14) × *P. purpureum* (2n=28) followed by chromosome doubling. The hybrid F_1s were crossed (2n=42) and then backcrossed with tetraploid *P. glaucum* (2n=4x=28), leading to a recovered backcross 3 (BC_3) plant showing near obligate apomixis and containing 29 chromosomes (DUJARDIN and HANNA 1989b). Additional apomictic backcross generations were subsequently produced. The transfer of apomixis from *P. squamulatum* to pearl millet has been hindered by drag of undesired characteristics such as poor seed set and low pollen fertility (SINGH *et al.* 2010). The morphological characteristics of apomictic BC_8 lines were not significantly different from BC_7 lines; however, BC_8 lines were found to have higher pollen viability, earlier flowering time, and higher seed set compared to BC_7 lines. While no meiotic embryo sacs were identified in BC_7 plants, the frequency of meiotic embryo sacs in BC_8 lines ranged from 0 to 37%. The frequency of germinating seedlings produced by apospory, as assayed by inheritance of apospory linked markers, showed that BC_7 germinating seedlings had a higher frequency of apospory (98%) than BC_8 (73-85%) (SINGH *et al.* 2010).

The genetics and evolution of apospory in *P. squamulatum* and closely related species *C. ciliaris* has been studied. A genetic study using randomly amplified polymorphic DNA (RAPD)

followed by the development of 12 genetically clustered sequence characterized amplified region (SCAR) markers in *Pennisetum* resulted in the designation of the apospory-specific genomic region (ASGR) (OZIAS-AKINS *et al.* 1998). The mapping population used was F_1s produced from a cross of induced tetraploid *P. glaucum* with apomictic donor *P. squamulatum*. The obligate apomictic parent *P. squamulatum* could only be used as the male parent, as chromosomally reduced gametes carrying potentially recombinant products of meiosis are produced during male gametogenesis. The scoring was based on embryo sac formation from microscopic examination of cleared ovules and the individuals in the population were scored as sexual or aposporous (obligate or facultative). The aposporous embryo sac in *Pennisetum* is panicum-type and can be distinguished from polygonum-type sexual embryo sac by the absence of antipodal cells. Mode of reproduction was confirmed by examining progeny produced after pollination of the F_1 plants with pollen from tetraploid pearl millet homozygous for a dominant "red" marker gene; the presence of red progeny indicated sexual reproduction, and the exclusive development of green progeny was a strong indication of the linkage between apomeiosis and parthenogenesis indicating apomictic reproduction. The ASGR shows hemizygosity at the locus as some molecular probes only hybridized to plants which are apomictic (OZIAS-AKINS *et al.* 1998). This result suggested that evolution of apomixis in *Pennisetum* did not appear to have resulted from a single genetic mutation in the sexual developmental pathways for *Pennisetum*, and that apomictic and sexual reproduction were not allelic alternatives. Another study was conducted to test those 12 tightly linked SCAR markers in a population of *C. ciliaris*, phenotyped for aposporous embryo sac formation (ROCHE *et al.* 1999). An ‘off type’ plant was discovered in a field of apomictic *C. ciliaris* and reproductive analysis found it to be a rare sexual *C. ciliaris* plant. As the sexual plant remained fertile and a tetraploid, it could be used as the female in

genetic crossing experiments. Six of the ASGR SCAR markers were found to be dominant markers between sexual and aposporous *C. ciliaris* genotypes. Five of the SCARs were always linked to apospory, with one marker showing a low level of recombination in the 84 F_1 plants tested. Restriction fragment length polymorphisms (RFLP) were observed in three of the remaining six ASGR markers from *P. squamulatum* when used as probes and were also linked to apospory in *C. ciliaris* (ROCHE et al. 1999).

Two Bacterial Artificial Chromosomes (BAC) clones library were developed to study the ASGR in these species (ROCHE *et al.* 2002). The libraries comprised a total of 118,272 clones with average insert size of 82 kb, and 68,376 clones with average inert size of 109 kb from a *P. squamulatum* x pearl millet apomictic polyhaploid genotype and buffelgrass (*C. ciliaris*), respectively. BAC clones containing the ASGR-linked SCAR markers were identified. BAC clones containing molecular markers (AFLPs, SCARs, and RFLPs) showing total genetic linkage to the apomictic phenotype are referred to as ASGR-linked BACs. One BAC clone containing an AFLP marker segregating 2 cM from the apomictic phenotype in *P. squamulatum* and is referred to as the ASGR-recombinant BAC (GOEL *et al.* 2006). Analysis of the BAC clones indicated that even SCAR markers that looked single copy via RFLP analysis could be duplicated at the ASGR and that ASGR-linked BAC clones between the two species were highly conserved at the restriction enzyme level. The ASGR-linked and ASGR-recombinant BAC clones were physically mapped to a chromosome by fluorescence *in situ* hybridization (FISH) in *P. squamulatum* and *C. ciliaris* (AKIYAMA *et al.* 2004; AKIYAMA *et al.* 2005; GOEL *et al.* 2006). In *P. squamulatum*, FISH analysis confirmed that the ASGR is physically large (~50Mb), has a telomeric location on the ASGR-carrier chromosome and is hemizygous and heterochromatic in nature (AKIYAMA *et al.* 2004), while in *C. ciliaris* the ASGR was found to be interstitial and

inverted in marker order relative to *P. squamulatum* (AKIYAMA *et al.* 2005; GOEL *et al.* 2006). The presence of a centromere-derived repeat signal at the distal end of ASGR in *P. squamulatum* suggests a possible inversion of the chromosome arm on the ASGR-carrier chromosome. Two regions with different FISH signals were observed from physical mapping of the ASGR-linked BAC clones. An approximately 13 Mb "low copy" region showed one or two intense FISH signal(s), while the "high copy" region within ASGR-linked BAC clones showed dispersed FISH signals flanking the low copy region (AKIYAMA *et al.* 2004). BACs with high copy FISH signals of the ASGR in *P. squamulatum* were abundant in Opie-2-like retrotransposons sequences. In *C. ciliaris* Opie-2 like sequences were associated with the centromeric regions of all chromosomes (AKIYAMA *et al.* 2004; AKIYAMA *et al.* 2005). Since the Opie-2 like repeat has been detected as part of the ASGR in seven out of 12 aposporous species, it is speculated that the association of this repeat to the ASGR was derived either by translocation of repeat-poor portion of the ASGR to repeat-rich region in the genome or by transposition and accumulation of retrotransposons in proximity to the ASGR (AKIYAMA *et al.* 2011). While the ASGR is shown to be highly conserved and macrosyntenic between *P. squamulatum* and *C. ciliaris*, subtle differences have been seen between the two species indicating local disruptions of microsynteny. The SCAR marker A14M was absent in *C. ciliaris* but the polyhaploid BAC containing this marker (P001) is syntenic in *C. ciliaris* relative to its position in *P. squamulatum*. Another SCAR marker, A10H, present in *C. ciliaris* did not recover a BAC that could be localized in *C. ciliaris* and the polyhaploid BAC (P1200), adjacent to A10H-containing BAC in *P. squamulatum*, also could not produce any detectable FISH signal in *C. ciliaris* (ROCHE *et al.* 1999; GOEL *et al.* 2006). These observations suggest a deletion of this region of the ASGR in *C. ciliaris* or an insertion of the region in the ASGR of *P. squamulatum*. The lack of synteny between the two species in this

region suggests that this region would not be required to confer the aposporous trait (GOEL *et al.* 2006).

Sequence analysis of bacterial artificial chromosome (BAC) clones from the ASGR of *Pennisetum* and *Cenchrus* has shown multiple small regions of shared synteny to the rice and sorghum genomes throughout the ASGR (CONNER *et al.* 2008). Non-repetitive BACs containing putative protein coding regions (PPCRs) were identified by Basic Local Alignment Search Tool (BlastX) analysis of the sample sequences to sorghum and rice protein databases. Small regions of microsynteny for the ASGR-linked BACs containing PPCRs were identified in rice and sorghum. However, large regions of collinearity between the ASGR and either rice or sorghum chromosome seems unlikely based on these results. These results mean that unless the size of the ASGR is greatly reduced, using sorghum or rice data to help identify candidate apomictic genes based on map positions or chromosome walking through the ASGR does not seem likely (CONNER *et al.* 2008).

FISH analysis of three BC_8 derived lines has revealed that the number of *P. squamulatum* chromosomes is reduced to one univalent chromosome, the ASGR-carrier chromosome (SINGH *et al.* 2010). Since BC_7 and BC_8 lines have the same number of chromosomes (28), but BC_7 contains two *P. squamulatum* derived chromosomes, the other alien univalent was replaced by a pearl millet chromosome in the BC_8 line 58. The replacement of the alien univalent chromosome by pearl millet chromosome could be the cause for the reduction of frequency of apospory in BC_8 (SINGH *et al.* 2010). Two transcriptomes derived from micro dissected ovules at the stage of aposporous initial formation between the apomictic donor parent, *P. squamulatum* (Ps26) and the apomictic derived BC_8 line 58 were compared in an attempt to discover candidate transcripts regulating aposporous initial specification in *P. squamulatum* (ZENG *et al.* 2011). 33,977 contigs

from the Ps26 ovule transcriptomes library and 26,576 contigs from the apomictic BC_8 line 58 ovule transcriptomes library resulted from assembly of reads obtained from Roche 454 high throughput sequencing. Analyses of both transcriptomes assemblies for biological function using Blast2GO resulted in 90% of the top BlastX hits, in order, to sorghum, maize, or rice proteins. The assembled contigs from the two libraries were analyzed by BlastN with Ps26 sequences as queries and BC_8 sequences as the database, which resulted in a total of 118 comparisons with 100% identity across and overlapping region ≥100 bp. Further analyses aligning the 118 Ps26/BC_8 contigs resulted in 61 inter-genotype contigs with no mismatches and having average overlapping regions of 241 bp. The remaining contigs initially identified by BlastN as having 100% identity over the region >100 bp did not continue to share sequence similarity outside this region. Linkage of the 61 contigs to the ASGR-carrier chromosome was tested using primer pairs designed from the Ps26 contig (ZENG *et al.* 2011). PCR screening of the contigs to the ASGR-carrier chromosome identified 45 contigs with amplification in Ps26 and apomictic BC_8 line 58 genotypes but no amplification from the BC_8 line 58 sexual genotypes. This result established the linkage of 45 transcripts to the ASGR-carrier chromosome. Four more contigs were linked to the ASGR-carrier chromosome using single-stranded conformation polymorphism (SSCP) analysis and cleavage amplified polymorphic sequences (CAPS) screening (ZENG *et al.* 2011). Among the 49 ASGR-carrier chromosome linked contigs, only one contig, Ps26_c9369, showed tight linkage to the ASGR based on PCR amplification of F1s using primers. While mapping the transcripts to the ASGR carrier chromosome was relatively straight forward, mapping to the ASGR was problematic and hampered by an efficient way of identifying polymorphism in the F_1s (ZENG *et al.* 2011).

In a recent study, ASGR-linked markers previously shown to be tightly linked to aposporous embryo sac development in *C. ciliaris* were found to segregate in the A8 plant derived from a cross using sexual *C. ciliaris* genotype B-2s and aposporous *C. ciliaris* genotype B-12-9 as female and pollen donor, respectively (CONNER *et al.* 2013). The A8 plant retained only two of nine tested ASGR-linked markers. Cleared ovaries of the A8 plants had an average of 2.3 unreduced aposporous embryo sacs showing that the loss of ASGR-linked SCAR markers did not interfere with the formation of aposporous embryo sac development. Flow cytometry of leaf tissue with sorghum as a genome size control confirmed that the open pollinated and crossed A8 seedlings were mostly 6x in ploidy level. Genotyping of three A8 derived seedlings with a 4x ploidy level determined that they were derived sexually and not by apomixis (CONNER *et al.* 2013). The increase of ploidy level in the A8 seedlings along with the loss of a section of the ASGR-linked markers, physically mapped to the ASGR via FISH, strongly suggests that recombination between apomixis components apomeiosis and parthenogenesis has taken place in the A8 plant. The four ASGR-linked SCAR markers found to be retained in A8 will help in physical mapping of the region associated with apomeiosis at the ASGR in *C. ciliaris* (CONNER et al. 2013).

Comparative Genomics and Collinearity in Grass Family

The idea of comparative studies of similarities and differences of structures and functions in plants actually precedes the identification of DNA as hereditary molecule (PATERSON *et al.* 2000). The idea behind carrying out comparative studies of morphological and functional similarities of grasses is that the knowledge gained in one species can be transferred to other grass species (PATERSON *et al.* 1995) for genes controlling key agronomic traits. Comparative studies can also result in an increase in our understanding of the evolutionary mechanisms that

have led to the current structure of grass genomes. Sequencing of plant genomes has provided insights into genome composition, organization and evolution. The comparative analysis of crop genomic sequences, the nature of chromosomes evolution, identification of syntenic blocks, detection of genomic re-arrangements and chromosome reshuffling have become possible with the analysis of genetic and physical maps and the use of molecular cytogenetic tools (SALSE *et al.* 2008). Consensus grass map created by aligning genomes of several grass species was used to describe different grass genomes in terms of "rice linkage blocks" and nine different grass genomes could be described by only 25 "rice linkage blocks". The comparative mapping of RFLP markers and comparative sequence analysis of cereal genomes revealed a remarkable conservation of gene content (synteny) and gene order (collinearity) across different grass species despite enormous differences in the chromosome number, the ploidy level and the genome size (GALE and DEVOS 1998; DUBCOVSKY *et al.* 2001).

Besides sequencing data from different species, availability of convenient sequence comparison tools is also essential for comparative study of grasses. Availability of powerful basic local alignment search tool (BLAST) program (ALTSCHUL *et al.* 1990) makes it possible to compare genes and sequences from one species to another species with numerous genomic information (like lists of genes, coordinates, annotations). Numerous tools have been developed and are now publicly available to compare plant genomes and tentatively identify orthologous genes (SALSE *et al.* 2009). In extensively studied grass species like maize, rice, sorghum, barley and wheat, conservation of gene order is found to the DNA sequence level (micro-collinearity). This collinearity of genes at the molecular level has allowed the study of both coding and non-coding regions and has resulting in better understanding of the molecular and genome evolution in the grasses. The understandings and knowledge obtained from the genome organization of

these different crop species is important to define and identify the best strategies and tools to isolate genes of agronomic and evolutionary importance from large and complex cereal (FEUILLET and KELLER 2002; BENNETZEN 2007). The reference genomes for sorghum and foxtail millet have been completed (PATERSON 2009; BENNETZEN *et al.* 2012). The annotated reference assembly of sorghum and foxtail millet can serve as a resource for comparative study of the orthologous gene pairs for the ASGR-carrier chromosome transcripts in sorghum and foxtail millet.

Next-Generation Sequencing Technology and Duplex Specific Nuclease

The Sanger enzymatic dideoxynucleotide DNA sequencing technique (SANGER *et al.* 1977) revolutionized the study of genetics and genomics by deciphering complete genes and later entire genomes. However, limitations of Sanger sequencing for routine large sequence output and difficulty in total automation of the sample preparation led to the demand for faster and more affordable sequencing technologies (ANSORGE 2009). The development of next-generation sequencing (NGS) technologies allowed parallel sequencing by means of a sensitive charge-coupled device (CCD) camera (ANSORGE 2009). The next-generation sequencers have had a major impact on genome wide research as they can sequence and process millions of sequence reads at a time from relatively small amount of fragmented DNA 'libraries' without the need of conventional vector-based cloning (MARDIS 2008). Current commercially available next-generation sequencers all produce shorter read lengths than conventional capillary sequencers, which impact the assembly of the reads, especially for *de novo* sequencing projects. Three major platforms of NGS in widespread use for massive parallel DNA sequencing are Roche/454 FLX Pyrosequencer, the Illumina/Solexa Genome Analyzer and the Applied Biosystems SOLiD™.

The tremendous amount of data from next generation sequencing has opened a platform for investigating deeper into the transcriptomes and genomes of several crop species. The presence of repetitive sequences, which can constitute a large proportion of a plant genome size, makes assembly of shotgun sequenced clones almost impossible (SANMIGUEL and BENNETZEN 1998). Limiting the redundant sequencing of repetitive elements with effective approaches can facilitate sequencing analyses of large and complex genomes. Several approaches have been used for the elimination of repetitive sequences and the enrichment of low copy sequences. In maize, approaches like Methylation Filtration (MF) (PALMER *et al.* 2003), development of hypomethylated partial restriction (HMPR) libraries (EMBERTON *et al.* 2005) and methylation-spanning linker libraries (MSLL) (YUAN *et al.* 2002) which utilize the tendency of repetitive sequences to be hypermethylated in higher plants for eliminating repetitive sequences and enriching the low copy sequences, have been reported. Another approach, high-C_0t analysis, is based on the renaturation kinetics of DNA. When a denatured genomic DNA solution is kept in an environment where it is allowed to renature, the rate of reassociation of a particular sequence is directly proportional to its repetitiveness in the genome. Since low copy sequences renature more slowly than high copy sequences, the double stranded repetitive DNA can be separated from the single-stranded (low copy) DNA by hydroxyapatite chromatography after the sheared and heat denatured DNA re-anneals slowly. This approach has been used in the study of the maize and sorghum genome (PETERSON *et al.* 2002; YUAN *et al.* 2003).

A more recent approach is the use of duplex specific nuclease (DSN) normalization technology, which is also based on hybridization kinetics, like high-C_0t DNA fractionation, but does not involve physical separation of single stranded (ss) DNA and double stranded (ds) DNA. DSN, obtained from the hepato-pancreas of the Kamchatka crab, is a thermostable enzyme that

displays a strong preference for cleaving dsDNA and DNA in RNA-DNA hybrid duplexes compared to ssDNA and RNA (SHAGIN *et al.* 2002). Initially this enzyme was used in a cDNA normalization technique for the enrichment of full length cDNA sequences to discover rare transcripts by the removal of intermediate and highly abundant cDNAs (ZHULIDOV *et al.* 2004; ZHULIDOV *et al.* 2005). The application of DSN normalization technology to whole-genome shotgun sequencing of the human genome which contain a large proportion of repetitive DNA was reported by Shagina *et al.* (2010). The results from the application of DSN in human genomic DNA indicated that DSN normalization can be highly effective for the elimination of evolutionarily young repetitive sequences from genomic DNA prior to sequencing (SHAGINA *et al.* 2010). Unlike most mammalian repeats, the genomes of many higher plants contain a large number of highly conserved repetitive elements that are evolutionarily young and share high sequence identity (SMITH and FLAVELL 1975; SANMIGUEL *et al.* 1998; RAMAKRISHNA *et al.* 2002).

Successful application of DSN normalization of plant genomic libraries has been reported in lettuce and Arabidopsis genome by Matvienko *et al* (2013). The study identifies that DSN normalization in larger plant genomes may enable efficient access to the low-copy fraction of genomes for identifying sequence and copy-number variants and also for mapping purposes. Furthermore, libraries for Illumina sequencing with short 300-500 bp fragments are reported to be ideally suited for reducing highly repeated sequences without concomitant loss of neighboring low-copy sequences (MATVIENKO *et al.* 2013).

References

AKIYAMA, Y., J. A. CONNER, S. GOEL, D. T. MORISHIGE, J. E. MULLET *et al.*, 2004 High-resolution physical mapping in Pennisetum squamulatum reveals extensive chromosomal

heteromorphism of the genomic region associated with apomixis. Plant Physiology **134:** 1733-1741.

AKIYAMA, Y., S. GOEL, J. A. CONNER, W. W. HANNA, H. YAMADA-AKIYAMA *et al.*, 2011 Evolution of the apomixis transmitting chromosome in Pennisetum. BMC Evolutionary Biology **11:** 289-289.

AKIYAMA, Y., P. OZIAS-AKINS and W. W. HANNA, 2005 High-resolution physical mapping reveals that the apospory-specific genomic region (ASGR) in Cenchrus ciliaris is located on a heterochromatic and hemizygous region of a single chromosome. Theoretical and Applied Genetics **111:** 1042-1051.

ALBERTINI, E., G. MARCONI, G. BARCACCIA, L. RAGGI and M. FALCINELLI, 2004 Isolation of candidate genes for apomixis in Poa pratensis L. Plant Molecular Biology **56:** 879-894.

ALBERTINI, E., G. MARCONI, L. REALE, G. BARCACCIA, A. PORCEDDU *et al.*, 2005 SERK and APOSTART. Candidate genes for apomixis in Poa pratensis. Plant Physiology **138:** 2185-2199.

ALBERTINI, E., A. PORCEDDU, F. FERRANTI, L. REALE, G. BARCACCIA *et al.*, 2001 Apospory and parthenogenesis may be uncoupled in Poa pratensis: a cytological investigation. Sexual Plant Reproduction **14:** 213-217.

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS and D. J. LIPMAN, 1990 Basic local alignment search tool. Journal Of Molecular Biology **215:** 403-410.

ANSORGE, W. J., 2009 Next-generation DNA sequencing techniques. New Biotechnology **25:** 195-203.

BARCACCIA, G., and E. ALBERTINI, 2013 Apomixis in plant reproduction: a novel perspective on an old dilemma. Plant Reproduction **26:** 159-179.

BARRELL, P. J., and U. GROSSNIKLAUS, 2005 Confocal microscopy of whole ovules for analysis of reproductive development: the elongate1 mutant affects meiosis II. Plant Journal **43:** 309-320.

BENNETZEN, J. L., 2007 Patterns in grass genome evolution. Current Opinion in Plant Biology **10:** 176-181.

BENNETZEN, J. L., J. SCHMUTZ, H. WANG, R. PERCIFIELD, J. HAWKINS *et al.*, 2012 Reference genome sequence of the model plant Setaria. Nature Biotechnology **30:** 555-561.

BICKNELL, R. A., 1997 Isolation of a diploid, apomictic plant of Hieracium aurantiacum. Sexual Plant Reproduction **10:** 168-172.

BICKNELL, R. A., and A. M. KOLTUNOW, 2004 Understanding Apomixis: Recent Advances and Remaining Conundrums, pp. S228. American Society of Plant Biologists.

BOUTILIER, K., R. OFFRINGA, V. K. SHARMA, H. KIEFT, T. OUELLET *et al.*, 2002 Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth, pp. 1737. American Society of Plant Biologists.

CALDERINI, O., S. B. CHANG, H. DE JONG, A. BUSTI, F. PAOLOCCI *et al.*, 2006 Molecular cytogenetics and DNA sequence analysis of an apomixis-linked BAC in Paspalum simplex reveal a non pericentromere location and partial microcolinearity with rice. Theoretical & Applied Genetics **112:** 1179-1191.

CARMAN, J. G., 1997 Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony [electronic resource]. Biological journal of the Linnean Society **61:** 51-94.

CATANACH, A. S., S. K. ERASMUSON, E. PODIVINSKY, B. R. JORDAN and R. BICKNELL, 2006 Deletion Mapping of Genetic Regions Associated with Apomixis in Hieracium, pp. 18650. National Academy of Sciences.

CECCATO, L., S. MASIERO, D. SINHA ROY, S. BENCIVENGA, I. ROIG-VILLANOVA *et al.*, 2013 Maternal Control of PIN1 Is Required for Female Gametophyte Development in Arabidopsis. PLoS ONE **8:** 1-7.

CHENG, C.-Y., D. E. MATHEWS, G. E. SCHALLER and J. J. KIEBER, 2013 Cytokinin-dependent specification of the functional megaspore in the Arabidopsis female gametophyte. The Plant Journal: For Cell And Molecular Biology **73:** 929-940.

COMAI, L., A. P. TYAGI, K. WINTER, R. HOLMES-DAVIS, S. H. REYNOLDS *et al.*, 2000 Phenotypic Instability and Rapid Gene Silencing in Newly Formed Arabidopsis Allotetraploids, pp. 1551. American Society of Plant Physiologists.

CONNER, J. A., S. GOEL, G. GUNAWAN, M. M. CORDONNIER-PRATT, V. E. JOHNSON *et al.*, 2008 Sequence analysis of bacterial artificial chromosome clones from the apospory-specific genomic region of Pennisetum and Cenchrus. Plant Physiology **147:** 1396-1411.

CONNER, J. A., G. GUNAWAN and P. OZIAS-AKINS, 2013 Recombination within the apospory specific genomic region leads to the uncoupling of apomixis components in Cenchrus ciliaris. Planta **238:** 51-63.

CORRAL, J. M., H. VOGEL, O. M. ALIYU, G. HENSEL, T. THIEL *et al.*, 2013 - A Conserved Apomixis-Specific Polymorphism Is Correlated with Exclusive Exonuclease Expression in Premeiotic Ovules of Apomictic Boechera Species. **- 163:** - 1672.

D'ERFURTH, I., S. JOLIVET, N. FROGER, O. CATRICE, M. NOVATCHKOVA *et al.*, 2009 Turning Meiosis into Mitosis. PLoS Biology **7:** 1-10.

DE MEEÛS, T., F. PRUGNOLLE and P. AGNEW, 2007 Asexual reproduction: genetics and evolutionary aspects. Cellular And Molecular Life Sciences: CMLS **64:** 1355-1372.

DUBCOVSKY, J., R. WUSIRIKA, P. J. SANMIGUEL, C. S. BUSSO, L. YAN *et al.*, 2001 Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. Plant Physiology **125:** 1342-1353.

DUJARDIN, M., and W. HANNA, 1984 Cytogenetics of double cross hybrids between Pennisetum americanum - P. purpureum amphiploids and P. americanum X Pennisetum squamulatum interspecific hybrids. Theoretical and Applied Genetics **69:** 97-100.

DUJARDIN, M., and W. W. HANNA, 1989a Crossability of pearl millet with wild Pennisetum species. Crop Science **29:** 77-80.

DUJARDIN, M., and W. W. HANNA, 1989b Developing apomictic pearl millet - characterization of a BC3 plant. Journal of Genetics & Breeding **43:** 145-149.

EBINA, M., H. NAKAGAWA, T. YAMAMOTO, H. ARAYA, S.-I. TSURUTA *et al.*, 2005 Co-segregation of AFLP and RAPD markers to apospory in Guineagrass (Panicum maximumJacq.). Grassland Science **51:** 71-78.

EDGER, P. P., and J. C. PIRES, 2009 Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes [electronic resource]. Chromosome Research **17:** 699-717.

EMBERTON, J., J. X. MA, Y. N. YUAN, P. SANMIGUEL and J. L. BENNETZEN, 2005 Gene enrichment in maize with hypomethylated partial restriction (HMPR) libraries. Genome Research **15:** 1441-1446.

FEUILLET, C., and B. KELLER, 2002 Comparative Genomics in the Grass Family: Molecular Characterization of Grass Genome Structure and Evolution. Annals of Botany **89:** 3-10.

FRISVOLD, G., K. BICKNELL and R. BICKNELL, 2005 A preliminary analysis of the benefits of introducing apomixis into rice, pp. in *AERU Discussion Paper*. Agribusiness & Economics Research Unit, Lincoln University, Canterbury; New Zealand.

GALE, M. D., and K. M. DEVOS, 1998 Comparative genetics in the grasses. Proceedings of the National Academy of Sciences of the United States of America **95:** 1971-1974.

GARCIA-AGUILAR, M., C. MICHAUD, O. LEBLANC and D. GRIMANELLI, 2010 Inactivation of a DNA Methylation Pathway in Maize Reproductive Organs Results in Apomixis-Like Phenotypes, pp. 3249. American Society of Plant Biologists.

GARCÍA, R., M. J. ASÍNS, J. FORNER and E. A. CARBONELL, 1999 Genetic analysis of apomixis in Citrus and Poncirus by molecular markers. Theoretical and Applied Genetics **99:** 511-518.

GOEL, S., Z. B. CHEN, Y. AKIYAMA, J. A. CONNER, M. BASU *et al.*, 2006 Comparative physical mapping of the apospory-specific genomic region in two apomictic grasses: Pennisetum squamulatum and Cenchrus ciliaris, pp. 389-400 in *Genetics*. Genetics Society of America, Bethesda; USA.

GRIMANELLI, LEBLANC, ESPINOSA, PEROTTI, GONZÁLEZ DE LEÓN *et al.*, 1998 Non-Mendelian transmission of apomixis in maize–Tripsacum hybrids caused by a transmission ratio distortion. Heredity **80:** 40-47.

GRIMANELLI, D., M. HERNÁNDEZ, E. PEROTTI and Y. SAVIDAN, 1997 Dosage effects in the endosperm of diplosporous apomictic Tripsacum (Poaceae). Sexual Plant Reproduction **10:** 279-282.

GRIMANELLI, D., O. LEBLANC, E. PEROTTI and U. GROSSNIKLAUS, 2001 Developmental genetics of gametophytic apomixis. Trends in Genetics **17:** 597-604.

GROSSNIKLAUS, U., 2001 From sexuality to apomixis: Molecular and genetic approaches., pp. 168-211 in *The Flowering of Apomixis: From Mechanisms to Genetic Engineering*, edited by J. G. C. Y. SAVIDAN, AND T. DRESSELHAUS. CIMMYT, IRD, European Commission DG, Mexico, D. F,.

GROSSNIKLAUS, U., and K. SCHNEITZ, 1998 The molecular and genetic basis of ovule and megagametophyte development. Seminars In Cell & Developmental Biology **9:** 227-238.

GUIMARÃES, L. A., D. M. DE A. DUSI, S. MASIERO, F. RESENTINI, A. C. M. M. GOMES *et al.*, 2013 BbrizAGL6 Is Differentially Expressed During Embryo Sac Formation of Apomictic and Sexual Brachiaria brizantha Plants. Plant molecular biology reporter **31:** 1397-1406.

GUO, M., D. DAVIS and J. A. BIRCHLER, 1996 Dosage effects on gene expression in a maize ploidy series. Genetics **142:** 1349-1355.

HAND, M. L., and A. M. G. KOLTUNOW, 2014 The Genetic Control of Apomixis: Asexual Seed Formation. Genetics **197:** 441-450.

HANNA, W. W., 1995 Use of Apomixis in cultivar development. Advances in agronomy **54:** 333-350.

HANNA, W. W., and E. C. BASHAW, 1987 Apomixis: its identification and use in plant breeding. Crop Science **27:** 1136-1139.

HEQIANG, H., J. A. CONNER and P. OZIAS-AKINS, 2009 Genetic mapping of the apospory-specific genomic region in Pennisetum squamulatum using retrotransposon-based molecular markers. Theoretical & Applied Genetics **119:** 199-212.

HÖRANDL, E., and E. M. TEMSCH, 2009 Introgression of apomixis into sexual species is inhibited by mentor effects and ploidy barriers in the Ranunculus auricomus complex. Annals of Botany **104:** 81-89.

JURANIC, M., K. O. SRILUNCHANG, N. G. KROHN, D. LELJAK-LEVANIC, S. SPRUNCK *et al.*, 2012 - Germline-Specific MATH-BTB Substrate Adaptor MAB1 Regulates Spindle Length and Nuclei Identity in Maize. **- 24:** - 4991.

KANTAMA, L., T. F. SHARBEL, M. E. SCHRANZ, T. MITCHELL-OLDS, S. DE VRIES *et al.*, 2007 Diploid Apomicts of the Boechera holboellii Complex Display Large-Scale Chromosome Substitutions and Aberrant Chromosomes, pp. 14026. National Academy of Sciences.

KAUSHAL, P., D. R. MALAVIYA and A. K. ROY, 2004 Prospects for breeding apomictic rice: a reassessment. Current Science **87:** 292-296.

KINDIGER, B., D. BAI and V. SOKOLOV, 1996 Assignment of a gene(s) conferring apomixis in Tripsacum to a chromosome arm: cytological and molecular evidence. Genome **39:** 1133-1141.

KOJIMA, A., and Y. NAGATO, 1997 Discovery of highly apomictic and highly amphimictic dihaploids in Allium tuberosum. Sexual Plant Reproduction **10:** 8-12.

KOLTUNOW, A. M., 1993 Apomixis: Embryo Sacs and Embryos Formed without Meiosis or Fertilization in Ovules, pp. 1425. American Society of Plant Physiologists.

KOLTUNOW, A. M., R. A. BICKNELL and A. M. CHAUDHURY, 1995 Apomixis: Molecular Strategies for the Generation of Genetically Identical Seeds without Fertilization, pp. 1345. American Society of Plant Physiologists.

KOLTUNOW, A. M. G., S. D. JOHNSON, J. C. M. RODRIGUES, T. OKADA, Y. HU *et al.*, 2011 Sexual reproduction is the default mode in apomictic Hieracium subgenus Pilosella, in

which two dominant loci function to enable apomixis. The Plant Journal: For Cell And Molecular Biology **66:** 890-902.

KOTANI, Y., S. T. HENDERSON, G. SUZUKI, S. D. JOHNSON, T. OKADA *et al.*, 2014 The LOSS OF APOMEIOSIS (LOA) locus in Hieracium praealtum can function independently of the associated large-scale repetitive chromosomal structure. New Phytologist **201:** 973-981.

LASPINA, N. V., J. P. A. ORTIZ, M. PODIO, V. C. ECHENIQUE, S. C. PESSINO *et al.*, 2008 Gene expression analysis at the onset of aposporous apomixis in Paspalum notatum [electronic resource]. Plant Molecular Biology **67:** 615-628.

LEBLANC, O., D. GRIMANELLI, D. GONZÁLEZ DE LEÓN and Y. SAVIDAN, 1995 Detection of the apomictic mode of reproduction in maize-Tripsacum hybrids using maize RFLP markers. Theoretical and Applied Genetics **90:** 1198-1203.

MARDIS, E. R., 2008 The impact of next-generation sequencing technology on genetics. Trends in Genetics **24:** 133-141.

MARTINEZ, E. J., J. P. A. ORTIZ, C. L. QUARIN, H. E. HOPP and J. STEIN, 2003 Genetic characterization of apospory in tetraploid Paspalum notatum based on the identification of linked molecular markers. Molecular breeding : new strategies in plant improvement **12:** 319-327.

MARTÍNEZ, E. J., C. L. QUARÍN and M. D. HAYWARD, 1999 Genetic control of apospory in apomictic Paspalum species. Cytologia **64:** 425-433.

MATVIENKO, M., A. KOZIK, L. FROENICKE, D. LAVELLE, B. MARTINEAU *et al.*, 2013 Consequences of Normalizing Transcriptomic and Genomic Libraries of Plant Genomes Using a Duplex-Specific Nuclease and Tetramethylammonium Chloride. PLoS ONE **8:** 1-17.

MATZK, F., A. MEISTER, R. BRUTOVSKÁ and I. SCHUBERT, 2001 Reconstruction of reproductive diversity in Hypericum perforatum L. opens novel strategies to manage apomixis. Plant Journal **26:** 275-282.

MATZK, F., S. PRODANOVIC, H. BÄUMLEIN and I. SCHUBERT, 2005 The inheritance of apomixis in Poa pratensis confirms a five locus model with differences in gene expressivity and penetrance. Plant Cell **17:** 13-24.

MEISTER, G., 2013 Argonaute proteins: functional insights and emerging roles. Nature Reviews Genetics **14:** 447-459.

MORGAN, R. N., P. OZIAS-AKINS and W. W. HANNA, 1998 Seed Set in an Apomictic BC 3 Pearl Millet, pp. 89. The University of Chicago Press.

NAKANO, M., T. SHIMADA, T. ENDO, H. FUJII, H. NESUMI *et al.*, 2012 Characterization of genomic sequence showing strong association with polyembryony among diverse Citrus species and cultivars, and its synteny with Vitis and Populus. Plant Science **183:** 131-142.

NAUMOVA, T. N., J. BERGERVOET, K. S. RAMULU, K. BOUTILIER, A. KRAVTCHENKO *et al.*, 2001 Reproductive development in apomictic populations of Arabis holboellii (Brassicaceae) [electronic resource]. Sexual Plant Reproduction **14:** 195-200.

NOGLER, G. A., 1984 Gametophytic apomixis, pp. 474-518 in *Embryology of angiosperms / edited by B.M. Johri*. Berlin : Springer-Verlag, 1984.

NONOMURA, K.-I., A. MOROHOSHI, M. NAKANO, M. EIGUCHI, A. MIYAO *et al.*, 2007 A Germ Cell-Specific Gene of the ARGONAUTE Family Is Essential for the Progression of Premeiotic Mitosis and Meiosis during Sporogenesis in Rice, pp. 2583. American Society of Plant Biologists.

NOYES, R. D., and L. H. RIESEBERG, 2000 Two independent loci control agamospermy (apomixis) in the triploid flowering plant Erigeron annuus. Genetics **155:** 379.

OGAWA, D., S. JOHNSON, S. HENDERSON and A. KOLTUNOW, 2013 Genetic separation of autonomous endosperm formation (AutE) from the two other components of apomixis in Hieracium. Plant Reproduction **26:** 113-123.

OKADA, T., Y. K. HU, M. R. TUCKER, J. M. TAYLOR, S. D. JOHNSON *et al.*, 2013 - Enlarging Cells Initiating Apomixis in Hieracium praealtum Transition to an Embryo Sac Program prior to Entering Mitosis. **- 163:** - 231.

OKADA, T., K. ITO, S. D. JOHNSON, K. OELKERS, G. SUZUKI *et al.*, 2011 Chromosomes Carrying Meiotic Avoidance Loci in Three Apomictic Eudicot Hieracium Subgenus Pilosella Species Share Structural Features with Two Monocot Apomicts, pp. 1327. American Society of Plant Biologists.

OLMEDO-MONFIL, V., N. DURÁN-FIGUEROA, M. ARTEAGA-VÁZQUEZ, E. DEMESA-ARÉVALO, D. AUTRAN *et al.*, 2010 Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature **464:** 628-632.

OZIAS-AKINS, P., 2006 Apomixis: Developmental Characteristics and Genetics. Critical Reviews in Plant Sciences **25:** 199-214.

OZIAS-AKINS, P., Y. AKIYAMA and W. W. HANNA, 2003 Molecular characterization of the genomic region linked with apomixis in Pennisetum/Cenchrus. Functional & Integrative Genomics **3:** 94-104.

OZIAS-AKINS, P., D. ROCHE and W. W. HANNA, 1998 Tight Clustering and Hemizygosity of Apomixis-Linked Molecular Markers in Pennisetum squamulatum Implies Genetic

Control of Apospory by a Divergent Locus that may have no Allelic Form in Sexual Genotypes, pp. 5127. National Academy of Sciences of the United States of America.

OZIAS-AKINS, P., and P. J. VAN DIJK, 2007 Mendelian Genetics of Apomixis in Plants. Annual Review of Genetics **41:** 509-537.

PALMER, L. E., P. D. RABINOWICZ, A. L. O'SHAUGHNESSY, V. S. BALIJA, L. U. NASCIMENTO *et al.*, 2003 Maize genome sequencing by methylation filtration. Science (New York, N.Y.) **302:** 2115-2117.

PATERSON, A. H., J. E. BOWERS, M. D. BUROW, X. DRAYE, C. G. ELSIK *et al.*, 2000 Comparative genomics of plant chromosomes. The Plant Cell **12:** 1523-1540.

PATERSON, A. H., Y.-R. LIN, Z. LI, K. F. SCHERTZ, J. F. DOEBLEY *et al.*, 1995 Convergent Domestication of Cereal Crops by Independent Mutations at Corresponding Genetic Loci, pp. 1714. American Society for the Advancement of Science.

PATERSON, A. H. B. J. E. B. R. I. J. H. G. U. T. A. J. M. H. X. T. A. K. C. J. F. A. U. I. V., 2009 The Sorghum bicolor genome and the diversification of grasses. Nature **457:** 551-556.

PESSINO, S. C., J. P. A. ORITZ, O. LEBLANC, C. B. D. VALLE, C. EVANS *et al.*, 1997 Identification of a maize linkage group related to apomixis in Brachiaria. Theoretical and Applied Genetics **94:** 439-444.

PETERSON, D. G., S. R. SCHULZE, E. B. SCIARA, S. A. LEE, J. E. BOWERS *et al.*, 2002 Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery. Genome Research **12:** 795-807.

PODIO, M., S. FELITTI, L. SIENA, L. DELGADO, M. MANCINI *et al.*, 2014 Characterization and expression analysis of SOMATIC EMBRYOGENESIS RECEPTOR KINASE (SERK) genes in sexual and apomictic Paspalum notatum. Plant Molecular Biology **84:** 479-495.

PUPILLI, F., and G. BARCACCIA, 2012 Cloning plants by seeds: inheritance models and candidate genes to increase fundamental knowledge for engineering apomixis in sexual crops. Journal of Biotechnology **159:** 291-311.

QUARIN, C. L., F. ESPINOZA, E. J. MARTINEZ, S. C. PESSINO and O. A. BOVO, 2001 A rise of ploidy level induces the expression of apomixis in Paspalum notatum. Sexual Plant Reproduction **13:** 243-249.

RAMAKRISHNA, W., J. EMBERTON, M. OGDEN, P. SANMIGUEL and J. L. BENNETZEN, 2002 Structural Analysis of the Maize Rp1 Complex Reveals Numerous Sites and Unexpected Mechanisms of Local Rearrangement. The Plant Cell**:** 3213.

RAVI, M., and S. W. L. CHAN, 2010 Haploid plants produced by centromere-mediated genome elimination. Nature **464:** 615-618.

RAVI, M., M. P. A. MARIMUTHU and I. SIDDIQI, 2008 Gamete formation without meiosis in Arabidopsis. Nature **451:** 1121-1124.

REISER, L., and R. L. FISCHER, 1993 The ovule and the embryo sac. Plant Cell **5:** 1291-1301.

ROCHE, D., P. CONG, Z. CHEN, W. W. HANNA, D. L. GUSTINE *et al.*, 1999 An apospory-specific genomic region is conserved between Buffelgrass (Cenchrus ciliaris L.) and Pennisetum squamulatum Fresen. Plant Journal **19:** 203-208.

ROCHE, D., J. CONNER, M. BUDIMAN, D. FRISCH, R. WING *et al.*, 2002 Construction of BAC libraries from two apomictic grasses to study the microcolinearity of their apospory-specific genomic regions. Theoretical & Applied Genetics **104:** 804-812.

SALSE, J., M. ABROUK, F. MURAT, U. M. QURAISHI and C. FEUILLET, 2009 Improved criteria and comparative genomics tool provide new insights into grass paleogenomics. Briefings in Bioinformatics **10:** 619-630.

SALSE, J., S. BOLOT, M. THROUDE, V. JOUFFE, B. PIEGU *et al.*, 2008 Identification and Characterization of Shared Duplications between Rice and Wheat Provide New Insight into Grass Genome Evolution. The Plant Cell**:** 11.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA Sequencing with Chain-Terminating Inhibitors, pp. 5463. National Academy of Sciences of the United States of America.

SANMIGUEL, P., and J. L. BENNETZEN, 1998 Evidence that a Recent Increase in Maize Genome Size was Caused by the Massive Amplification of Intergene Retrotransposons. Annals of Botany **82:** 37-44.

SANMIGUEL, P., B. S. GAUT, A. TIKHONOV, Y. NAKAJIMA and J. L. BENNETZEN, 1998 The paleontology of intergene retrotransposons of maize. Nature Genetics **20:** 43-45.

SAVIDAN, Y., 2000 Apomixis: genetics and breeding. Plant breeding reviews **18:** 13-86.

SAVIDAN, Y., J. CARMAN, T. DRESSELHAUS and G. H. B. J. C. C. F. N. T. N. V.-C. J. P. S. R. T. G. D. T. J. G.-D.-L. TOENNIESSEN, 2001 The flowering of apomixis: from mechanisms to genetic engineering, pp. in *The flowering of apomixis: from mechanisms to genetic engineering*, edited by Y. SAVIDAN, J. CARMAN and T. DRESSELHAUS. Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT) (International Maize and Wheat Improvement Center), Mexico City; Mexico.

SCHALLAU, A., F. ARZENTON, A. J. JOHNSTON, U. HÄHNEL, D. KOSZEGI *et al.*, 2010 Identification and genetic analysis of the APOSPORY locus in Hypericum perforatum L. Plant Journal **62:** 773-784.

SHAGIN, D. A., D. V. REBRIKOV, V. B. KOZHEMYAKO, I. M. ALTSHULER, A. S. SHCHEGLOV *et al.*, 2002 A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. Genome Research **12:** 1935-1942.

SHAGINA, I., E. BOGDANOVA, I. Z. MAMEDOV, Y. LEBEDEV, S. LUKYANOV *et al.*, 2010 Normalization of genomic DNA using duplex-specific nuclease. BIOTECHNIQUES **48:** 455-459.

SHARBEL, T. F., M.-L. VOIGT, J. M. CORRAL, G. GALLA, J. KUMLEHN *et al.*, 2010 Apomictic and Sexual Ovules of Boechera Display Heterochronic Global Gene Expression Patterns, pp. 655. American Society of Plant Biologists.

SILVEIRA, R. D., M. M. DO CARMO COSTA, M. Ł. ALVES-FERREIRA, V. T. DE CAMPOS CARNEIRO, N. Ł. F. MARTINS *et al.*, 2012 Expressed sequence-tag analysis of ovaries of Brachiaria brizantha reveals genes associated with the early steps of embryo sac differentiation of apomictic plants [electronic resource]. Plant cell reports **31:** 403-416.

SINGH, M., J. A. CONNER, Y. J. ZENG, W. W. HANNA, V. E. JOHNSON *et al.*, 2010 Characterization of apomictic BC7 and BC8 pearl millet: meiotic chromosome behavior and construction of an ASGR-carrier chromosome-specific library. Crop Science **50:** 892-902.

SINGH, M., S. GOEL, R. B. MEELEY, C. DANTEC, H. PARRINELLO *et al.*, 2011 Production of Viable Gametes without Meiosis in Maize Deficient for an ARGONAUTE Protein, pp. 443. American Society of Plant Biologists.

SMITH, D., and R. FLAVELL, 1975 Characterisation of the wheat genome by renaturation kinetics. Chromosoma **50:** 223.

TAS, I. C. Q., and P. J. VAN DIJK, 1999 Crosses between sexual and apomictic dandelions (Taraxacum). I. The inheritance of apomixis. Heredity **83:** 707-714.

TUCKER, M. R., A.-C. G. ARAUJO, N. A. PAECH, V. HECHT, D. L. S. ED *et al.*, 2003 Sexual and Apomictic Reproduction in Hieracium Subgenus Pilosella Are Closely Interrelated Developmental Pathways, pp. 1524. American Society of Plant Biologists.

TUCKER, M. R., T. OKADA, S. D. JOHNSON, F. TAKAIWA and A. M. G. KOLTUNOW, 2012 Sporophytic ovule tissues modulate the initiation and progression of apomixis in Hieracium. Journal of Experimental Botany **63:** 3229-3241.

VAN DIJK, P. J., I. C. Q. TAS, M. FALQUE and T. BAKX-SCHOTMAN, 1999 Crosses between sexual and apomictic dandelions (Taraxacum). II. The breakdown of apomixis. Heredity **83:** 715-721.

VAN DIJK, P. J., and K. VIJVERBERG, 2005 The significance of apomixis in the evolution of the angiosperms: a reappraisal. REGNUM VEGETABILE **143:** 101-116.

WANG, Y., Y. HOU, H. GU, D. KANG, Z.-L. CHEN *et al.*, 2013 The Arabidopsis anaphase-promoting complex/cyclosome subunit 1 is critical for both female gametogenesis and embryogenesis(F). Journal Of Integrative Plant Biology **55:** 64-74.

YUAN, Y., P. J. SANMIGUEL and J. L. BENNETZEN, 2002 Methylation-spanning linker libraries link gene-rich regions and identify epigenetic boundaries in Zea mays. Genome Research **12:** 1345-1349.

YUAN, Y., P. J. SANMIGUEL and J. L. BENNETZEN, 2003 High-Cot sequence analysis of the maize genome. The Plant Journal: For Cell And Molecular Biology **34:** 249-255.

ZENG, Y., J. CONNER and P. OZIAS-AKINS, 2011 Identification of ovule transcripts from the Apospory-Specific Genomic Region (ASGR)-carrier chromosome. BMC Genomics **12:** 206-206.

ZENG, Y. J., 2009 Identification and characterization of apospory candidate genes in *Pennisetum* and *Cenchrus*, pp. 101-120 in *PhD Dissertation*. University of Georgia, Athens, GA.

ZHULIDOV, P. A., E. A. BOGDANOVA, A. S. SHCHEGLOV, I. A. SHAGINA, L. L. VAGNER *et al.*, 2005 [A method for the preparation of normalized cDNA libraries enriched with full-length sequences]. Bioorganicheskaia Khimiia **31:** 186-194.

ZHULIDOV, P. A., E. A. BOGDANOVA, A. S. SHCHEGLOV, L. L. VAGNER, G. L. KHASPEKOV *et al.*, 2004 Simple cDNA normalization using kamchatka crab duplex-specific nuclease. Nucleic Acids Research **32:** e37-e37.

CHAPTER 2

PHYSICAL MAPPING OF THE ASGR-CARRIER CHROMOSOME REVEALS COLLINEARITY TO CHROMOSOME 2 OF SORGHUM AND FOXTAIL MILLET OUTSIDE THE ASGR[1]

[1]Sapkota, S. K, J. A. Conner, and P. Ozias-Akins. To be submitted to *BMC Genomics*

Abstract

Apomixis is a process of asexual reproduction through seed, in which an embryo develops from an unreduced embryo sac of the ovule bypassing meiosis and fertilization of the gamete. Apomixis can be classified as sporophytic or gametophytic. Gametophytic apomixis can be further subdivided as diplospory or apospory. In *Pennisetum squamulatum*, apospory is controlled by a single genetic locus consisting of a physically large, hemizygous, and heterochromatic chromosomal block called the apospory-specific genomic region (ASGR). In *P. squamulatum*, the ASGR is located near the telomere on the short arm of the ASGR-carrier chromosome. Partially sequenced BACs linked to the ASGR did not identify large regions of collinearity with the sequenced genomes of sorghum and rice. However, enrichment for chromosome 2 of sorghum was identified in apomictic BC_8 line 58, a line containing only the ASGR-carrier chromosome from *P. squamulatum*, in comparative genomic hybridization (CGH) experiments. BlastN results using ovule transcripts genetically mapped to the ASGR-carrier chromosome against the sorghum and foxtail millet reference genome exhibited a collinear distribution of orthologous genes on chromosome 2 of sorghum and foxtail millet. Bacterial artificial chromosome (BAC) clones containing the ASGR-carrier chromosome SCAR markers were identified. Physical mapping of eight BACs carrying the ASGR-carrier chromosome SCAR markers by fluorescence *in situ* hybridization (FISH) onto the ASGR-carrier chromosome revealed large scale collinearity to chromosome 2 of both sorghum and foxtail millet, outside the ASGR. Based on this analysis, the ASGR-carrier chromosome of *P. squamulatum*, outside of the ASGR, is evolutionarily similar to chromosome 2 of sorghum and foxtail millet.

Introduction

Apomixis is an asexual form of reproduction through seeds (NOGLER 1984). In apomictic reproduction, unlike sexual reproduction, the embryo develops from cells in the ovule without meiotic division and fertilization of gametes (HAND and KOLTUNOW 2014). The endosperm in apomictic seeds can, however, be generated either autonomously (autogamy) or through fertilization of polar nuclei (pseudogamy). Because apomictic plants are genetically identical to the maternal plant, apomixis has the potentiality to fix hybrid vigor in crop species (OZIAS-AKINS 2006; HAND and KOLTUNOW 2014). Although apomixis is widespread in angiosperms and has been observed in over 400 species spanning ~40 different families, apomixis is rarely found in major crop species (CARMAN 1997). Several efforts have been made to transfer the apomictic trait to major grains like maize (LEBLANC *et al.* 1995) and pearl millet (DUJARDIN and HANNA 1989b) by interspecific hybridization. These efforts have not yet resulted in commercially viable germplasm (SAVIDAN 2000; OZIAS-AKINS and VAN DIJK 2007; HAND and KOLTUNOW 2014).

Apomixis can be classified as sporophytic or gametophytic. Gametophytic apomixis can be further subdivided into diplospory and apospory. In diplospory, the embryo sac forms from the megaspore mother cell which does not undergo meiosis but instead undergoes mitotic divisions to form a non-reduce embryo sac. In apospory, the non-reduced embryo sac develops from mitotic divisions of aposporous initial cells which are formed from one or more nucellar cells of the ovule (OZIAS-AKINS *et al.* 2003; HAND and KOLTUNOW 2014). In gametophytic apomictic species such as *Taraxacum officinale* (VAN DIJK *et al.* 1999), *Poa pratensis* (ALBERTINI *et al.* 2001), *Erigeron annuus* (NOYES and RIESEBERG 2000), *Hieracium caespitosum* (CATANACH *et al.* 2006) and *Cenchrus ciliaris* (CONNER *et al.* 2013), formation of unreduced embryo sacs and parthenogenesis were confirmed to be controlled by independent

loci. In *Pennisetum squamulatum* and *Cenchrus ciliaris*, apospory is controlled by hemizygous genetic locus named the apospory-specific genomic region (ASGR) (OZIAS-AKINS *et al.* 1998; GOEL *et al.* 2003; AKIYAMA *et al.* 2004). Two bacterial artificial chromosomes (BAC) clone libraries were developed to study the ASGR in *P. squamulatum* and *C. ciliaris* (ROCHE *et al.* 2002). The ASGR in *P. squamulatum* was identified to be a large (>50 Mb), heterochromatic chromosomal block, consisting of two repetitive blocks flanking a single copy region, near the telomere on the short arm of the ASGR-carrier chromosome by fluorescence *in situ* hybridization (FISH) of ASGR-linked BACs (GOEL *et al.* 2003; AKIYAMA *et al.* 2004). ASGR-linked BAC clones were partially sequenced and the sequence analysis showed the presence of multiple regions of microsynteny but did not exhibit any large-scale collinearity with the rice or sorghum genome (CONNER *et al.* 2008). An apomictic backcross 8 (BC_8) line 58 plant, originating from a cross between *P. squamulatum* and *P. glaucum*, was identified by FISH to carry only the ASGR-carrier chromosome from *P. squamulatum* (SINGH *et al.* 2010). Two transcriptomes derived from microdissected ovules of *P. squamulatum* and apomictic BC_8 line 58 were sequenced and compared. Forty nine transcripts were verified to be ASGR-carrier chromosome specific and only one transcript was found to be tightly linked to the ASGR (ZENG *et al.* 2011).

The angiosperm lineage is thought to be ~200 million years old, but most of the current species of grasses are found to be derived from a common ancestor that lived about 50-80 million years ago. The grass genomes show tremendous divergence in genome size and chromosome number, despite its relatively recent and monophyletic origin (GAUT 2002; BENNETZEN 2007). Genome sizes of grasses range from species with smaller genomes like rice (430 Mb) and sorghum (750 Mb), to large genome size grasses like maize (2,400 Mb), barley (5,000 Mb), rye (8,000 Mb) and wheat (17 Gb). Although the grass genomes show tremendous

divergence in genome size and chromosome number, comparative mapping of rice, sorghum, barley, maize and various grass species have shown remarkable conservation of marker and gene order (collinearity) among species (FEUILLET and KELLER 2002; BENNETZEN 2007). The comparative analysis of crop genomic sequences, the nature of chromosomes evolution, identification of syntenic blocks, detection of genomic re-arrangements and chromosome reshuffling have become possible with the analysis of genetic and physical maps, and the use of molecular cytogenetic tools (SALSE *et al.* 2008).

Pennisetum belongs to the subfamily Panicoideae in the grass family (Poaceae). Pearl millet (*P. glaucum*) shared a common ancestor with foxtail millet (*Setaria italica*) ~8.3 million years ago and with maize and sorghum ~26 million years ago (BENNETZEN *et al.* 2012). A comparative genomic hybridization (CGH) experiment using apomictic and sexual BC_8 line 58 DNA against the sorghum genome chip identified enrichment of signals from sorghum chromosome 2 in the apomictic sample. The presence of the ASGR-carrier chromosome in the apomictic BC_8 line 58 but not in the sexual, suggests that the sequences enriched for sorghum chromosome 2 are most likely derived from the ASGR-carrier chromosome. Furthermore, a basic local alignment search tool (Blast) search of nucleotide sequences from the 49 ASGR-carrier chromosome transcripts from Zeng *et al.* (2011) against the sorghum and foxtail millet reference genome were found to have the highest similarity to different regions residing on sorghum and foxtail millet chromosome 2. The regions with similarities were collinear across chromosome 2 in the two species. Orthologous gene pair comparisons in nine foxtail millet chromosome and 10 sorghum chromosomes have shown that chromosome 2 of foxtail millet is collinear with sorghum chromosome 2 (BENNETZEN *et al.* 2012). In this study, we sought to identify bacterial artificial chromosome (BAC) clones carrying the ASGR-carrier chromosome

SCAR markers from Zeng *et al.* (2011) and physically mapping the ASGR-carrier chromosome-linked BACs on the ASGR-carrier chromosome to unravel the collinearity of the ASGR-carrier chromosome to chromosome 2 of sorghum and foxtail millet outside of the ASGR.

Materials and Methods

Plant materials

P. glaucum apomictic backcross 8 (BC_8) line 58 carries only the ASGR-carrier chromosome from *P. squamulatum* and is a facultative apomict. Therefore seedlings from BC_8 line 58 plants can be derived via sexual or apomictic reproduction (SINGH *et al.* 2010). Seedlings derived from an apomictic BC_8 line 58 plant were genotyped and confirmed as sexual or apomictic using the ASGR-linked SCAR primer set 787 (5′-ATACATGAGATGAGCAGGCA-3′) and 788 (5′-GCACTACCTACCTATCAAACA-3′). All seedlings were used for DNA extraction and plants with apomictic genotypes were used for collection of root tips for mitotic chromosome preparation.

DNA extraction

DNA extraction was done using a modified cetyltrimethylammonium bromide (CTAB) method. Fifty to 100 mg of young leaf tissue was placed in 150 µl 2× CTAB buffer [100 mM Tris-Cl (pH 8), 20 mM EDTA, 1.4 M NaCl, 2% CTAB, and 0.2% β-mercaptoethanol] and disrupted. An additional 350 µl 2× CTAB buffer was added, mixed well and incubated for 15-30 minutes at 65ºC. 500 µl of 24:1 chloroform:iso-amyl alcohol was added, mixed well and centrifuged for 5 minutes at 14,000 rpm. 400 µl of the supernatant was transferred to a new tube; an equal volume of iso-propanol was added, mixed well and centrifuged for 10 minutes at 14,000 rpm. DNA pellets were washed in 70% ethanol, air dried and suspended in 50 µl 0.5 X Tris-EDTA with 1 mg/ml RNase A (Sigma-Aldrich Co. LLC, St. Louis, MO, USA).

ASGR-carrier chromosome transcripts and identification of BAC clones

Ovule transcripts derived from the ASGR-carrier chromosome have been previously identified (ZENG *et al.* 2011) . The ASGR-carrier chromosome linked bacterial artificial chromosome (BAC) clones were identified through screening the polyhaploid BAC library with α-^{32}P-labeled SCAR probes from the ASGR-carrier chromosome ovule transcripts (ZENG *et al.* 2011). The polyhaploid BAC library reported in ROCHE *et al.* (2002) was derived from the polyhaploid apomictic line MS228-20 that germinated from seed of an open pollinated apomictic F_1 line derived from the cross between *P. glaucum* and *P. squamulatum* (DUJARDIN and HANNA 1986). ~25ng of purified probe was labeled with α-^{32}P dCTP (3000 Ci/mmol) using the Random Primed DNA Labeling kit (Invitrogen, Carlsbad, CA, USA) and purified by passing the reaction through a homemade Sephadex G-50 (Sigma, St. Louis, MO, USA) column, assembled using a Ultrafree®-MC Centrifugal Filter Units (Millipore, Bedford, MA, USA). The polyhaploid BAC filters were pre-hybridized in hybridization buffer (0.5 M sodium phosphate, 7% SDS, 1 mM EDTA, pH 8.0) containing 0.1 mg ml^{-1} denatured salmon sperm DNA at 65°C for at least 4 hours before addition of the labeled, denatured probe. Hybridization was conducted at 65°C overnight followed by two washes at 65°C for 30 minutes each with 1× SSC, 0.1% SDS. After the final wash, the membranes were wrapped with plastic film and exposed to x-ray film overnight at -80°C prior to manual development with Kodak® GBX Developer and Fixer (Thermo Fisher Scientific Inc., Wilmington, DE, USA). Autoradiographs were aligned with the respective filter and BAC addresses manually identified. If the original SCAR primers from Zeng *et al*. (2011) did not identify any BAC clones, additional primers were designed to make probes of longer length for screening the BAC library (Table 2.1). The isolated ASGR-carrier

chromosome BACs were confirmed by PCR with respective ASGR-carrier chromosome SCAR primers with PCR conditions of Zeng *et al*. (2011).

Two highly repetitive *Kpn*I fragments of ~140 and 160 bp, differing only by an 18 bp deletion in the 140-bp *Kpn*I family, has been reported in *Pennisetum* (INGHAM *et al.* 1993). An 137-bp *Hae*III repeat from *P. glaucum* cv. Massue has been shown to be localized to the centromere (KAMM *et al.* 1994). Centromeric BAC clones for *Pennisetum* were identified by screening the polyhaploid BAC library using the 160-bp *Kpn*I fragment used as a probe and hybridization protocol as described previously. DNA from four BAC clones showing the strongest hybridization signals were digested with *Hae*III (New England Biolabs, Ipswich, MA) along with the *Cenchrus ciliaris* centromere BAC (c035B23) as control. The BAC clone containing the largest proportion of the DNA fragments with a 160 bp laddering pattern relative to the *Cenchrus ciliaris* centromere BAC clone was chosen as a centromeric probe for FISH analysis.

Southern hybridization

Southern hybridization of BAC DNA was done to identify the amount of repetitive DNA in the BAC clone based on hybridization with the apomictic BC_8 line 58 genome. DNA from ASGR-carrier chromosome linked BAC clones was extracted using an alkaline lysis BAC mini prep. 10 µl of BAC DNA was digested with 40 units of *Hind*III-HF (New England Biolabs) in 1× NEBuffer 4 at 37ºC from 3 hours to overnight. The digested DNA was run in 1% (w/v) agarose gel. The DNA fragments were transferred from the agarose gel to Genescreen Plus nylon membrane (NEN Life Sciences, Boston, MA, USA) through alkaline transfer. The membrane was rinsed in 2× SSC, air-dried and baked at 80ºC in a vacuum oven for 2 hours. The membrane with BAC DNA fragments was hybridized with α-^{32}P-labeled BC_8 line 58 genomic DNA

fragments and washed as described previously. The hybridized membrane was imaged using Storm phosphor imager system (Amersham Biosciences, Pittsburgh, PA, USA)

Chromosome preparation

For mitotic chromosome spreads, root tips were collected early in the morning from greenhouse-grown apomictic BC_8 line 58 plants, cleaned of soil by rinses in water and pre-treated in a nitrous oxide chamber for 3 to 4 hours at 1 to 1.5 Mpa pressure by placement of the root tips in a 0.5 ml tube with 300 µl distilled water. After the nitrous oxide treatment, the root tips were removed from the tube and fixed in 3:1 ethanol:acetic acid (v/v) for at least 2 days at 4º C prior to use or stored at 4º C in 3:1 ethanol:acetic acid. Fixed root tips were rinsed in distilled water for 10-15 minutes and briefly rinsed in 10 mM citrate buffer (pH 4.5). Root caps were removed from the root tips and 2-3 mm of the meristematic region was incubated in an enzyme mix containing 2% (w/v) cellulose RS (Karlan Research, Santa Rosa, CA), 1% (w/v) pectolyase Y23 (Karlan Research, Santa Rosa, CA), 1% (w/v) macerozyme R 10 (Desert Biologicals, Phoenix, AZ) in citrate buffer (10 mM sodium citrate, 10mM sodium EDTA, pH 5.5) (IWATA *et al.* 2013) for 90 minutes at 37ºC. Slide preparation for chromosome spreads after digestion was done either through enzymatic maceration/air-drying method (FUKUI 1996) or "SteamDrop" method (KIROV *et al.* 2014).

BAC DNA labeling for FISH probes

BAC DNA used for FISH probes were labeled by nick translation. BAC DNA for nick translation was extracted using an alkaline lysis method for high quality BAC DNA extraction based on the protocol from the Section on Neural Gene Expression (SNGE), National Institute of Mental Health (http://www.nimh.nih.gov) with the following modifications. The RNase treatment was done with 10 µl Ambion® RNase cocktail (Life Technologies, Grand Island, NY)

consisting of 5 U of RNase A and 200 U of RNase T1 and the pelleted DNA was suspended in Buffer EB (QIAGEN Inc., Valencia, CA, USA). 1-2 µg of BAC DNA was labeled with either 0.016 mM biotin (bio)-11-dUTP (Roche, Indianapolis, IN) or 0.016 mM digoxigenin (dig)-11-dUTP (Roche), using the nick translation kit (Roche) with 0.05 mM dATP, 0.05 mM dCTP, 0.05 mM dGTP and 0.033 mM dTTP. Centromere probes were prepared by PCR labeling with primer set 3184 (5'-GGTACCCCGTAATAGTGCATTC-3') and 3185 (5'-GAAAATGGTTTCGCAACAAAAG-3') designed from the 160-bp *Kpn*I repeat family sequence (INGHAM *et al.* 1993). PCR labeling was done with 0.25 mM each of dATP/dCTP/dGTP, 0.167 mM dTTP, 0.083 mM of biotin-11-dUTP, 0.25 µM primers, 0.5 units of *Jumpstart™ Taq* DNA polymerase (Sigma-Aldrich), 15-30 ng *P. squamulatum* DNA in a total volume of 20 µl in 1× PCR buffer. The nick translated DNA probe was checked on a 2% (w/v) agarose gel by electrophoresis to obtain desired probe DNA fragments of 200-750 bp in size.

Labeled probes were purified by ethanol precipitation in the presence of 3M sodium acetate. Pelleted DNA was resuspended in 50% formamide in 2× SSC and stored at -20ºC. The labeling efficiency of probes was tested by dot blot. 0.1 to 0.00002 µl of probe were hybridized on Genescreen Plus nylon membrane (NEN Life Sciences) along with a tested positive control probe, fixed under an UV light for 30 seconds and blocked in 1× DIG blocking solution (Roche) in 1× PBS for 20 minutes. The dig-labeled probes were incubated with Anti-dig-AP fragments (Roche) and biotin-labeled probes were incubated with Streptavidin-AP conjugate (Roche) in buffer 1 (0.1 M Tris-HCl (pH 7.5), 1 M NaCl, 0.05% Triton-x100, 2mM $MgCl_2$) for 10 minutes. The hybridized membrane was washed twice in buffer 1 for 5 minutes each, twice in buffer 2 (0.1 M Tris-HCl (pH 9.5), 1 M NaCl, 5 mM $MgCl_2$) for 5 minutes each. The signals from the

probes were detected using NBT/BCIP solution (Roche) in 0.1 M Tris-HCl (pH 9.5), 0.1 M NaCl, 5 mM $MgCl_2$. All the incubation and washes were done on a gentle shaking platform.

Chromosome pretreatment and hybridization for FISH

The slides were screened using phase contrast light microscopy. Only the slides with at least 25 cytoplasm free metaphase chromosome spreads were selected for FISH experiments. Slides were stored at room temperature up to a week, and then stored at 4ºC. Selected slides were dried at 60ºC for 30 minutes in an oven and the subsequent pretreatment of chromosome preparations was done as described by Zhong *et al.* (1996) with some modifications. Unless a volume is noted, the slides were placed in a coplin jar filled with solution. Slides were covered with 24 x 40 mm parafilm strips if small volumes were used. Unless specified all the washes and incubations were done at room temperature. Slides were incubated in 2× SSC for 5 minutes prior to treatment with 150 µl of 100 µg/ml RNase A (Sigma-Aldrich) in 2× SSC at 37ºC for 45 minutes. The slides were washed in 2× SSC three times for 3 minutes, incubated in 0.01 M HCl for 2 minutes, incubated in 5 µg/ml pepsin in 0.01 M HCl solution at 37ºC for 5 to 10 minutes, washed in distilled water for 2 minutes, washed in 2× SSC two times for 3 minutes, incubated in 1× PBS buffer (pH 7.5) with 50 mM $MgCl_2$ for 2 minutes, fixed in 1% formaldehyde with 50 mM $MgCl_2$ in 1× PBS buffer (pH 7.5) for 10 minutes, washed in 1× PBS buffer for 3 minutes, washed in 2× SSC for 3 minutes, dehydrated in 70%, 90% and 100% (v/v) ethanol for 3 minutes each, and let air-dry.

The probes were mixed in a pairwise combination for hybridization to different targets in the experiments. The hybridization mix consisted of 1-5 ng/µl of probe for each target, 50% formamide, 10% dextran sulfate, 75-85 ng/µl salmon sperm DNA and 2× SSC in a final volume of 18-20 µl. Blocking DNA was prepared from *P. squamulatum* by incubating DNA (in

microtube) at 0.4M NaOH concentration for 40 minutes in boiling water. 10-50 ng/µl of blocking DNA, with DNA fragments 100-1000 bp in size, was used to block minor signals from BACs showing repetitive signals. The hybridization mixtures were denatured at 80ºC in a heat block for 5 minutes and snap chilled on ice. The hybridization mixture was applied to the chromosome preparations, covered with 22 × 30 mm coverslip and denatured at 80ºC for 3 minutes on a heat block. Cover slips were sealed with rubber cement, incubated in a moist chamber at 60ºC for 1.5 hours and then allowed to cool down to 37ºC where they subsequently were incubated for 64-67 hours.

Post-hybridization washes and probe detection

The coverslips were removed in 2× SSC and two post-hybridization washes were performed by gently shaking in 50% formamide in 2× SSC at 37ºC for 10 minutes each. Formamide washes were followed by two washes in 1× SSC for 5 minutes at room temperature. The slides were then washed in TNT (100 mM Tris-HCl, pH 7.5; 150 mM NaCl; 0.05% Tween-20) buffer for 5 minutes at room temperature, blocked in TNB (100 mM Tris-HCl, pH 7.5; 150 mM NaCl; 1× DIG blocking solution, Roche) for 30 minutes at 37ºC, and blocked again in 5% (w/v) IgG-free bovine serum albumin (Sigma) in TN (100 mM Tris-HCl, pH 7.5; 150 mM NaCl) for 30 minutes at 37ºC.

Two-color detection was carried out according to Zhong *et al*. (1996) with modifications. The biotin-labeled probes were detected with Texas red in three steps amplification and dig-labeled probes were detected with FITC by two step amplification, with anti-bodies diluted in TNB. All the slides were incubated at 37ºC for at least an hour during anti-body detection steps and each step was followed by three washes in TNT buffer each for 5 minutes at room temperature under dark conditions. The incubation steps for detection were carried out in

following order: 5 µg/ml Texas red®-conjugated avidin (Vector Laboratories, Burlingame, CA), 1 µg/ml anti-dig from sheep conjugated with FITC (Roche) and 1 µg/ml Biotinylated anti-streptavidin (Vector Laboratories), 1.875 µg/ml FITC-conjugated anti-sheep from rabbit (Jackson ImmunoResearch, West Grove, PA) and 5 µg/ml Texas red®-conjugated avidin (Vector Laboratories). The slides were dehydrated in 70%, 90% and 100% (v/v) ethanol each for 5 minutes, and let air-dry. The slides were mounted in Vectashield (Vector Laboratories) containing 4',6-diamidino-2-phenylindole (DAPI; 1.5 µg/ml).

Slides were examined with a Zeiss Axioskop 2 *plus* fluorescence microscope. Fluorescent signals were detected for DAPI (l_{ex} = 360 nm, l_{em} = 420 nm), FITC (l_{ex} = 480 nm, l_{em} = 515 nm), and Texas red (l_{ex} = 560 nm, l_{em} = 645), and monochrome digital images were captured with a charge-coupled device AxioCam camera. Images were pseudo-colored with blue for DAPI, green for FITC, and red for Texas red. Images were compiled using AxioVision, version 4.8 for Windows. Images were merged and adjusted using Adobe® Photoshop CS2 version 9.0.

Results

BlastN of the ASGR-carrier chromosome transcripts

A BlastN search was done using sequences from the ASGR-carrier chromosome ovule transcripts from Zeng *et al*. (2011) as query against the sorghum reference genomic sequence (refseq_genomic; NC_012877.1) and foxtail millet reference genomic sequence (refseq_genomic; NW_004675962.1) at the National Center for Biotechnology Information (NCBI) with a cutoff of e^{-20} (Table 2.2). Only the top result was recorded except when hits were within an *e*-value difference of less than e^{-15} or when the transcript had chromosome 2 as the second highest hit. Transcripts with more than three hits close in *e*-value on multiple chromosomes were considered as multiple hits. Forty out of 49 ASGR-carrier chromosome

sequences had hits to sorghum chromosome 2, while 44 out of 49 transcripts had hits to foxtail millet chromosome 2. Thirty-nine transcripts that hit sorghum chromosome 2 also hit foxtail millet chromosome 2 in a collinear order. The reported hits were distributed collinearly along the length of chromosome 2 in both sorghum and foxtail millet (Table 2.2). One transcript (Ps26_c2838) with the top result on sorghum chromosome 2 had a top hit on chromosome 9 in foxtail millet using the United States Department of Energy Joint Genome Institute (JGI) database. However, this transcript gave a hit to a scaffold placed on chromosome 2 in the Beijing Genomics Institute (BGI) foxtail millet database (http://foxtailmillet.genomics.org.cn). Transcript Ps26_c9369, which mapped to the ASGR (ZENG *et al.* 2011), gave hits to chromosome 4, 6, and 10 in sorghum and chromosome 1 in foxtail millet but not to chromosome 2 in either species.

Isolation of ASGR-carrier chromosome BAC clones

From previous work done in the lab, a total of 98 overgo probes corresponding to the 49 ASGR-carrier chromosome sequences were prepared by taking 40-50 overlapping oligomers from the transcript and probing the polyhaploid BAC library from Roche *et al*. (2002). A total of 152 BAC clones hybridizing with the oligo probes under high probing/washing stringency were identified and two 96-well plates containing this sub pool of BAC clones was made. DNA from the sub-pooled BACs were extracted and used to screen with 19 of the 49 ASGR-carrier chromosome SCAR primers from Zeng *et al.* (2011) (unpublished results). BACs identified to be linked to those 19 ASGR-carrier chromosome SCAR primers were reconfirmed by PCR prior to FISH analysis (Table 2.2).

To identify additional ASGR-carrier chromosome linked BAC clones corresponding to specific locations in sorghum chromosome 2 which were not identified in the BAC PCR

isolation, a hybridization screen of the 152 sub-pooled BAC clones using the ASGR-carrier chromosome SCAR probes from the ASGR-carrier chromosome transcripts was done. BAC clones p039F15 and p039C19 were identified and confirmed by PCR with primers 1482/1483 and 1583/1681 to be linked to the transcripts Ps26_c13655 and Ps26_c1406, respectively (Table 2.2). A hybridization screen using the entire polyhaploid library (ROCHE *et al.* 2002) was used to identify BAC clones containing the ASGR-carrier chromosome SCAR probe from transcript Ps26_c3993. BAC clones with the strongest hybridization signals were grown in culture and colony PCR was done to identify the BAC clones containing SCAR primers 1502/1713. BAC clones p049K20, p188P16, and p258L05 were identified and confirmed to be linked to transcript Ps26_c3993 (Figure 2.1). When SCAR probes from Zeng *et al.* (2011) did not result in distinguishable hybridization signals in polyhaploid BAC library, additional primers were designed to create probes of longer length (Table 2.1). BAC clones p117P06, p024E15 and p181G17 were identified and confirmed by PCR with the SCAR primers 1692/1693, 1542/1543 and 1530/1531 to be linked to transcripts Ps26_c19109, Ps26_c1312 and Ps26_c1279, respectively (Table 2.2).

The centromeric 160 bp *Kpn*I repeat family sequence was used as a probe to screen the polyhaploid BAC library to identify centromeric BAC clones from *Pennisetum*. DNA from BAC clones with the strongest signals on hybridization with the *Kpn*I repeat sequences probe were isolated and digested with *Hae*III restriction enzyme overnight at 37ºC. *Hae*III repeats have been identified to be present in the centromeric sequences of *Pennisetum* (KAMM *et al.* 1994).The digested DNA fragments were run on a 1% (w/v) agarose gel (Figure 2.2). The digested BAC DNA showed a laddering pattern of DNA fragments comparable to *C. ciliaris* centromere BAC

(c035B23) with fragments that are multiples of the 160 bp repeat. We selected BAC p095P07 as the centromeric BAC probe for the FISH experiments.

Southern hybridization of BAC clones

As repetitive DNA within a BAC clone can cause nonspecific FISH hybridization, the amount of repetitive DNA residing in the ASGR-carrier chromosome BAC clones was assayed through Southern blots of the BAC DNA hybridized with labeled total genomic DNA from apomictic BC_8 line 58. BAC DNA from either p104 or p109 was used as an ASGR low copy control while BAC DNA from p800 was used as an ASGR high copy control (Figure 2.3 and Figure 2.4). The repetitiveness within the BAC DNA was determined by the relative number of hybridizing bands and the signal strength of those bands. BACs showing multiple bands with very strong signals (lanes 4, 5, 6, 11, 12, 37 in figure 2.3b; p117P06 in figure 2.4) and BACs showing multiple bands of moderate strength (lanes 21, 30, 33 in figure 2.3b; p024E15 and p181G17 in figure 2.4) are likely to give multiple FISH signals due to the repetitiveness in the genome and considered as high copy. BACs having three to four bands with less strong signals (lanes 13, 17, 18 in figure 2.3b) were considered to be medium copy BACs. BACs with two or fewer bands with very weak hybridization signals (lane 24, 31, 35, 36 in figure 2.3b) were considered as low copy BACs and had priority as FISH probes. This selection and determination of repetitiveness is not sufficient to identify location of the repetitiveness, as there may be BACs (like p800; figure 2.3b and 2.4) with sequences that are repetitive in a specific region in a specific chromosome of the genome compared to more abundant and ubiquitous sequences.

Physical mapping of ASGR-carrier chromosome linked BAC clones

With the objective to physically map ASGR-carrier chromosome linked BAC clones to the ASGR-carrier chromosome to identify their physical order corresponding to relative

locations in the sorghum and foxtail millet chromosome 2, we selected low copy ASGR-carrier chromosome linked BAC clones. Selection of the BACs was based on the corresponding position of BlastN result of each transcript in sorghum/foxtail millet chromosome 2. If necessary we also resorted to using medium and high copy BACs for physical mapping. Six low copy, one medium copy and one high copy BAC clones were physically mapped onto the ASGR-carrier chromosome in apomictic BC_8 line 58 through dual-color FISH (Table 2.2). All unknown ASGR-carrier chromosome linked BAC clones chosen mapped outside of the ASGR on the ASGR-carrier chromosome.

BACs p142D19 (Ps26_c5851), p057M05 (Ps26_c5080), p258L05 (Ps26_c3993) and p236E19 (Ps26_c10535) gave a single signal on the long, non-ASGR arm of the ASGR-carrier chromosome (Figure 2.5a-d). BAC p142D19 was positioned telomeric in location to p057M05 (Figure 2.5g). Similarly, p057M05 was positioned telomeric to both p236E19 and p258L05 in the long arm of ASGR-carrier chromosome (Figure 2.5e-f). The intensity of the centromere signal in the chromosomes differed based on the centromere probe used. Centromere probe from BAC p095P07 gave strong signal on most of the chromosomes in the spreads. However, the ASGR-carrier chromosome was found to have the weakest centromere signal among all the chromosomes in most of the spreads when p095P07 was used as a centromere probe (Figure 2.5a, c, d). When PCR labeled centromere probe was used, the centromere on the ASGR-carrier chromosome had a strong centromere signal as compared most of the other chromosomes in BC_8 line 58 (Figure 2.5b). As BACs p220A02 (Ps26_c583) and p036L06 (Ps26_c 2448) produced stronger and weaker signals on either side of the centromere, minor signals were blocked using *P. squamulatum* DNA (Figure 2.6). After blocking, p036L06 was placed near the centromere on the long arm of the ASGR carrier chromosome (Figure 2.6b), while p220A02 was placed near

the centromere on the short arm of the ASGR-carrier chromosome (Figure 2.6d). Since the centromere is highly condensed in the chromosome, the centromere in the ASGR-carrier chromosome could also be identified as the region with highest DAPI staining region in the chromosome when a centromeric FISH signal was low or not used (Figure 2.6d). Similarly, BAC p285J18 (Ps26_c 2838) was mapped outside the ASGR on the short arm of the ASGR-carrier chromosome (Figure 2.7a). BAC p181G17 (Ps26_c1279), which showed multiple hybridizing bands on the southern blot (Figure 2.4), gave multiple signals in and around the centromere region on the ASGR-carrier chromosome (Figure 2.7b-d).

Discussion

ASGR-carrier chromosome and ovule transcripts

In *P. squamulatum* the apospory is controlled by a single genetic locus transmitted as a large (>50 Mb), hemizygous, heterochromatic chromosomal block called the apospory-specific genomic region (OZIAS-AKINS *et al.* 1998; GOEL *et al.* 2003; AKIYAMA *et al.* 2004). The ASGR in *P. squamulatum* is located near the telomere of the short arm of the ASGR-carrier chromosome and consists of two large repetitive blocks flanking a ~13 Mb single copy region (AKIYAMA *et al.* 2004). Initially, a *P. glaucum* x *P. squamulatum* F_1 was crossed with a *P. glaucum* x *P. purpureum* F_1 and hybrid apomictic individuals with good male fertility were selected (DUJARDIN and HANNA 1989b). Subsequent backcrosses with tetraploid *P. glaucum* yielded a BC_8 line 58 that was shown by FISH to contain only the ASGR-carrier chromosome from *P. squamulatum* (SINGH *et al.* 2010). Ovule transcriptomes derived from microdissected ovule from *P. squamulatum* and BC_8 line 58 were compared and 49 ASGR-carrier chromosome specific transcripts were identified. When SCAR primers from 49 ASGR-carrier chromosome transcripts were mapped on to the *P. squamulatum* x *P. glaucum* F_1s, only one SCAR showed

ASGR specificity while primer specificity was not seen for rest of the 48 transcripts (ZENG *et al.* 2011). Because of the lack of detectable polymorphisms in the ASGR-carrier chromosome, SCARs from the ASGR-carrier chromosome transcripts could not be mapped genetically to the ASGR.

CGH and nucleotide similarity of the ASGR-carrier chromosome

Array-based comparative genomic hybridization has been used to identify copy number variation in maize and rice (BELÓ *et al.* 2010; YU *et al.* 2011). Enrichment of signals from sorghum chromosome 2 in the apomictic BC_8 line 58 genome versus the sexual genome was identified by comparative genomic hybridization against sorghum genome chip (unpublished and in collaboration with Pioneer; Figure 2.8). Since the ASGR-carrier chromosome is present in apomictic BC_8 line 58 plants but not in the sexually derived BC_8 line 58 plants, the sequences showing enrichment for sorghum chromosome 2 were inferred to be derived from the ASGR-carrier chromosome. In addition, BlastN search of the 49 ASGR-carrier chromosome transcripts against the sorghum and foxtail millet genome sequences identified that 82% and 90% of the transcripts shared high sequence similarity to chromosome 2 of sorghum and foxtail millet, respectively. Thirty-nine of the 40 transcripts that shared high sequence similarity to sorghum chromosome 2 also shared high sequence similarity to foxtail millet chromosome 2 in a collinear order. Orthologous gene pair comparisons between the nine foxtail millet chromosomes and 10 sorghum chromosomes showed that chromosome 2 of foxtail millet is collinear with sorghum chromosome 2 (BENNETZEN *et al.* 2012). Collinear distribution of the orthologous genes of ASGR-carrier chromosome transcripts in chromosome 2 of sorghum and foxtail millet genome suggested similar evolution and potential collinearity between the ASGR-carrier chromosome and chromosome 2 of sorghum and foxtail millet.

Identification and mapping of ASGR-carrier chromosome linked BACs

BAC clones linked to 19 ASGR-carrier chromosome transcripts were identified from a sub pool of BAC clones prior to this thesis study (unpublished results). Eight additional BAC clones were isolated from six SCAR probes containing six of the ASGR-carrier chromosome transcripts by screening either the sub pool of BAC clones or the entire polyhaploid library. BAC clones could not be identified for transcripts Ps26_c6373 and Ps26_c1372 when the sub pool of BAC clones was screened, while for transcript Ps26_c2339 no BAC clones could be isolated from the polyhaploid library.

The pearl millet genome has about 69% repetitive DNA (WIMPEE and RAWSON 1979) and the presence of repetitive DNA within a BAC clone can result in non-specific FISH hybridization. Therefore, we assayed the repetitiveness of the ASGR-carrier chromosome BAC clones through Southern blots of BAC DNA hybridized with labeled total genomic DNA from apomictic BC_8 line 58. Although actual repetitiveness of the repeats that are evolutionarily diverged and localized to a specific chromosome or locus in the genome is less likely to be accurately identified through the Southern blot, BAC clones with highly abundant and ubiquitous repeats were identified and avoided. Six low copy BACs, one medium copy BAC and one high copy BAC were selected for physical mapping by FISH. The selected ASGR-carrier chromosome BAC clones mapped outside the ASGR on the ASGR-carrier chromosome. The short arm of the ASGR-carrier chromosome contains the hemizygous and heterochromatic ASGR and therefore the intensity of the DAPI signals was always higher on the short arm as compared to the long arm in the ASGR-carrier chromosome. The intensity of the DAPI signal from the long arm of the ASGR-carrier chromosome gradually declined from the centromere towards the telomeric region (AKIYAMA *et al.* 2004; AKIYAMA *et al.* 2011). When the BACs

classified as low and medium copy were used as probes without blocking DNA, all but two gave single FISH signal and were mapped to a specific location on the ASGR-carrier chromosome. After blocking minor FISH signals using *P. squamulatum* blocking DNA, the two BAC clones p220A02 and p036L06 were successfully positioned near the centromere on the short and long arm of ASGR-carrier chromosome, respectively (Figure 2.6). When the high copy ASGR-carrier chromosome BAC p181G17 was used as probe without blocking, it gave multiple signals in and around the centromere of the ASGR-carrier chromosome (Figure 2.7b, d). The orthologous putative transcribed genes in chromosome 2 of sorghum and foxtail millet for the transcripts linked to the BAC clones p220A02, p036L06 and p181G17 were located at the peri-centromeric region of the chromosome. Major and minor hybridization signals observed around the peri-centromeric region from these BACs suggest the presence of peri-centromeric repeats in these BAC clones. Peri-centromeric regions in the chromosome are heterochromatic in nature and consist of several retrotransposons and tandem repeats (SAUNDERS and HOUBEN 2001; TOPP and DAWE 2006; LIM *et al.* 2007). The centromeric satellite DNA sequences are amongst the most rapidly evolving sequences and there are substantial differences in the tandemly repeated DNA sequences that make up centromeres between species. Genes may remain recognizably similar while the repeats diverge even within a genus (HESLOP-HARRISON *et al.* 2003). The lack of hybridization signals in other chromosomes in BC_8 line 58 from the BACs located in the peri-centromeric region of ASGR-carrier chromosome indicates the divergence of the repeats found in the mapped ASGR-carrier chromosome BAC clones from peri-centromeric repeats in *P. glaucum*.

Collinearity of the ASGR-carrier chromosome

The physically mapped ASGR-carrier chromosome BACs show large-scale collinearity between the ASGR-carrier chromosome and chromosome 2 of sorghum and foxtail millet, outside the ASGR (Figure 2.9). The centromeres in chromosome 2 of sorghum and foxtail millet are located at ~30-35 Mb and ~17-20 Mb, respectively (http://ensembl.gramene.org). BAC p285J18 linked to the Ps26_c2838 transcript having orthologous gene in sorghum chromosome 2 at 11.45 Mb was mapped between the ASGR and the centromere on the short arm of the ASGR-carrier chromosome. In foxtail millet the Ps26_c2838 sequences showed similarity to chromosome 9 in United States Department of Energy JGI database, but when this transcript was searched against the Beijing Genomic Institute foxtail millet database, this transcript gave a hit to a scaffold placed at ~10 Mb on chromosome 2. BAC p142D19 linked to transcript Ps26_c5851 was ordered and found to be telomeric to p057M05. And, BAC clone p057M05 linked to transcript Ps26_c5080 was ordered separately with BACs p236E19 and p258L05, and in both cases p057M05 was mapped telomeric to either of the BAC clones. The orthologous transcribed gene for transcript Ps26_c3993 in foxtail millet chromosome 2 was located at 36.46 Mb, but in sorghum it only gave hit to a gene in chromosome 6 with an *e*-value $>e^{-20}$. When the transcripts located on 5' and 3' of the Ps26_c3993 ortholog in foxtail millet was used as query for BlastN search against the sorghum reference genomic sequences in NCBI, they gave hits to sorghum chromosome 2 at ~64 Mb regions. This suggests that the gene orthologous to Ps26_c3993 in sorghum chromosome 2 might have been subjected to deletion, translocation or divergence overtime in the course of evolution. Based on our mapping data the synteny starts from ~11 Mb and ~10 Mb in chromosome 2 of sorghum and foxtail millet and from about mid-point of the centromere and the ASGR on the ASGR-carrier chromosome. Whether this synteny extends to

the very beginning of the chromosome 2 and right outside the ASGR is still unknown. Since the *in silico* data shows the homology between transcript Ps26_c194 and orthologous genes at ~5 Mb in chromosome 2 of sorghum and foxtail millet, we expect the collinearity between these chromosomes extends outside the ASGR in the short arm.

Evolution of the ASGR-carrier chromosome

The level of variation detected for the ASGR suggests a relatively recent origin of the ASGR (AKIYAMA *et al.* 2011). The possibility of the hemizygous ASGR on the ASGR-carrier chromosome being derived by introgression of alien chromatin from another plant species has previously been discussed (AKIYAMA *et al.* 2004). In light of our observation that the ASGR-carrier chromosome is collinear to chromosome 2 of sorghum and foxtail millet outside of the ASGR, this theory holds strong stance in the evolution of the ASGR on the ASGR-carrier chromosome in *P. squamulatum*. However, the results from this study can only decipher the evolution of the ASGR-carrier chromosome in *P. squamulatum*, while the evolution of the ASGR-carrier chromosome in other *Pennisetum* and *Cenchrus* species is yet to be determined. The mapping of the ASGR-linked BAC P208, chromatin density, and distribution and intensity of the Opie-2 signals in 13 different *Pennisetum* and *Cenchrus* species show morphological differences in the ASGR-carrier chromosome across species (AKIYAMA *et al.* 2011). Physical mapping of the ASGR-carrier chromosome BAC clones from this study in other apomictic *Pennisetum* species could be used to study the evolution of the ASGR-carrier chromosome across different species in the genus *Pennisetum*.

Location of the ASGR to the telomeric and condensed region of the chromosome is thought to be an ancestral state (AKIYAMA *et al.* 2011). The ASGR in *P. squamulatum* could have arisen from the introgression of chromatin from chromosomes of other species into the

telomeric regions of ASGR-carrier chromosome during interspecific hybridization events. In all *Pennisetum* species, the ASGR is located near the telomere of the chromosome, while it is interstitial in *Cenchrus* species and inverted relative to *P. squamulatum* (GOEL *et al.* 2003; GOEL *et al.* 2006). A repeat-poor region is flanked by Opie-2-rich regions in the ASGR of both *P. squamulatum* and *C. ciliaris*. Opie-2-like sequences were found to be abundant in *P. squamulatum* only at the ASGR (AKIYAMA *et al.* 2004), while in *C. ciliaris* these sequences were associated with the centromeric regions of all chromosomes (AKIYAMA *et al.* 2005). Since the Opie-2 like repeat has been detected as part of ASGR in seven out of 12 aposporous species, it is speculated that the association of this repeat to the ASGR was derived either by translocation of repeat-poor portion of the ASGR to a repeat-rich region in genome or by transposition and accumulation of retrotransposons in proximity to the ASGR (AKIYAMA *et al.* 2011). Additional sequencing information from the locus along with FISH data for the ASGR in other *Pennisetum* and *Cenchrus* species will be required to understand more about the evolution and origin of the ASGR on the ASGR-carrier chromosome.

In the present study we identified ASGR-carrier chromosome BAC clones linked to transcripts with orthologous genes on chromosome 2 of sorghum and foxtail millet. Eight ASGR-carrier chromosome BAC clones were physically mapped across the ASGR-carrier chromosome through FISH. The physical mapping result shows large-scale collinearity between ASGR-carrier chromosome and chromosome 2 of sorghum and foxtail millet outside of the ASGR. This suggests that the ASGR was derived by introgression of chromatin from other chromosomes into the chromosome evolutionarily similar to chromosome 2 of sorghum and foxtail millet, resulting in the present day ASGR-carrier chromosome for *P. squamulatum*. This study will support as a key in the evolutionarily study of ASGR in the *Pennisetum* and *Cenchrus*

species. In addition, the lack of synteny of the ASGR to chromosome 2 of sorghum and foxtail millet, will be substantial in narrowing down the search for putative apomixis candidate genes based on sequencing of BC_8 line 58 sexual and apomictic plants.

References

AKIYAMA, Y., J. A. CONNER, S. GOEL, D. T. MORISHIGE, J. E. MULLET *et al.*, 2004 High-resolution physical mapping in Pennisetum squamulatum reveals extensive chromosomal heteromorphism of the genomic region associated with apomixis. Plant Physiology **134:** 1733-1741.

AKIYAMA, Y., S. GOEL, J. A. CONNER, W. W. HANNA, H. YAMADA-AKIYAMA *et al.*, 2011 Evolution of the apomixis transmitting chromosome in Pennisetum. BMC Evolutionary Biology **11:** 289-289.

AKIYAMA, Y., P. OZIAS-AKINS and W. W. HANNA, 2005 High-resolution physical mapping reveals that the apospory-specific genomic region (ASGR) in Cenchrus ciliaris is located on a heterochromatic and hemizygous region of a single chromosome. Theoretical and Applied Genetics **111:** 1042-1051.

ALBERTINI, E., A. PORCEDDU, F. FERRANTI, L. REALE, G. BARCACCIA *et al.*, 2001 Apospory and parthenogenesis may be uncoupled in Poa pratensis: a cytological investigation. Sexual Plant Reproduction **14:** 213-217.

BELÓ, A., M. K. BEATTY, D. HONDRED, K. A. FENGLER, B. LI *et al.*, 2010 Allelic genome structural variations in maize detected by array comparative genome hybridization. TAG. Theoretical And Applied Genetics. Theoretische Und Angewandte Genetik **120:** 355-367.

BENNETZEN, J. L., 2007 Patterns in grass genome evolution. Current Opinion in Plant Biology **10:** 176-181.

BENNETZEN, J. L., J. SCHMUTZ, H. WANG, R. PERCIFIELD, J. HAWKINS *et al.*, 2012 Reference genome sequence of the model plant Setaria. Nature Biotechnology **30:** 555-561.

CARMAN, J. G., 1997 Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony [electronic resource]. Biological journal of the Linnean Society **61:** 51-94.

CATANACH, A. S., S. K. ERASMUSON, E. PODIVINSKY, B. R. JORDAN and R. BICKNELL, 2006 Deletion Mapping of Genetic Regions Associated with Apomixis in Hieracium, pp. 18650. National Academy of Sciences.

CONNER, J. A., S. GOEL, G. GUNAWAN, M. M. CORDONNIER-PRATT, V. E. JOHNSON *et al.*, 2008 Sequence analysis of bacterial artificial chromosome clones from the apospory-specific genomic region of Pennisetum and Cenchrus. Plant Physiology **147:** 1396-1411.

DUJARDIN, M., and W. HANNA, 1986 An apomictic polyhaploid obtained from a pearl millet x Pennisetum squamulatum apomictic interspecific hybrid. Theoretical & Applied Genetics **72:** 33.

DUJARDIN, M., and W. W. HANNA, 1989 Developing apomictic pearl millet - characterization of a BC3 plant. Journal of Genetics & Breeding **43:** 145-149.

FEUILLET, C., and B. KELLER, 2002 Comparative Genomics in the Grass Family: Molecular Characterization of Grass Genome Structure and Evolution. Annals of Botany **89:** 3-10.

FUKUI, K., 1996 Plant chromosome at mitosis, pp. 1-17 in *Plant chromosomes laboratory methods*, edited by K. FUKUI and S. NAGAYAMA. CRC Press, Boca Raton, FL.

GAUT, B. S., 2002 Evolutionary dynamics of grass genomes. New Phytologist **154:** 15-28.

GOEL, S., Z. B. CHEN, Y. AKIYAMA, J. A. CONNER, M. BASU *et al.*, 2006 Comparative physical mapping of the apospory-specific genomic region in two apomictic grasses: Pennisetum squamulatum and Cenchrus ciliaris, pp. 389-400 in *Genetics*. Genetics Society of America, Bethesda; USA.

GOEL, S., C. ZHENBANG, J. A. CONNER, Y. AKIYAMA, W. W. HANNA *et al.*, 2003 Delineation by Fluorescence in Situ Hybridization of a Single Hemizygous Chromosomal Region Associated With Aposporous Embryo Sac Formation in Pennisetum squamulatum and Cenchrus ciliaris. Genetics **163:** 1069.

HAND, M. L., and A. M. G. KOLTUNOW, 2014 The Genetic Control of Apomixis: Asexual Seed Formation. Genetics **197:** 441-450.

HESLOP-HARRISON, J. S., A. BRANDES and T. SCHWARZACHER, 2003 Tandemly repeated DNA sequences and centromeric chromosomal regions of Arabidopsis species. Chromosome Research **11:** 241-253.

INGHAM, L. D., W. W. HANNA, J. W. BAIER and L. C. HANNAH, 1993 Origin of the main class of repetitive DNA within selected Pennisetum species. Molecular and General Genetics **238:** 350-356.

IWATA, A., A. L. TEK, M. M. S. RICHARD, B. ABERNATHY, A. FONSÊCA *et al.*, 2013 Identification and characterization of functional centromeres of the common bean. Plant Journal **76:** 47-60.

KAMM, A., J. S. HESLOP-HARRISON and T. SCHMIDT, 1994 Molecular and physical organization of highly repetitive, undermethylated DNA from Pennisetum glaucum. Molecular & general genetics : MGG **244:** 420-425.

KIROV, I., M. DIVASHUK, K. VAN LAERE, A. SOLOVIEV and L. KHRUSTALEVA, 2014 An easy "SteamDrop" method for high quality plant chromosome preparation. Molecular Cytogenetics (17558166) **7:** 1-20.

LEBLANC, O., D. GRIMANELLI, D. GONZÁLEZ DE LEÓN and Y. SAVIDAN, 1995 Detection of the apomictic mode of reproduction in maize-Tripsacum hybrids using maize RFLP markers. Theoretical and Applied Genetics **90:** 1198-1203.

LIM, K.-B., T.-J. YANG, Y.-J. HWANG, J. S. KIM, J.-Y. PARK *et al.*, 2007 Characterization of the centromere and peri-centromere retrotransposons in Brassica rapa and their distribution in related Brassica species. The Plant Journal: For Cell And Molecular Biology **49:** 173-183.

NOGLER, G. A., 1984 Gametophytic apomixis, pp. 474-518 in *Embryology of angiosperms / edited by B.M. Johri*. Berlin : Springer-Verlag, 1984.

NOYES, R. D., and L. H. RIESEBERG, 2000 Two independent loci control agamospermy (apomixis) in the triploid flowering plant Erigeron annuus. Genetics **155:** 379-390.

OZIAS-AKINS, P., 2006 Apomixis: Developmental Characteristics and Genetics. Critical Reviews in Plant Sciences **25:** 199-214.

OZIAS-AKINS, P., Y. AKIYAMA and W. W. HANNA, 2003 Molecular characterization of the genomic region linked with apomixis in Pennisetum/Cenchrus. Functional & Integrative Genomics **3:** 94-104.

OZIAS-AKINS, P., D. ROCHE and W. W. HANNA, 1998 Tight Clustering and Hemizygosity of Apomixis-Linked Molecular Markers in Pennisetum squamulatum Implies Genetic Control of Apospory by a Divergent Locus that may have no Allelic Form in Sexual Genotypes, pp. 5127. National Academy of Sciences of the United States of America.

OZIAS-AKINS, P., and P. J. VAN DIJK, 2007 Mendelian Genetics of Apomixis in Plants. Annual Review of Genetics **41:** 509-537.

ROCHE, D., J. CONNER, M. BUDIMAN, D. FRISCH, R. WING *et al.*, 2002 Construction of BAC libraries from two apomictic grasses to study the microcolinearity of their apospory-specific genomic regions. Theoretical & Applied Genetics **104:** 804-812.

SALSE, J., S. BOLOT, M. THROUDE, V. JOUFFE, B. PIEGU *et al.*, 2008 Identification and Characterization of Shared Duplications between Rice and Wheat Provide New Insight into Grass Genome Evolution. The Plant Cell**:** 11.

SAUNDERS, V. A., and A. HOUBEN, 2001 The pericentromeric heterochromatin of the grass Zingeria biebersteiniana (2n = 4) is composed of Zbcen1-type tandem repeats that are intermingled with accumulated dispersedly organized sequences. Genome **44:** 955-961.

SAVIDAN, Y., 2000 Apomixis: genetics and breeding. Plant breeding reviews **18:** 13-86.

SINGH, M., J. A. CONNER, Y. J. ZENG, W. W. HANNA, V. E. JOHNSON *et al.*, 2010 Characterization of apomictic BC7 and BC8 pearl millet: meiotic chromosome behavior and construction of an ASGR-carrier chromosome-specific library. Crop Science **50:** 892-902.

TOPP, C. N., and R. K. DAWE, 2006 Reinterpreting pericentromeric heterochromatin. Current Opinion in Plant Biology **9:** 647-653.

VAN DIJK, P. J., I. C. Q. TAS, M. FALQUE and T. BAKX-SCHOTMAN, 1999 Crosses between sexual and apomictic dandelions (Taraxacum). II. The breakdown of apomixis. Heredity **83:** 715-721.

WIMPEE, C. F., and J. R. RAWSON, 1979 Characterization of the nuclear genome of pearl millet. Biochimica Et Biophysica Acta **562:** 192-206.

YU, P., C. WANG, Q. XU, Y. FENG, X. YUAN *et al.*, 2011 Detection of copy number variations in rice using array-based comparative genomic hybridization. BMC Genomics **12:** 372-372.

ZENG, Y., J. CONNER and P. OZIAS-AKINS, 2011 Identification of ovule transcripts from the Apospory-Specific Genomic Region (ASGR)-carrier chromosome. BMC Genomics **12:** 206-206.

ZHONG, X., A. V. KAMMEN, J. HANS DE JONG, P. ZABEL, P. F. FRANSZ *et al.*, 1996 High-resolution mapping on pachytene chromosomes and extended DNA fibres by fluorescence in-situ hybridisation. Plant molecular biology reporter **14:** 232-242.

Table 2.1. Primers designed for isolation of ASGR-carrier chromosome BACs. Tm denotes annealing temperature.

Transcript Name	Primer Number	Sequence (5'-3')	Tm (°C)
Ps26_11544	3106	ATGGCAATGCTGAGCTTCGGCG	60
	3107	TGCTGGACTAGATGGGTCGAAACC	
Ps26_c1312	3108	CCTGCAGTGCCTGGTTTGCGTA	67
	3109	GGCACTGGAGAGTTCGGCTTGC	
Ps26_c1279	3110	TCCTCTTCCTCTCCCTGGGCCT	60
	3111	CAGTCCTACCAGCTCTGGCCGT	
Ps26_c19109	3112	GCAATTCTTTCTCCTCTGCAGCC	65
	3113	CGTCCCATCATGTCTGTGTCTGT	

Table 2.2. BlastN analysis of the ASGR-carrier chromosome transcripts with corresponding BAC information. Table headings from left to right. Transcript name denotes the name of the Ps26 transcript according to Zeng *et al*. (2011); hits refer to the chromosome number or scaffold number of the BlastN hit; location denotes the position of the hit (s) on the chromosome in Mb; FTM stands for foxtail millet; multiple hits refers to three or more significant chromosome or scaffold hits for the transcript; the underlined BAC clones represent BACs isolated prior to the beginning of this thesis study; amount of repetitiveness identified in the BAC clones is denoted by (H): High, (M): Medium, (L): Low; * represents BACs physically mapped via FISH.

Transcript Name	Sorghum Hits	Sorghum e-value	Sorghum Location (Mb)	FTM Hits	FTM e-value	FTM Location (Mb)	BACs Linked (Repetitiveness)
Ps26_c17388	Multiple hits	-------	-------	Multiple hits	-------	-------	
Ps26_c9369	chr-4	6.00E-51	40.13	chr-1	3.00E-46	23.80	p197C01(L), p079H10 (L), p033B01 (L)
	chr-10	6.00E-51	44.05				
	chr-6	3.00E-49	29.79				
Ps26_c2653	Multiple hits	-------	-------	chr-7	0	Multiple locations	
Ps26_c194	chr-2	2.00E-92	5.78	chr-2	0	5.55	p039E23 (L), p131J01 (M), p132I10 (H), p164O24 (L)
	chr-6	4.00E-87	47.12				
	chr-3	2.00E-85	12.53				
Ps26_c1422	chr-2	4.00E-114	8.17	chr-2	0	7.32	
Ps26_c6744	chr-2	1.00E-90	8.18	chr-2	2.00E-118	7.33	
Ps26_c28392	No hit	-------	-------	chr-2	9.00E-30	8.41	

Ps26_c1472	chr-2	2.00E-40	9.31	chr-2	4.00E-139	8.57	
Ps26_c2838	chr-2	2.00E-77	11.45	chr-9	2.00E-121	44.51	p056D06 (L), p068G08 (H), p105C21 (M), p285J18 (L)*
Ps26_c19109	chr-2	3.00E-78	20.72	chr-2	2.00E-88	15.55	p117P06 (H)
Ps26_c583	chr-2	7.00E-114	21.62	chr-2	0	15.73	p220A02 (L)*
Ps26_c1312	chr-10	5.00E-171	1.61	chr-2	0	21.71	p024E15 (H)
Ps26_c9776	chr-7	0	18.75	chr-2	0	22.23	p118I15 (H), p230J07 (H), p276O23 (H)
Ps26_c1279	chr-2	1.00E-88	45.16	chr-2	3.00E-140	24.21	p181G17 (H)*
Ps26_c14318	chr-2	1.00E-83	49.44	chr-2	0	24.92	
Ps26_c2785	chr-2	5.00E-33	52.12	chr-2	4.00E-104	26.19	p141G14 (H)
Ps26_c2448	chr-2	4.00E-102	59.16	chr-2	0	30.62	p036L06 (L)*, p182P03 (H), p182P04 (H)
Ps26_c11544	chr-1	8.00E-79	30.66	chr-9	7.00E-92	20.44	
				chr-2	4.00E-25	32.08	
Ps26_c6131	chr-2	0	62.31	chr-2	0	34.47	p020J22 (L)
Ps26_c13157	chr-2	4.00E-30	62.94	chr-2	6.00E-111	34.86	p118L09 (M), p151I17 (L)
Ps26_c8165	chr-2	4.00E-128	63.65	chr-2	0	35.73	p172K05 (H)
Ps26_c3993	chr-6	1.00E-91	55.23	chr-2	3.00E-102	36.46	p049K20 (M), p188P16 (L), p258L05 (L)*
Ps26_33813	chr-2	1.00E-81	64.44	chr-2	5.00E-70	36.74	
Ps26_c2807	chr-2	8.00E-32	65.12	chr-2	3.00E-125	37.44	
Ps26_c13922	chr-2	1.00E-63	65.55	chr-2	3.00E-128	37.91	
Ps26_c10535	chr-2	6.00E-115	65.96	chr-5	6.00E-172	40.71	p021H05 (M), p236E19 (L)*
				chr-2	3.00E-93	38.32	

Ps26_c13655	chr-7	5.00E-156	59.79	chr-2	0	38.79	p039F15 (L)
Ps26_c6373	chr-2	1.00E-48	67.25	chr-2	4.00E-100	39.73	
Ps26_c4150	chr-2	0	67.73	chr-2	0	40.26	
	chr-10	0	8.80				
Ps26_c7587	chr-2	0	68.39	chr-2	0	40.98	p236E36 (L)
Ps26_c32589	chr-2	4.00E-41	69.08	chr-2	4.00E-136	41.65	
Ps26_c4364	chr-2	3.00E-122	69.98	chr-2	0	42.48	
Ps26_c338	chr-2	0	70.34	chr-2	0	42.86	p124L14 (L), p080J04 (M)
Ps26_c1878	chr-2	0	70.62	chr-2	0	43.14	
Ps26_c2388	chr-2	0	71.14	chr-2	0	43.64	p097O04 (L)
Ps26_c10331	chr-2	6.00E-35	71.18	chr-2	1.00E-63	36.48	
	chr-9	4.00E-38	59.31	chr-6	4.00E-113	5.50	
Ps26_c5080	chr-2	3.00E-144	72.07	chr-2	4.00E-169	44.45	p053C12 (L), p057M05 (L)*
Ps26_c3546	chr-2	1.00E-125	72.26	chr-2	3.00E-140	44.62	p110L05 (L)
Ps26_c3609	chr-2	3.00E-57	72.61	chr-2	3.00E-103	44.94	
Ps26_c5210	chr-2	0	73.14	chr-2	0	45.46	p138D03 (L), p275N19
Ps26_c2552	chr-2	5.00E-111	73.71	chr-2	0	45.96	
	chr-8	1.00E-105	39.54	chr-3	5.00E-117	46.54	
Ps26_c1406	chr-2	0	73.76	chr-2	0	46.01	p035C19 (L)
Ps26_c30198	chr-2	0	74.46	chr-2	0	46.60	
	chr-3	0	10.63	chr-5	0	5.13	
Ps26_c1372	chr-2	0	75.03	chr-2	0	47.03	
Ps26_c3455	chr-2	0	75.58	chr-2	0	47.43	p161K15 (M), p068E18 (H), p073M07 (H)

Ps26_c5851	chr-2	2.00E-36	75.87	chr-2	2.00E-55	47.64	p142D19 (M)*, p233B10 (M)
	chr-1	2.00E-29	58.89				
Ps26_c704	chr-2	0	77.65	chr-2	0	49.01	
Ps26_c2339	chr-2	0	77.78	chr-2	0	49.09	
Ps26_c22381	chr-2	7.00E-35	77.86	chr-2	3.00E-53	49.11	

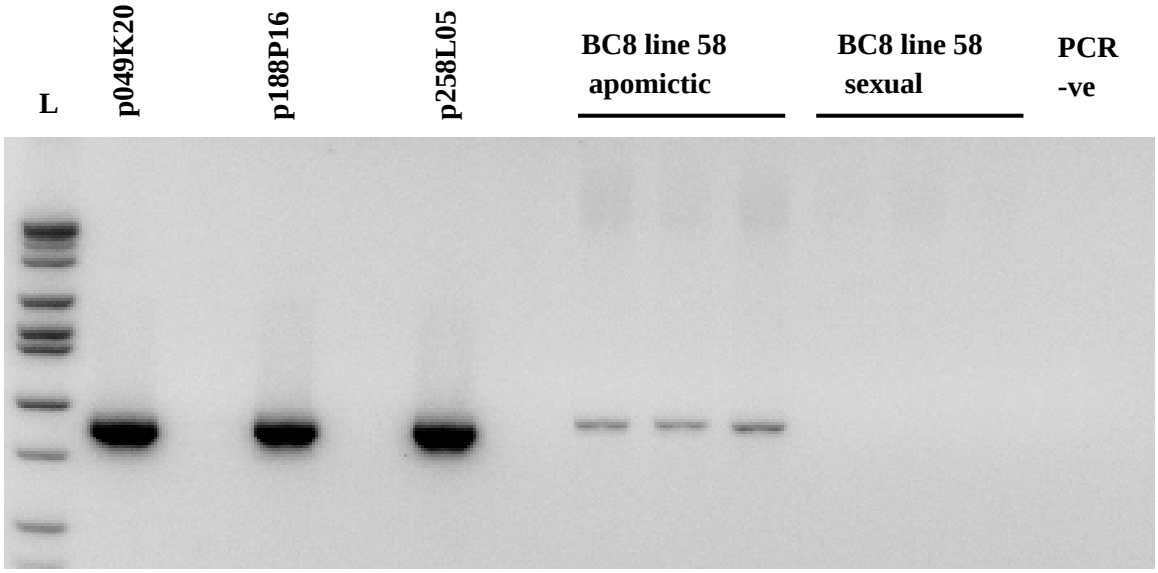

Figure 2.1. Example of PCR confirmation of BAC clones containing the ASGR-carrier chromosome SCAR primers. This figure shows a gel image for PCR amplification of the ASGR-carrier chromosome SCAR 1502/1713 linked to transcript Ps26_c3993. L stands for DNA size marker/ladder, and PCR –ve is a reaction without DNA template.

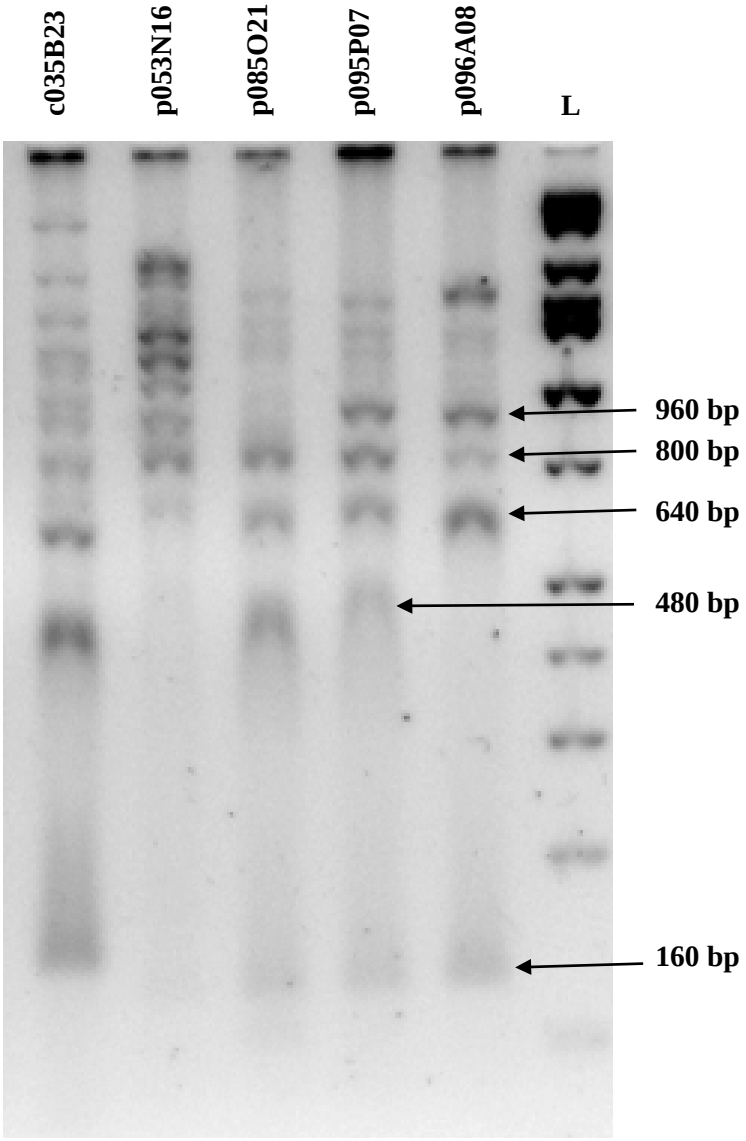

Figure 2.2. DNA fragments generated from a *Hae*III digestion of BAC clones with strong hybridization signals with the 160 bp *Kpn*I centromeric repeat sequences. *Cenchrus ciliaris* centromere BAC c035B23 was used as control, L stands for DNA size marker/ladder.

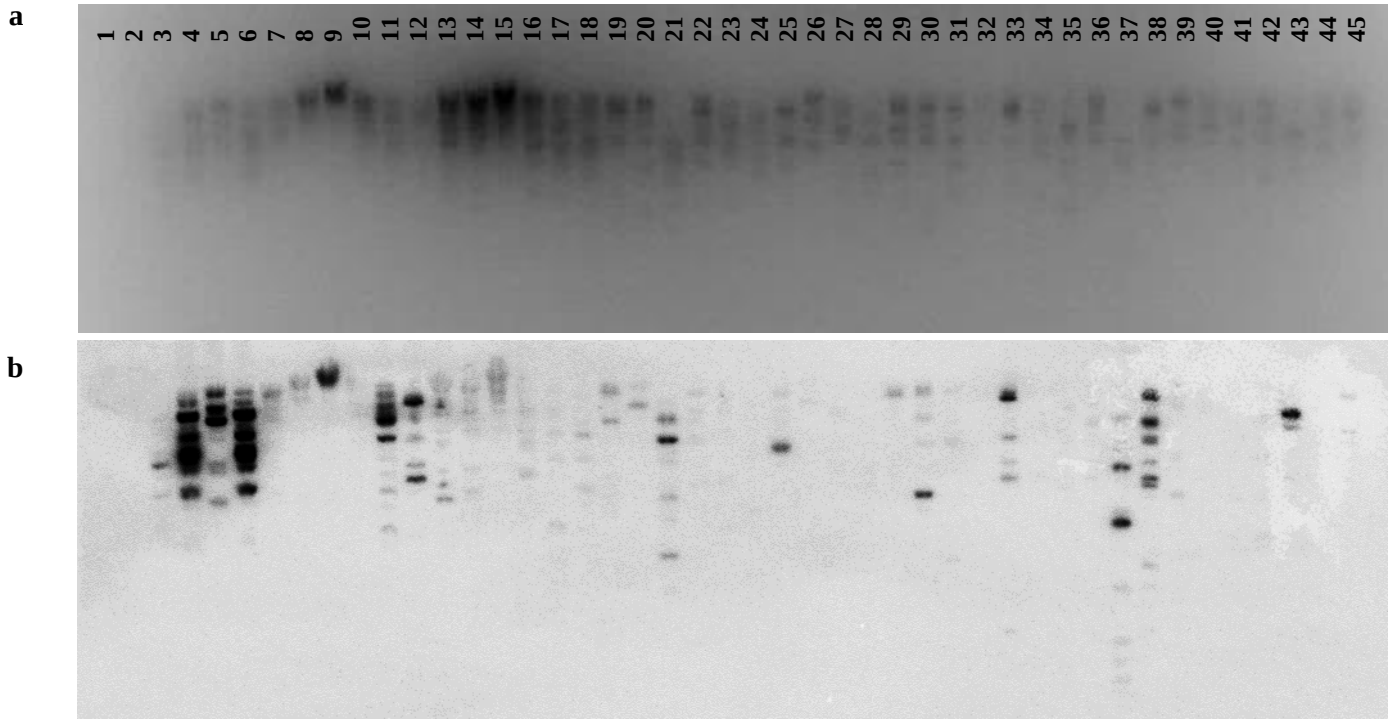

Figure 2.3. Assessment of repetitiveness of the ASGR-carrier chromosome BACs by Southern blot. **a**, Ethidium Bromide stained gel image of BAC DNA fragments generated from a *Hind*III digestion . **b**, results from the southern hybridization of the *Hind*III digested BAC DNA fragments with α-^{32}P-labeled BC_8 line 58 genomic DNA (Lane 1-45: p039E23, p110L05, p131J01, p172K05, p132I10, p182P03, p117P06, p138D03, p182P04, p020J22, p068E18, p118I15, p142D19, p021H05, p068G08, p097O04, p118L09, p161K15, p073E12, p124L14, p073M07, p164O24, p151I17, p036L06, p205G14, p056D06, p079H10, p057M05, p080J04, p216K11, p220A02, p231N09, p276O23, p236E36, p285J18, p236E19, p105C08, p230J07, p233B10, p197C01, p111A16, p141C03, p141G14, p104, p800). Underlined BACs were used for physical mapping in FISH experiments.

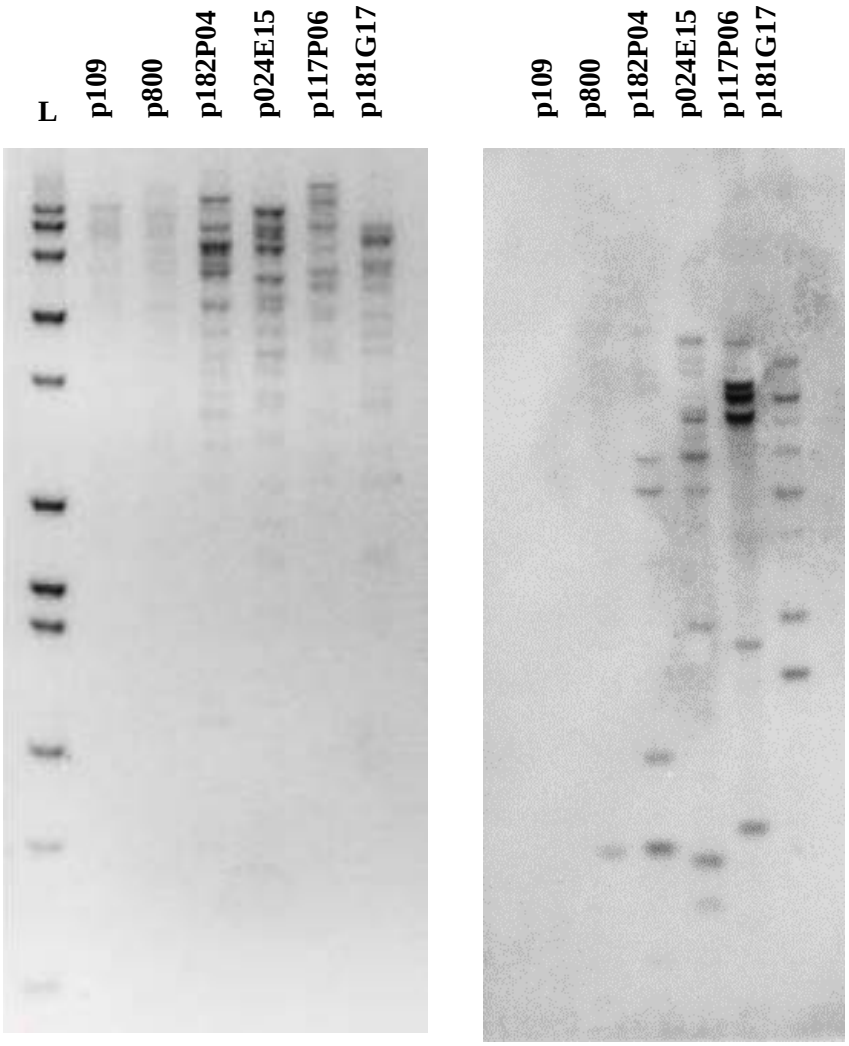

Figure 2.4. Southern blot of the ASGR-carrier chromosome BAC clones with apomictic BC_8 line 58 DNA. Left: Ethidium Bromide stained gel image of BAC DNA fragments generated from a *Hind*III digestion. Right: results from the southern hybridization of the *Hind*III digested BAC DNA fragments with α-^{32}P-labeled BC_8 line 58 genomic DNA.

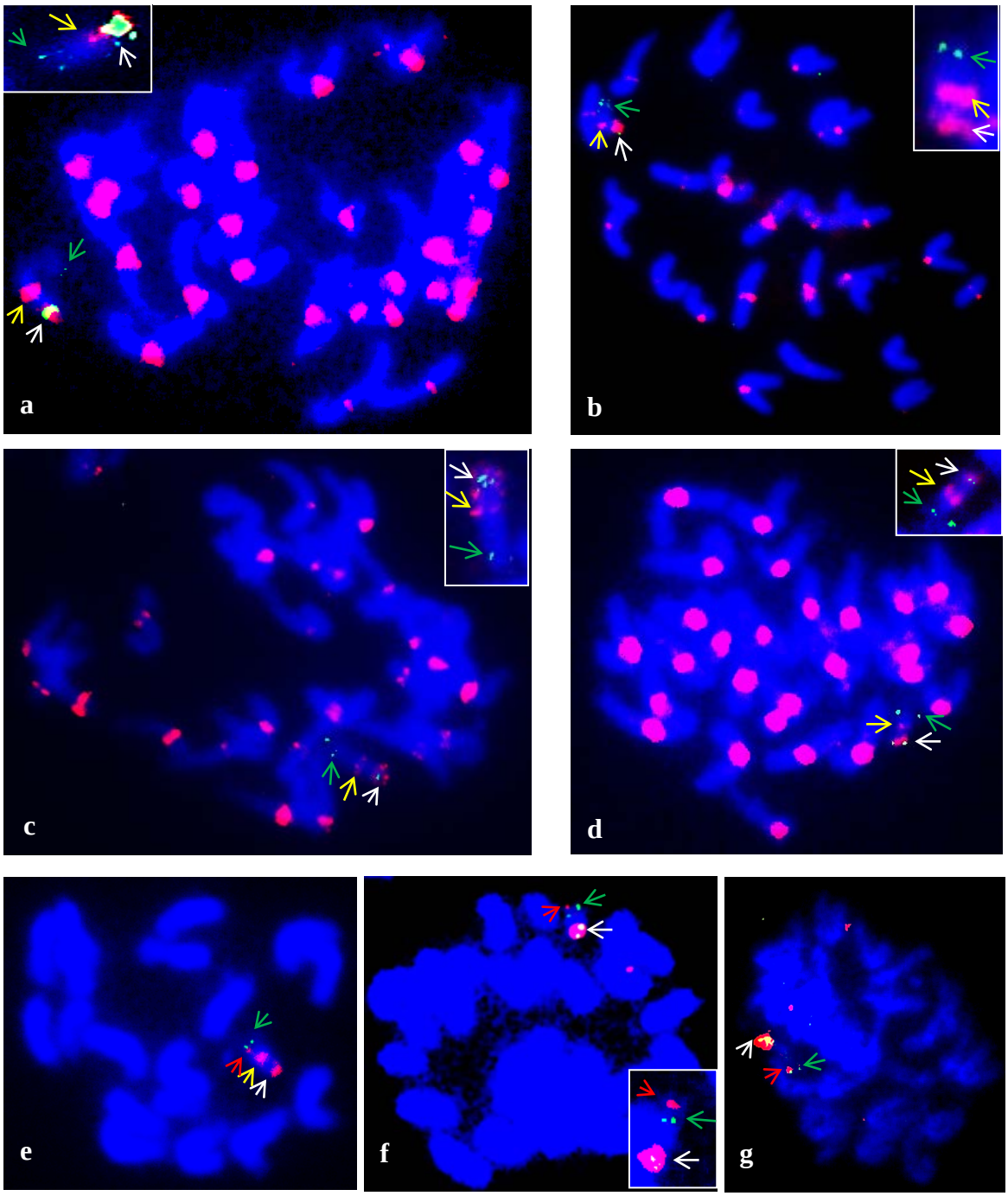

Figure 2.5. Physical mapping of the ASGR-carrier chromosome BAC clones p142D19, p236E19, p258L05 and p057M05 on metaphase chromosome spreads of apomictic BC_8 line 58 root tips. All inset images show an enlarged ASGR-carrier chromosome from a separate spread.

Green arrow indicates: **a** & **g**, p142D19 (green signal); **b** & **e**, p057M05 (green signal); **c**, p236E19 (green signal); **d** & **f**, p258L05 (green signal). **a**-**g**: white arrow indicates the ASGR (p800 + p109 or p303) signals (red or red + green). **a**-**e**: yellow arrow indicates centromere signal (red). **e**, red arrow indicates p236E19 signal (red). **f** & **g**, red arrow indicates p057M05 signal (red).

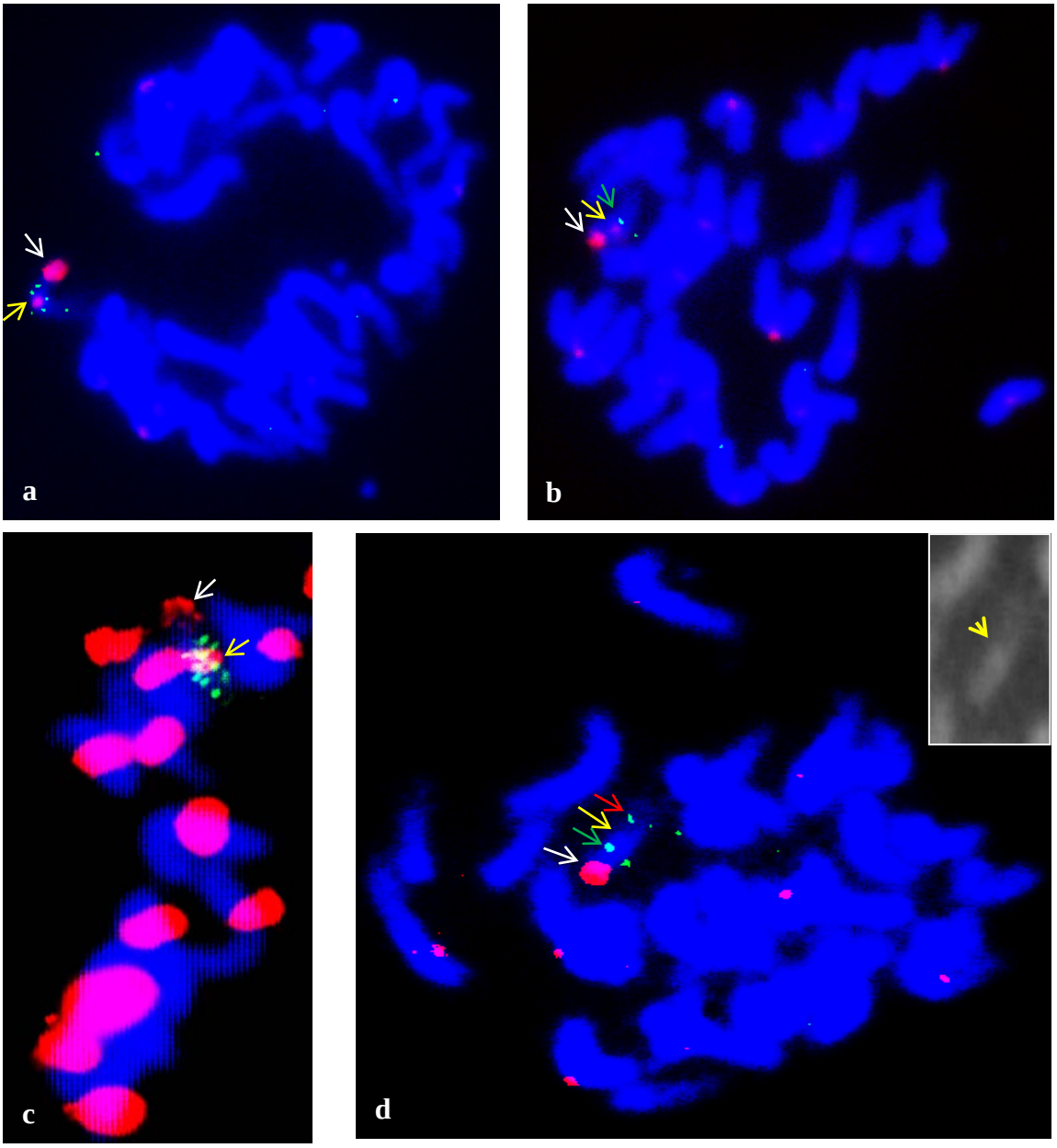

Figure 2.6. Physical mapping of the ASGR-carrier chromosome using BACs p036L06 and p220A02 on metaphase chromosome spreads of apomictic BC_8 line 58 root tips.. The inset image represents an inverted DAPI image of the ASGR-carrier chromosome. **a-d**, white arrow indicates p800/ASGR signal (red). **a-c**, yellow arrow indicates centromere signal (red). **a**, unblocked p036L06 signals (green). **b**, green arrow indicates p036L06 signals (green) after blocking with

Ps26 DNA. **c**, unblocked p220A02 signals (green). **d**, green arrow indicates p220A02 signal (green) after blocking with Ps26 DNA, yellow arrow indicates possible location of centromere based on strength of DAPI, red arrow indicates p236E19 signal (green).

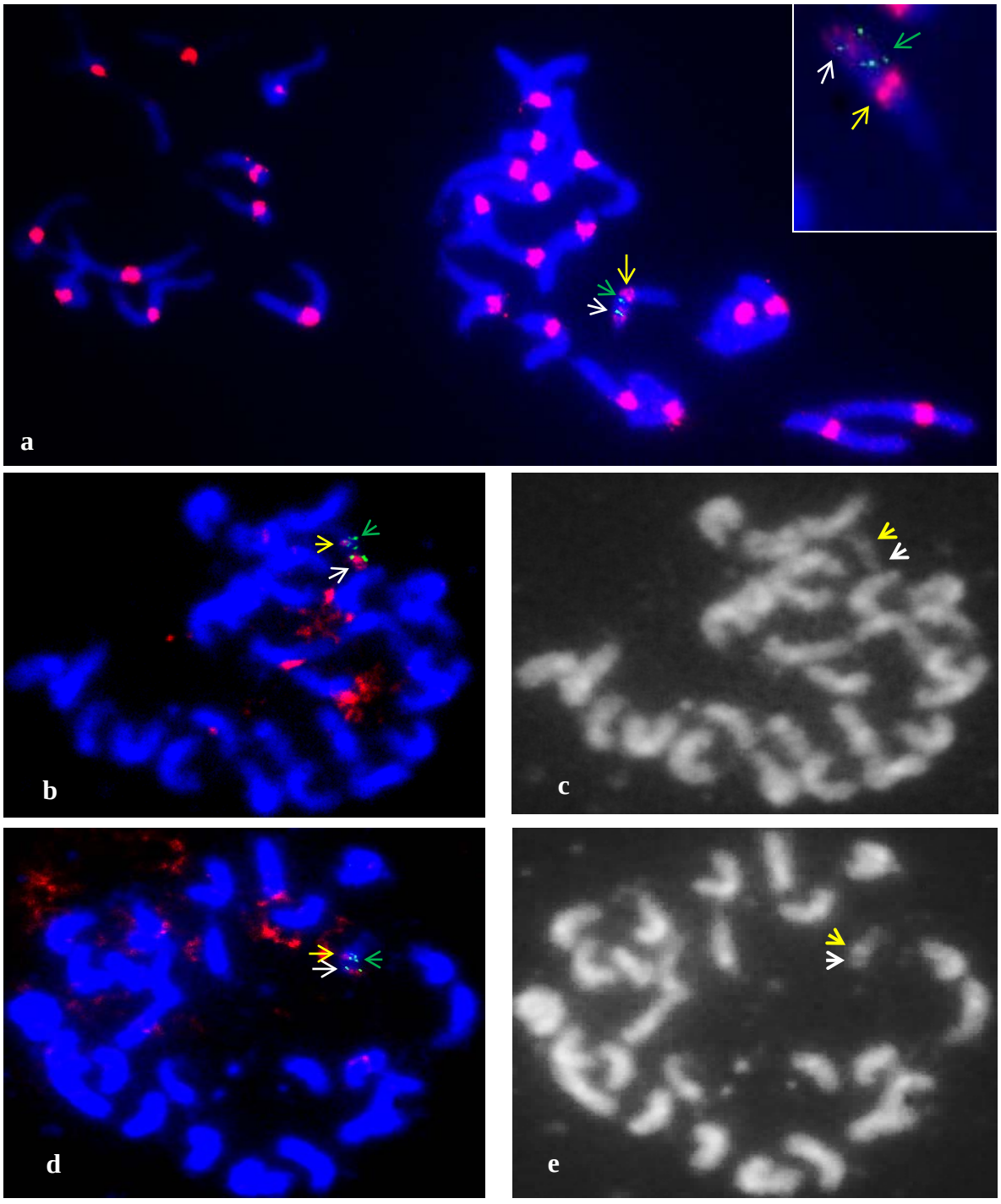

Figure 2.7. Physical localization of the ASGR-carrier chromosome BACs p285J18 and p181G17 on metaphase chromosome spreads from apomictic BC_8 line 58 root tips. Inset image shows the enlarged ASGR-carrier chromosome from a separate spread. In **a-e**, white arrow indicates ASGR

(p800: red, p109/p303: green) signals, yellow arrows indicates centromere (red signals). **a**, green arrow indicates p285J18 signals (green). **b** & **d**, green arrow indicates p181G17 signals (green).

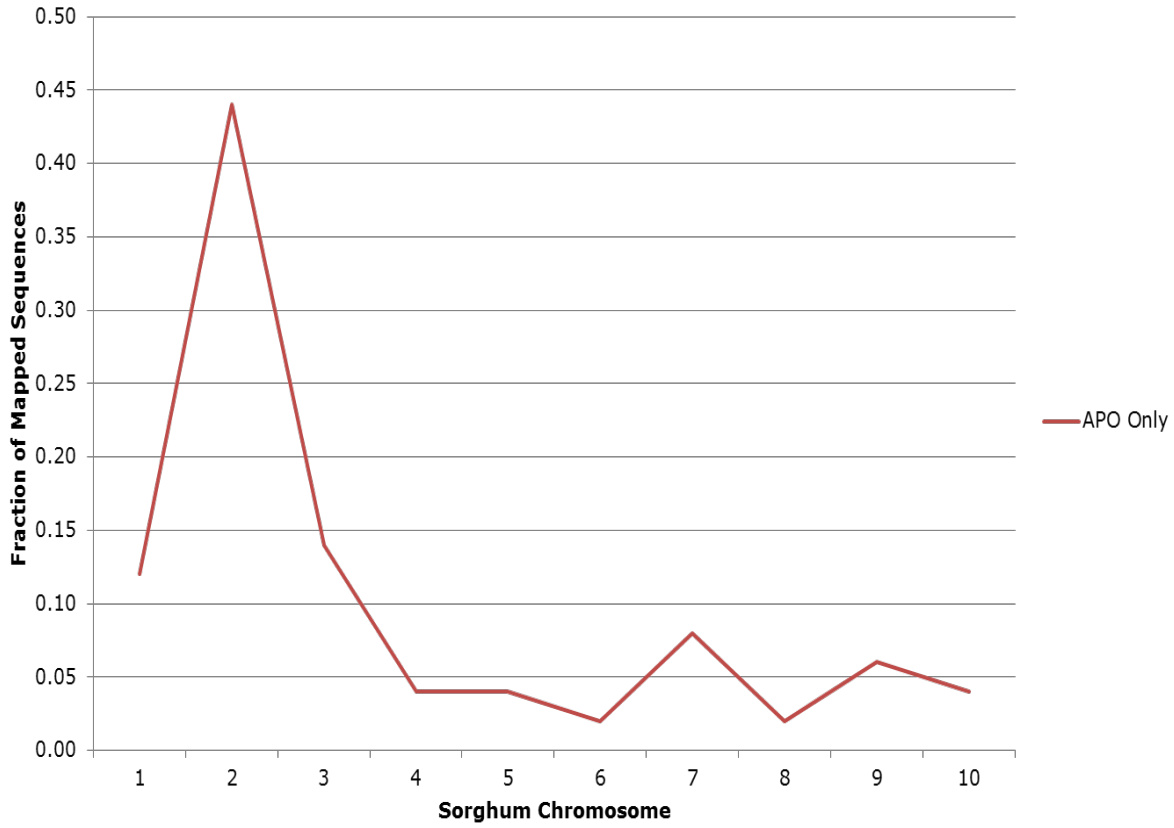

Figure 2.8. Graph showing enrichment of apomictic but not sexual signal on sorghum chromosome 2 during CGH analysis. Data provided by Pioneer.

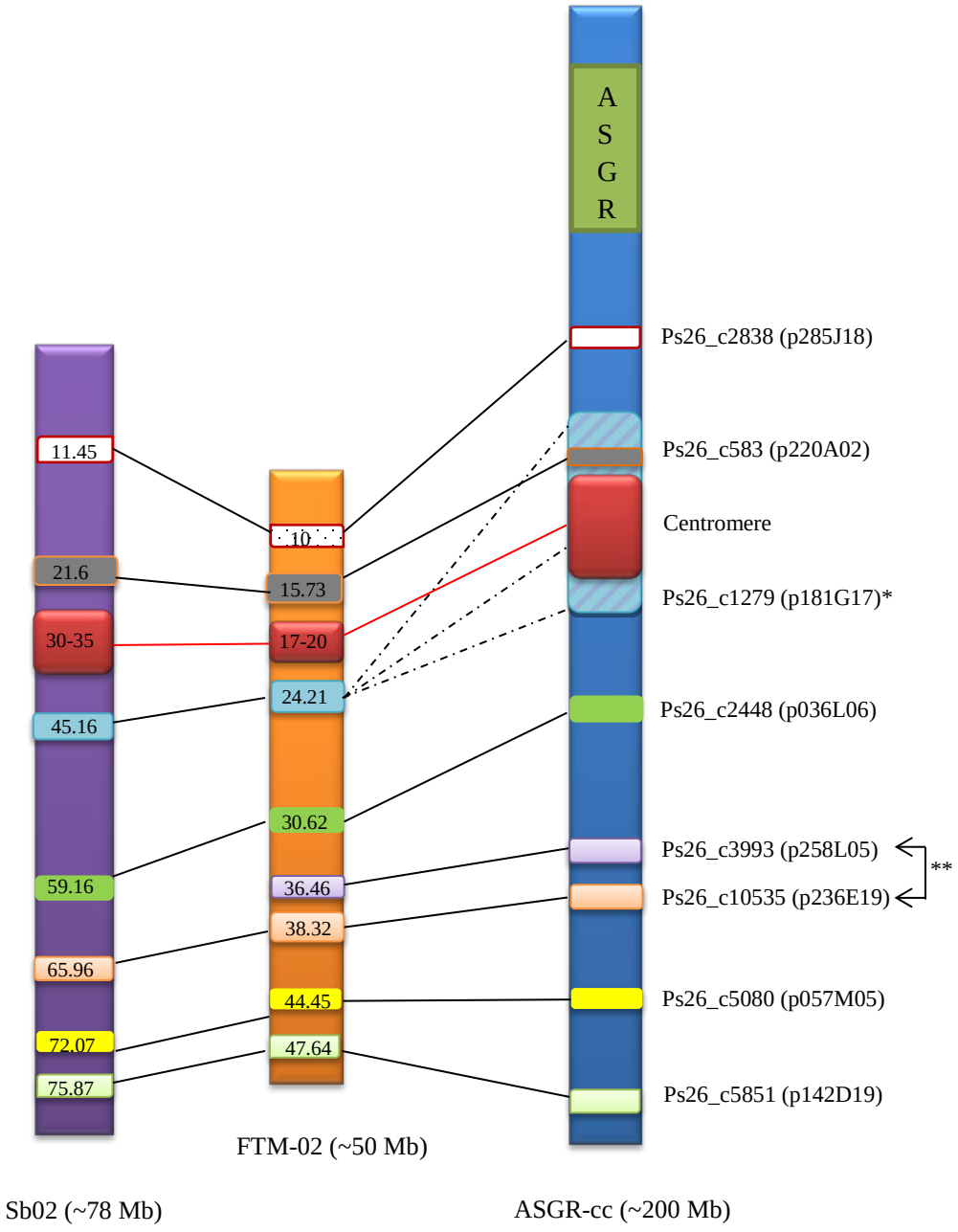

Figure 2.9. Collinear alignments between the ASGR-carrier chromosome and chromosome 2 of sorghum and foxtail millet. Sb02, FTM-02 and ASGR-cc represent sorghum chromosome 2, foxtail millet chromosome 2 and the ASGR-carrier chromosome, respectively. To the right of the ASGR-carrier chromosome are the ASGR-carrier chromosome transcripts and the BACs

carrying the transcripts are in the parentheses at the position they were mapped on the ASGR-carrier chromosome; * represents the BAC with multiple signals in and around the centromere (hatched tab) in the ASGR-carrier chromosome. Black lines connect the genes orthologous in the three chromosomes; red line shows the relative positions of centromere; dashed lines show the regions with BAC p181G17 signals; ** show that the order of the two BACs p258L05 and p236E19 is not determined by mapping but inferred from positions of orthologs in foxtail millet chromosome 2. The values in the tab in the chromosomes represent the location (in Mb) of orthologous genes in respective chromosomes. Orthologous position of transcript Ps26_c2838 in foxtail millet (dotted tab) was determined based on Beijing Genomic Institute database, while Joint Genome Initiative assembly shows orthologous relationship to a gene in chromosome 9 of foxtail millet.

CHAPTER 3

GENETIC MAPPING OF SEQUENCE CHARACTERIZED AMPLIFIED REGION MARKERS IN *PENNISETUM SQUAMULATUM* AND *CENCHRUS CILIARIS* USING DUPLEX SPECIFIC NUCLEASE NORMALIZED ILLUMINA SEQUENCES[1]

[1]Sapkota, S. K, J. A. Conner, and P. Ozias-Akins. To be submitted to *BMC Genomics*

Abstract

Apospory is a form of gametophytic apomixis in which embryo develops from an unreduced embryo sac derived from the nucellar cells of ovule bypassing meiosis and fertilization of gametes. Apospory in *Pennisetum squamulatum* and *Cenchrus ciliaris* is controlled by a physically large, hemizygous, heterochromatic chromosomal block called the apospory-specific genomic region (ASGR), which is abundant in Opie-2 like retrotransposons. Sequence analysis of the *ASGR-BBM-like* gene along with FISH analysis of low copy BAC containing this gene has suggested a relatively recent origin of the ASGR. Duplex specific nuclease (DSN) was used to normalize the libraries of genomic DNA from apomictic and sexual backcross 8 (BC_8) line 58, and apomictic and sexual F_1s from the cross between *P. squamulatum* and *P. glaucum*. A total of 39 out of 44 sequence characterized amplified region (SCAR) markers from potential ASGR-specific contigs mapped to the ASGR in a *Pennisetum* F_1 mapping population. In a limited study eighteen SCARs mapping to ASGR in *P. squamulatum* also showed apomictic specific amplification in *Cenchrus ciliaris*. Four SCAR markers mapped to the A8 plant, a *C. ciliaris* recombinant plant that is capable of producing unreduced aposporous embryo sacs but has lost the ability to undergo parthenogenesis. Genetic mapping of SCARs from putative exonic regions of the ASGR-specific contigs with significant homology to proteins did not display specificity to the ASGR. By successfully mapping more than 85% of the attempted SCAR markers to the ASGR, we show that DSN normalization and Illumina sequencing can be used as an effective strategy for targeted mapping of a physically large locus rich in repetitive sequences, like that of the ASGR.

Introduction

Apomixis is an asexual mode of reproduction in which viable embryos are produced from unreduced egg cells of the ovule without fertilization of gametes (NOGLER 1984). Endosperm of apomictic plants can be derived either autonomously or through fertilization of polar nuclei (OZIAS-AKINS 2006). Because apomictic plants produce seeds that are genetically identical to the maternal plant, apomixis holds as a promising mechanism to fix heterosis in major crops (SAVIDAN 2000; OZIAS-AKINS 2006; HAND and KOLTUNOW 2014). Although it has been documented in more than 120 angiosperm genera, apomixis is rarely found in major crop species (CARMAN 1997; HAND and KOLTUNOW 2014). Several efforts have been made to transfer the apomictic trait to major grains like maize (LEBLANC *et al.* 1995) and pearl millet (DUJARDIN and HANNA 1989b) by interspecific hybridization. These efforts have not yet resulted in commercially viable germplasm (SAVIDAN 2000; OZIAS-AKINS and VAN DIJK 2007; HAND and KOLTUNOW 2014). Alternatively, the transfer of one or more well-characterized apomictic genes to sexual plants could confer the trait of apomixis to this plant and the utilization of molecular tools to identify molecular mechanisms and genes governing components of apomixis can enable application of apomixis to crop breeding (KOLTUNOW *et al.* 1995; GRIMANELLI *et al.* 2001; OZIAS-AKINS 2006).

Apomixis is often associated with polyploidy (NOGLER 1984; VOIGT-ZIELINSKI *et al.* 2012) and is thought of as a deregulation of the sexual pathway by mechanism comprising genetic and/or epigenetic components (CARMAN 1997; GROSSNIKLAUS 2001). Apomixis can be classified as sporophytic or gametophytic. Gametophytic apomixis is subdivided into diplospory and apospory. In diplospory, the embryo sac forms from the megaspore mother cell which does not undergo meiosis but undergoes mitotic divisions to form a non-reduced embryo sac. In

apospory, the non-reduced embryo sac develops from mitotic divisions of aposporous initial cells which are formed from one or more nucellar cells of the ovule (OZIAS-AKINS *et al.* 2003; HAND and KOLTUNOW 2014). In *Pennisetum squamulatum*, apospory is controlled by a single genetic locus transmitted as single large (>50 Mb), hemizygous chromosomal block named the apospory-specific genomic region (ASGR) (OZIAS-AKINS *et al.* 1998; GOEL *et al.* 2003; AKIYAMA *et al.* 2004). Whereas in other gametophytic species such as *Taraxacum officinale* (VAN DIJK *et al.* 1999), *Poa pratensis* (ALBERTINI *et al.* 2001), *Erigeron annuus* (NOYES and RIESEBERG 2000) and *Hieracium caespitosum* (CATANACH *et al.* 2006), formation of unreduced embryo sacs and parthenogenesis were confirmed to be controlled by independent loci. Sequence analysis of ASGR-linked and ASGR-recombining BAC clones covering about 2.7 Mb of the DNA, showed that the ASGR has both gene rich and gene poor regions with several genes identified that could be postulated to play role in apomictic development (CONNER *et al.* 2008). Sequence analysis of the *ASGR-BBM-Like* gene in *Pennisetum*/*Cenchrus* species along with FISH analysis of the low copy BAC containing this gene suggests that this region of the ASGR is of relatively recent origin (AKIYAMA *et al.* 2011).The ASGR is also characterized as highly heterochromatic and contains abundant repetitive elements such as an Opie-2 like retrotransposon (AKIYAMA *et al.* 2004; CONNER *et al.* 2008).

Whole genome shotgun sequencing, a powerful approach capable of generating enough sequence data to cover the genome several times over, can be used for assembly of sequence reads into contigs by aligning and orienting reads based on regions of shared identity (SHAGINA *et al.* 2010). The tremendous amount of data from next generation sequencing has opened a platform for investigating deeper into the transcriptomes and genomes of several crop species. However, the presence of repetitive sequences, which constitutes a large proportion of the plant

genome size, makes assembly of shotgun sequenced clones almost impossible (SANMIGUEL and BENNETZEN 1998). Limiting the redundant sequencing of repetitive elements with effective approaches can facilitate sequencing analyses of large and complex genomes. Several approaches have been used for the elimination of repetitive sequences and the enrichment of low copy sequences. In maize, approaches like Methylation Filtration (MF) (PALMER *et al.* 2003), development of hypomethylated partial restriction (HMPR) libraries (EMBERTON *et al.* 2005) and methylation-spanning linker libraries (MSLL) (YUAN *et al.* 2002) which utilize the tendency of repetitive sequences to be hypermethylated in higher plants for eliminating repetitive sequences and enriching the low copy sequences, have been reported. Another approach, high-C_0t analysis, is based on the renaturation kinetics of DNA. When a denatured genomic DNA solution is kept in an environment where it is allowed to renature, the rate of reassociation of a particular sequence is directly proportional to its repetitiveness in the genome. Since low copy sequences renature more slowly than high copy sequences, the double stranded repetitive DNA can be separated from the single-stranded (low copy) DNA by hydroxyapatite chromatography after the sheared and heat denatured DNA re-anneals slowly. This approach has been used in the study of the maize and sorghum genome (PETERSON *et al.* 2002; YUAN *et al.* 2003).

A more recent approach is the use of duplex specific nuclease (DSN) normalization technology, which is also based on hybridization kinetics, like high-C_0t DNA fractionation, but does not involve physical separation of single stranded (ss) DNA and double stranded (ds) DNA. DSN, obtained from the hepato-pancreas of the Kamchatka crab, is a thermostable enzyme that displays a strong preference for cleaving dsDNA and DNA in RNA-DNA hybrid duplexes compared to ssDNA and RNA (SHAGIN *et al.* 2002). Initially this enzyme was used in a cDNA normalization technique for the enrichment of full length cDNA sequences to discover rare

transcripts by the removal of intermediate and highly abundant cDNAs (ZHULIDOV *et al.* 2004; ZHULIDOV *et al.* 2005). The application of DSN normalization technology to whole-genome shotgun sequencing of the human genome which contain a large proportion of repetitive DNA was reported by Shagina *et al.* (2010). The results from the application of DSN in human genomic DNA indicated that DSN normalization can be highly effective for the elimination of evolutionarily young repetitive sequences from genomic DNA prior to sequencing (SHAGINA *et al.* 2010). Unlike most mammalian repeats, the genomes of many higher plants contain a large number of highly conserved repetitive elements that are evolutionarily young and share high sequence identity (SMITH and FLAVELL 1975; SANMIGUEL *et al.* 1998; RAMAKRISHNA *et al.* 2002).

Successful application of DSN normalization of plant genomic libraries has been reported in the lettuce and Arabidopsis genome by Matvienko *et al*. (2013). The study identifies that DSN normalization in larger plant genomes may enable efficient access to the low-copy fraction of genomes for identifying sequence and copy-number variants and also for mapping purposes. Furthermore, libraries for Illumina sequencing with short 300-500 bp fragments are reported to be ideally suited for reducing highly repeated sequences without concomitant loss of neighboring low-copy sequences (MATVIENKO *et al.* 2013). In *P. squamulatum* 12 sequence characterized amplified region (SCAR) and 7 amplified fragment length polymorphism (AFLP) makers have been mapped successfully to the ASGR (OZIAS-AKINS *et al.* 1998; GOEL *et al.* 2006). Using the sequence from a long terminal repeat region of an ASGR-abundant retrotransposon, 45 sequence specific amplified polymorphism (SSAP) markers were genetically mapped to a 9cM region containing other ASGR-linked specific markers. Five SSAP markers were successfully converted into SCAR markers and also shown to be tightly linked to the ASGR (HEQIANG *et al.*

2009). Here we present the use of DSN normalization as a technique for the elimination of highly repetitive sequences from *Pennisetum* sequences and the data produced by the DSN normalized Illumina libraries for targeted mapping to the ASGR. SCAR markers were developed from DSN normalized Illumina sequences for mapping to a segregating F_1 population to identify ASGR specific sequences.

Materials and Methods

Plant materials

Five sexual plants (R5-33, R5-44, R5-59, R5-77, and R5-87) and seven apomictic plants (R5-04, R5-19, R5-41, R5-46, R5-65, R5-99, and R5-102) from a segregating F_1 population (R5-F_1), obtained from a cross between *Pennisetum squamulatum* (Ps26; 2n=56) and *Pennisetum glaucum* (2n=4X=28), were used for nuclei DNA isolation and sequencing. The F_1 population contains apomictic and sexual plants which show recombination outside of the ASGR but suppressed recombination within the ASGR, as shown by the tight clustering of the molecular markers at the ASGR (GOEL *et al.* 2006).

Four apomictic and five sexual plants from line 58 of backcross 8 (BC_8) were used for nuclei DNA isolation and sequencing. The BC_8 line 58 plants were genotyped and confirmed as sexual or apomictic using the ASGR-linked SCAR primer set 787 (5′-ATACATGAGATGAGCAGGCA-3′) and 788 (5′-GCACTACCTACCTATCAAACA-3′). BC_8 line 58 carries only the ASGR-carrier chromosome from Ps26 and is a facultative apomict. Therefore seedlings from BC_8 line 58 plants can be derived via sexual or apomictic reproduction (SINGH *et al.* 2010).

Nuclei DNA isolation

High molecular weight genomic DNA was extracted from the plants by physical homogenization of plant tissues and nuclei isolation. Isolation of genomic DNA from nuclei minimizes chloroplast and mitochondrial DNA contamination, and produces high quality and quantity of genomic DNA (LUTZ *et al.* 2011). Two to five grams of leaf tissue from 5 BC_8 line 58 sexual plants (sample 1) , 4 BC_8 line 58 apomictic plants (sample 2), 5 F_1 sexual plants (sample 3) and 7 F_1 apomictic plants (sample 4) were used in separate nuclei DNA preparations as described in Roche *et al*. (2002) . Briefly, leaf tissue for each sample was ground in liquid nitrogen. A single homogenization/wash buffer (ZHANG *et al.* 1995) containing 0.5% triton X-100 was used throughout the nuclei DNA preparation. The powdered tissue was suspended in homogenization/wash buffer, filtered through miracloth and nylon sieves, and nuclei were collected via centrifugation. The nuclei were washed two times prior to the final centrifugation to collect the nuclei pellet. DNA was extracted from the isolated nuclei and purified using QIAprep® Spin Miniprep Kit (QIAGEN Inc., Valencia, CA, USA). DNA from each purified sample was quantitated with PicoGreen® (Life Technologies, Grand Island, NY, USA). DNA isolated for each sample type was equally pooled based on the picogreen reading to contain 11 to 12.5 µg of DNA. The samples were sent to Pioneer (Johnston, IA, USA) for normalization and sequencing.

DSN normalization of a complex genome

The complexity of BC_8 line 58 and F_1 genomes, which consist of repetitive sequences, would pose difficulties in assembling and analyses of sequenced reads from the whole genome. Therefore, treatment of denatured/renatured DNA with duplex specific nuclease (DSN) was used to normalize DNA samples, by reducing repetitive sequences, prior to sequencing. Normalization

and sequencing was done in collaboration with Pioneer (Johnston, IA). The following protocol for normalization and sequencing was provided by Pioneer.

DNA was digested with restriction enzyme *Bts*CI (GGATG^NN) and purified with MinElute columns (Qiagen). *Bts*CI specific universal adapters were ligated to the digested DNA fragments and purified with Agencourt AMPure XP beads (Beckman Coulter, Brea, CA). The purified DNA solution was denatured and allowed to renature for 8 hours, followed by 25 minute incubation with 2 Units of DSN. The DSN treatment was inactivated and the DNA was purified with Agencourt AMPure XP beads (Beckman Coulter). The normalized single stranded DNA fraction remaining after DSN treatment was amplified by 20 cycles of PCR and purified with Agencourt AMPure XP (Beckman Coulter) beads. The PCR amplified DNA was digested with restriction enzyme *Bts*CI and purified with Agencourt AMPure XP (Beckman Coulter). The DNA was randomly sheared (Covaris Inc., Woburn, MA) and purified with Agencourt AMPure XP beads (Beckman Coulter). Treatment conditions were established for samples 1 and 2, and consisted of untreated control DNA and either a 4 hour or 8 hour renaturation step followed by DSN treatment.

Illumina sequencing

Paired-end Illumina libraries were prepared for untreated control, 4 and 8 hour renatured/DSN treated DNA and sequenced at a read length of 50 bp on the Illumina® HiSeq2000. After analysis of the three treatment conditions, samples 3 and 4 were processed with 8 hour renatured/DSN treated DNA and paired-end Illumina libraries were generated. All 4 samples were then sequenced as paired ends at a read length of 100 bp on the Illumina® HiSeq2000.

Assembly and reference sequence

Adapter and poor quality sequences were removed from all sequencing runs prior to assembly. Velvet *de novo* assembler (ZERBINO 2010) was used to assemble contiguous sequences from the 100 bp paired end dataset of Illumina sequencing for sample 2 (BC_8 line 58 apomictic plants). Five Velvet assemblies using K-mer values of 31, 39, 47, 55 and 63 were created by Pioneer (Johnston, IA). Basic Local Alignment Search Tool (Blast; (ALTSCHUL *et al.* 1990)) analysis of the five assemblies was used to identify unique contigs from each assembly to build a reference sequence assembly. To begin, the 55 k-mer assembly was used as the database for BlastN analysis against the 63 K-mer assembly with a cutoff of e^{-10}. Results were parsed and any contigs from the 63 K-mer assembly with no BlastN hit to the 55 K-mer database were considered unique. The unique 63 K-mer contigs were added to the 55 K-mer base assembly to create a new BC_8 line 58 DSN contig database for BlastN analysis. This process was repeated with 47, 39 and 31 K-mer base assembly, leading to an *A777s_Uniques.fa* file.

Sequence alignment

The file *A777s_Uniques.fa* was used as the reference sequence to align short reads from the sequencing libraries of apomictic and sexual BC_8 and F_1 samples using the Bowtie version 1.1.0 program on the *zcluster* at the Georgia Advanced Computing Resource Center (GACRC) at the University of Georgia (http://www.gacrc.uga.edu). Bowtie uses an indexing strategy to create an ultrafast, memory-efficient short read aligner to align short reads to reference sequences or contigs (LANGMEAD *et al.* 2009).

The *A777s_Uniques.fa* file was used to build the base/reference for read alignment using the bowtie –build command.

time /usr/local/bowtie/latest/bin/bowtie -build –f A777s_Uniques.fa A777U_base

The short reads from each sequence libraries were aligned to *A777U_base* using Bowtie with the following script parameters *bowtie –v0 –a –best –strata –f A777U_base –S 'filename'.sam*

The *–v0* parameter allow no mismatches during alignment between the individual sequences and the base contigs, *-a* parameter reports all the alignments per read, *-best* asks Bowtie to report the best alignment in terms of number of mismatches, *-strata* keeps the hits in sub-optimal strata from being reported, *-f* is for query input file format as fasta (.fa/.mfa), and *–S* is to write the hits in SAM format.

The SAM files were converted to BAM format using the following samtools script:

samtools view -bS 'filename'.sam > 'filename'.bam

The BAM files were sorted using the following samtools script:

samtools sort 'filename'.bam > 'filename'_sorted

The sorted files were indexed using the following script:

samtools index 'filename'_sorted.bam

The sorted BAM files were converted to tab delimited text files with reported number of read counts using the following samtools script:

samtools idxstats 'filename'_sorted.bam > 'filename'_sorted.txt

All scripts were compiled as a shell script file (.sh). The job queries were submitted to the *rcc-30d* node of the *zcluster* at GACRC using the following script:

qsub –q rcc-30d 'filename'.sh

Screening of the Bowtie alignment and contigs

Since the BC_8 line 58 apomictic plants have the ASGR-carrier chromosome which is lacking in the BC_8 line 58 sexual plants (SINGH *et al.* 2010), contigs from *A777s_Uniques.fa* with

reads aligning to the BC_8 line 58 apomictic but not BC_8 line 58 sexual should be enriched for sequences derived from ASGR-carrier chromosome. Therefore the results from the Bowtie alignment were first sorted based on contigs having at least one hundred BC_8 line 58 apomictic hits and zero BC_8 line 58 sexual hits. Among those contigs, the contigs with Bowtie reads aligning to F_1 apomictic but not to the F_1 sexual sequences were considered to be potential ASGR-derived contigs.

A BlastX search of the potential ASGR-derived contigs was done against the NCBI (National Center for Biotechnology Information; http://www.ncbi.nlm.nih.gov) non-redundant protein sequence (nr) database with a cutoff of e^{-5} to screen the contigs for repetitive elements/proteins, organelle proteins and non-plant hits. Contigs with low copy protein or no protein hits were of high interest, while contigs with similarity to repetitive elements, transposable elements, non-plant proteins and organelle proteins were avoided. BlastN search against the sorghum reference genome sequences (refseq_genomic) in NCBI with a cutoff of e^{-20} was used to identify homology of contigs to sequences/regions in sorghum. The potential ASGR-derived contigs were also screened against the *A777s_Uniques.fa* database using BlastN, and contigs with nearly 100% identity to multiple contigs were considered to be repetitive and therefore avoided. Contigs were also selected to vary the number of F_1 apomictic read counts from 36 to 10,592 sequencing reads. Thirty five aligned F_1 apomictic reads was chosen as a 'base' low number for screening as a contig containing a gene known to reside at the ASGR had 37 F_1 apomictic hits.

Primers design

The contigs with BC_8 line 58 apomictic and F_1 apomictic Bowtie hits but with no BC_8 line 58 sexual and F_1 sexual Bowtie hits; and with F_1 apomictic reads covering at least 50% of

the contig length were selected first to design primers for mapping. Some primers for mapping were designed from contigs that had a significant low copy protein hit in BlastX even if they did not have 50% contig length covered and/or had F_1 apomictic read counts lower than 37. Primers were designed using Geneious (http://www.geneious.com) or primer3 (http://bioinfo.ut.ee/primer3-0.4.0/primer3). Primer information is listed in Table 3.1.

Mapping population used in primer screening

P. glaucum (IA4X, an induced sexual tetraploid millet), sexual *Cenchrus ciliaris* genotype B-2s, apomictic *C. ciliaris* genotype B-12-9, recombinant *C. ciliaris* A8 (CONNER *et al.* 2013), Ps26, segregating population of R5-F_1s (17 apomictic and 22 sexual) and BC_8 line 58 (3 sexual and 3 apomictic) were used for mapping the sequence characterized amplified region (SCAR) primers to the ASGR and ASGR-carrier chromosome. DNA extraction, when needed, was done from approximately 100 mg of young leaf tissue using DNeasy® Plant mini kit (Qiagen). All DNA used in the experiment were genotyped using ASGR-SCAR 787/788 and 792/793 (CONNER *et al.* 2013).

PCR conditions and screening method

PCR was done using 0.375 Units of *Jumpstart™ Taq* DNA polymerase (Sigma-Aldrich Co. LLC, St. Louis, MO, USA) with 0.25 mM dNTP and 0.25 µM primers, or 0.5 units of *TaKaRa™ Ex Taq* DNA polymerase Hot Start Version (Clontech Laboratories, Mountain View, CA, USA) with 0.2 mM dNTP and 0.5 µM primers, in a total reaction volume of 20 µl with 1X PCR buffer (Sigma or Clontech), 25-50 ng template DNA and amplified for 30-35 cycles (Table 3.6). The amplified PCR products were analyzed through a 1.5% (w/v) agarose gel using electrophoresis.

The screening was done in two steps. The SCAR primers were first pre-screened with DNA from 4 F_1 apomictic (R5-04, R5-41, R5-46, R5-65), 4 F_1 sexual (R5-33, R5-59, R5-77, R5-87), 3 BC_8 line 58 apomictic, 3 BC_8 line 58 sexual, Ps26, and IA4X. The primers from pre-screening which only amplified apomictic DNA were screened on a larger F_1 population which consisted of 17 F_1 apomictic (R5-15, R5-19, R5-29, R5-32, R5-37, R5-47, R5-48, R5-51, R5-67, R5-69, R5-74, R5-78, R5-94, R5-04, R5-41, R5-46, and R5-65), 22 F_1 sexual (R5-26, R5-27, R5-28, R5-34, R5-36, R5-38, R5-44, R5-45, R5-54, R5-66, R5-72, R5-73, R5-90, R5-92, R5-93, R5-95, R5-97, R5-101, R5-33, R5-59, R5-77, and R5-87), BC_8 line 58 apomictic, BC_8 line 58 sexual, Ps26, B-12-9, B-2s, A8 and IA4X.

Results

DSN treatment and sequencing

50 bp paired-end reads of BC_8 line 58 apomictic and sexual DNA for untreated control, 4 hours and 8 hours normalized samples were sequenced. K-mers 31bp in length of the paired-end reads from all three treatments of apomictic and sexual BC_8 line 58 were analyzed for count number and abundance by graphing the results. The graph shows that the DSN treatment enriched the abundance of the low copy number 31 K-mer sequences. At lower abundance both DSN treated samples (4 hours and 8 hours) had higher counts of 31 K-mers than the untreated control sample (Figure 3.1 and 3.2). Differences between the samples renatured for 4 hours and 8 hours were slight. Given this data, the remaining samples (3 and 4) were allowed to renature for 8 hours before DSN treatment and library preparation. All 4 sample libraries were sequenced as 100 bp paired-end reads. The number of reads generated from sequencing of samples 1, 2, 3, and 4 are listed in Table 3.2.

Assembly and reference sequences/contigs

The sequence reads from the BC_8 line 58 apomictic samples were assembled using K-mer lengths of 31, 39, 47, 55, and 63 using the Velvet assembler. Velvet used approximately 30% of total library reads to assemble contigs that are a minimum 400 bp and on an average about 750 bp in length. Detailed information on the assembly is listed in Table 3.3. The assemblies created using different K-mer length produced different number and size of contigs and each assembly had different unique contigs in them. Unique contigs were identified through BlastN of query assemblies against the database assembly with cutoff of e^{-10}. BlastN of the 63 K-mer length assembly against 55 K-mer length assembly identified 12,464 unique contigs from the 63 K-mer assembly not present within the 55 K-mer assembly. The unique 63 K-mer contigs were added to 55 K-mer assembly to make new database for BlastN with the 31 K-mer assembly as query. At each step, the unique contigs identified from BlastN of an assembly were added on to form new database for BlastN to next query assembly. A total of 9938, 1191 and 667 unique contigs were identified from the 31, 39 and 47 K-mer assemblies, respectively, and added on step by step to form the new database. Consequently, *A777s_Uniques.fa* file was created with a total of 135,324 contigs after addition of unique contigs respectively from 63, 31, 39 and 47 K-mer assemblies. The *A777s_Uniques.fa* sequence contigs were searched against the maize database in "RepeatMasker" (http://repeatmasker.org) by using 'wublast'. A total of 8% of the bases from the *A777s_Uniques.fa* file were masked with the repeats from maize database. Among 8% masked bases, 6.2% were retroelements comprising of 3.27% Gypsy/DIRS1 like long terminal repeats (LTR), 1.7% Typ1/Copia like LTR elements and 1.2% long interspersed elements (LINEs). Approximately 1% of the bases consisted of DNA transposons, and remaining ~1% were simple repeats and low complexity DNA.

Screening of contigs

Results from the Bowtie alignment were analyzed and sorted for potential ASGR-carrier chromosome derived contigs. 2,660 contigs with at least 100 reads aligning to the BC_8 line 58 apomictic sequences but no reads aligning to the BC_8 line 58 sexual sequences were recorded to a separate file. The 2,660 contigs were used for BlastN analysis against the ASGR-linked BACs and ASGR-recombinant BAC sequences (CONNER *et al.* 2008) to identify contigs with similarity to the ASGR and ASGR-recombinant BAC sequences contigs. A total of 322 hits resulted from the BlastN with an e-value cutoff of e^{-20}. When the hits were parsed for ≥ 99% sequence identity, 101 contigs were identified to have almost complete identity to different ASGR-BACs and ASGR-recombinant BAC sequences (Table 3.5). 7 contigs had similarity to ASGR-BACs (p800, p801 and p900) which gave high copy fluorescence *in situ* hybridization (FISH) signal, 5 contigs had similarity to the ASGR-recombinant BAC (p1300) sequences, while the remaining contigs hit the ASGR-BACs with low copy FISH signals found to be distributed throughout the ASGR low copy regions.

BlastX against the NCBI nr database with the 2,660 potential ASGR-carrier chromosome contigs identified 297 contigs with significant similarity to repetitive elements, non-plant proteins or organelle proteins. A contig containing a portion of the *PsASGR-BBMLike* gene (CONNER *et al.* 2008) was identified through BlastX to the NCBI nr database. Based on a total number of 37 F_1 apomictic sequence alignments but no F_1 sexual read aligning to the *PsASGR-BBML* gene sequence, contigs with at least 35 F_1 apomictic alignments and no F_1 sexual alignments were considered as a base for screening the potential ASGR-derived contigs for mapping. 243 of the 2,660 contigs had at least 35 F_1 apomictic reads aligned with no reported alignments to F_1 sexual. However, two contigs with 23 and 29 F_1 apomictic reads alignment and

covering 170 to 180 bp (9% and 38%, respectively) of the contig length were also chosen for screening based on similarity to low copy protein. Results from BlastN of the 245 selected contigs against the sorghum reference genomic sequences was checked and contigs with multiple and repetitive hits to the sorghum genome were avoided. A total of 44 contigs were selected for primer design and mapping (Table 3.4) with one set of primers.

PCR amplification and mapping

Primers were designed from 44 contigs to test specificity of these Sequence Characterized Amplified Regions (SCARs) markers to the ASGR-carrier chromosome and ASGR. The specificity of the SCAR primers was tested by PCR amplification using either *Jumpstart Taq* or *TaKaRa Ex Taq* HS DNA polymerase.

Pre-screening showed amplification of 39 SCAR primers specific to the ASGR (Figure 3.3a), 4 SCAR primers specific to ASGR-carrier chromosome but not to the ASGR (Figure 3.3b), and one SCAR primer (2891/92) amplifying all tested DNA including sexual BC_8 line 58, IA4X, and sexual F_1 DNA (Figure 3.3c). The ASGR-specific SCAR markers identified from pre-screening were mapped to a larger F_1 population by 32 cycles of PCR amplification to confirm their location. All 39 markers showed ASGR specific amplification in the larger F_1 population (Figure 3.4). 3 SCAR primers 2897/2898, 2931/2932 and 3036/3037 specific to the apomictic F_1s did not amplify any *C. ciliaris* genotypes (Figure 3.4d). 18 SCAR primers specific to apomictic F_1s showed amplification in all *C. ciliaris* DNA regardless of genotype (Figure 3.4a). 4 SCAR primers amplifying the *C. ciliaris* apomictic genotype B-12-9 also showed amplification in the recombinant *C. ciliaris* genotype A8 but not in sexual *C. ciliaris* genotype B-2s (Figure 3.4b). 14 SCAR primers specific to F_1 apomicts were also specific to apomictic genotype B-12-9

and didn't amplify sexual B-2s or recombinant A8 (Figure 3.4c). The results of mapping of the SCAR primers with PCR conditions are summarized in Table 3.6.

Additional analysis of contigs of interest

Contig *A7770001.min7_55_Ctg112355* showed similarity to the 3' UTR of a *histone-lysine N-methyltransferase SUVH2-like* protein from *Setaria italica*. As the first set of designed SCAR primers for this contig, 3050/3051, were ASGR-carrier chromosome specific (Figure 3.3b), an additional set of SCAR primers 3190/3191 were designed. These primers also showed only ASGR-carrier chromosome specificity.

Three other contigs with ASGR-specific SCAR primers and BlastX similarity to a *xylosyltransferase 1-like* gene, a *G-type lectin S-receptor-like serine/threonine-protein kinase RLK1-like* gene and a transcriptional corepressor *LEUNIG-like* gene were investigated for possible gene expression analysis (Table 3.7). As the ASGR-specific SCAR primers for these contigs did not cover the predicted transcriptional regions of these contigs, two additional sets of primers were designed from each. SCAR primers 3192/3193 and 3194/3195 from contig *A7770001.min7_55_Ctg76235 (xylosyltransferase 1-like)* and SCAR primers 3196/3197 and 3198/3199 from contig *A7770001.min7_55_Ctg77495 (G-type lectin S-receptor-like serine/threonine-protein kinase RLK1-like)* were designed to amplify from the last predicted exon region into the predicted 3' UTR. SCAR primers 3188/3189 were designed from contig *A7770001.min7_55_Ctg232696* (*LEUNIG-like*) to amplify a predicted exonic region of the gene. Pre-screening showed that none of the five newly designed 'genic region' SCAR primers specifically amplified the ASGR region. SCAR primers derived from the *G-type lectin S-receptor-like serine/threonine-protein kinase RLK1-like* contig amplified apomictic BC_8 line 58, apomictic and sexual F_1 but did not amplify any BC_8 line 58 sexual and IA4X DNA. SCAR

primers derived from the *xylosyltransferase 1-like* contig amplified all tested apomictic and sexual DNA (Table 3.5). Pre-screening of the SCAR primers derived from the *LEUNIG-like* contig showed amplification from apomictic F_1, BC_8 line 58 apomictic, Ps26 DNA, and one out of four F_1 sexual but not in BC_8 line 58 sexual or IA4X DNA. When mapped to the larger F_1 population *LEUNIG-like* contig SCAR primers 3188/3189 amplified the entire F_1 apomictic population, Ps26, BC_8 line 58 apomictic, B-12-9, A8 and 9 of 22 sexual F_1s (Figure 3.5a). The SCAR primer 2915 from the predicted intron sequences and SCAR primer 3189 from the predicted exon region of the *LEUNIG-like* contig were used as primer sets for PCR amplification. Amplification from this SCAR primer set was found to be specific to the F_1 apomictic DNA (Figure 3.5b).

Discussion

The ASGR and rationale of DSN treatment

The ASGR is characterized as a complex, hemizygous, single genetic locus that controls apospory in *Pennisetum* and *Cenchrus* species as shown by tight clustering of molecular marker (OZIAS-AKINS *et al.* 1998; ROCHE *et al.* 1999; GOEL *et al.* 2006). The ASGR is rich in several repetitive elements such as an Opie-2 like retrotransposons (AKIYAMA *et al.* 2004; CONNER *et al.* 2008). The ability to position a phenotype of interest with respect to molecular markers based on recombination distances is crucial for map-based cloning. However, as recombination at the ASGR (as with many other apomictic loci) is heavily suppressed, map based cloning is not a feasible strategy for identification and isolation of apomictic genes from the locus (OZIAS-AKINS *et al.* 1998; KOTANI *et al.* 2014). Saturation of the ASGR with molecular markers and physical isolation of the BACs linked to these markers can be a potential strategy for the isolation of apomictic genes from the ASGR (HEQIANG *et al.* 2009).

Assembly of shotgun sequences of whole plant genomes is hindered by repetitive elements (SANMIGUEL and BENNETZEN 1998; SHAGINA *et al.* 2010). DSN has been used as a successful strategy for normalization of complex genomes by eliminating repetitive elements (SHAGINA *et al.* 2010; MATVIENKO *et al.* 2013). Evolutionarily young repetitive elements that contain low complexity repeats are depleted more efficiently by DSN than ancient and diverged repeats (MATVIENKO *et al.* 2013). Phylogenetic analysis of the *ASGR-BBM-like* gene in *Pennisetum*/*Cenchrus* along with FISH analysis of the low copy BAC containing this gene have suggested a relatively recent origin of the ASGR in these species (AKIYAMA *et al.* 2011). Using the reassociation kinetics of DNA during C_0t fractionation, Wimpee and Rawson (1979) estimated that the pearl millet (*Pennisetum americanum*) genome consists of 69% repeated DNA, and much of these repeats were present in long tandem arrays consisting of shorter elements varying slightly from one another in their nucleotide sequences. The high degree of sequence homogeneity in repeated DNA of millet implies that these sequences are either highly conserved or that they are relatively recent and have undergone little evolutionary divergence (WIMPEE and RAWSON 1979). This observation suggests repeated DNA in the pearl millet genome should be an efficient target for DSN normalization from both the genome and the ASGR of apomictic samples, and through the comparison of sequences generated from DSN normalized Illumina libraries of apomictic and sexual plants, and effectively target contigs mapping to the ASGR.

DSN normalization and enrichment of low copy sequences

We attempted to use DSN as a normalization strategy to eliminate the repetitive sequences from the genome of BC_8 line 58 (SINGH *et al.* 2010) and F_1 progeny from a cross of *P. squamulatum* and *P. glaucum* (GOEL *et al.* 2006). Analysis of the count number and abundance

of 31 K-mer sequences of the 4 and 8 hours DSN treated BC_8 line 58 sexual and apomictic DNA showed enrichment of low abundance 31 K-mers, compared to the untreated control (Figure 3.1 and 3.2). This result is in agreement with the result seen in DSN treatment of the lettuce genome where sequences with up to 43 copies were enriched rather than depleted in the normalized libraries (MATVIENKO *et al.* 2013). Although DSN treated samples were shown to be enrichment for low abundance 31 K-mer sequences, the level at which repeats were eliminated by DSN treatment of the sampled genomes could not be assessed or quantified with bioinformatics tools as the sequence data of untreated DNA was not provided to us by Pioneer. Masking of repeats in the contigs to the "RepeatMasker" maize database by 'wublast' showed about 6% bases were masked by LTR and LINE retroelements, and about 1% bases were masked by DNA transposons from all the sequences present in the *A777s_Uniques.fa* file. While repeats unique to the *Pennisetum* genomes may not have been identified in the maize database used, the identified repeats represent evolutionarily conserved repeats between maize and *Pennisetum*. Since approximately half of the single copy sequences in pearl millet are between 750 and 1400 bp with another half being 1400 to 8600 bp in length (WIMPEE and RAWSON 1979), at least half of the low copy sequences from the pearl millet genome (BC_8 line 58) are likely to be assembled effectively into a single contig from the enriched low copy sequences of DSN normalized libraries.

Significant BlastN scores of the 2,660 potential ASGR-carrier chromosome derived contigs against the contigs derived from shot-gun cloned ASGR-BAC sequences identified most of the contigs showing similarity to sequences derived from BACs from the "low copy" region of the ASGR rather than the flanking "high copy" region (Table 3.5). Although more BAC sequence was generated from the "low copy" region of the ASGR rather than the "high copy"

region the result does suggest that the assembled contigs have sequences distributed widely along the low copy region of the ASGR and that high copy repeats in the ASGR were mostly removed using the DSN strategy.

Reference assembly and screening of contigs

In the absence of a reference genomic assembly for *Pennisetum*, we used the Velvet assembly of BC_8 line 58 apomictic DSN treated DNA as the reference for Bowtie alignment of the sequenced reads. BlastN of assemblies created using different K-mers against one of the K-mer assembly used as database proved to be a good approach to identify contigs unique to the query assemblies. A total of 24,260 contigs were successfully added to the 111,064 contigs from the 55 K-mer assembly from 63, 31, 39 and 47 K-mer assemblies of BC_8 line 58 apomictic. Since the apomictic BC_8 line 58 has one alien chromosome, the ASGR-carrier chromosome, from the apomictic donor *P. squamulatum* (SINGH *et al.* 2010), it makes a good reference assembly to screen for ASGR and ASGR-carrier chromosome sequences through Bowtie alignment of the reads from apomictic and sexual F_1 and sexual BC_8 line 58. With the objective of finding and mapping low copy sequences to the ASGR, we carefully screened against remaining high copy sequences in the contig assembly by avoiding sequences with homology to repetitive proteins, with homology to multiple chromosomal regions in the sorghum reference genome, with contigs showing multiple hits to nucleotide Blast against the whole assembly, and with repetitive elements from the search against the panicoid database in RepeatMasker (http://repeatmasker.org). BlastN search against the sequences of the ASGR-BACs (CONNER *et al.* 2008) showed that the potential ASGR contig sequences are distributed throughout the ASGR, and that any identified ASGR-specific sequences could be used to isolate BACs at the ASGR.

Mapping the SCARs to the ASGR

The targeted mapping of the ASGR using the DSN normalized Illumina sequences was successful. 44 potential ASGR-specific contig sequences were used to design SCAR primers and the 39 SCARs, shown to be ASGR-specific from pre-screening (Figure 3.3a), successfully mapped specific to apomicts in a F_1 mapping population consisting of 39 individuals (22 sexual and 17 apomicts) from the cross between *P. squamulatum* (2n =56) and *P. glaucum* (Induced tetraploid, 2n=4x=28) (Figure 2.4). None of our mapped SCARs showed any recombination with the ASGR, which is consistent with the previous results that the ASGR is a hemizygous region with suppressed recombination (OZIAS-AKINS *et al.* 1998; GOEL *et al.* 2006). The ASGR is shown to be highly conserved and macrosyntenic between *P. squamulatum* and *C. ciliaris*, however; subtle differences have been observed between the two species indicating local disruptions of microsynteny (GOEL *et al.* 2006). Roche *et al*. (1999) reported 10 of the 12 SCAR markers present in *P. squamulatum* to be present in *C. ciliaris*, and six of them were scored as dominant markers present in aposporous but not in sexual *C. ciliaris*. Similarly in our mapping results, 36 of the 39 SCAR markers specific to the ASGR in *Pennisetum squamulatum* were also present in *C. ciliaris*, and 18 of them were scored as dominant markers present in aposporous but not in sexual *C. ciliaris* (Figure 3.4b and Figure 3.4c). A8, *C. ciliaris* plant showing recombination within the ASGR, has been found to retain the ability to make aposporous embryo sacs but has lost the ability to undergo parthenogenesis. A8 has lost a portion of the ASGR containing SCAR markers UGT197 and Q8M which are located on either side of the "low copy" region of the ASGR. (CONNER *et al.* 2013). 4 SCAR markers in this study that were specific to F_1 apomictic mapped to the aposporous *C. ciliaris* B-12-9 and recombinant A8 but not to the

sexual *C. ciliaris* B-2s. These newly identified A8 retained SCARs markers could be used to identify and isolate more A8 retained BACs and help in the identification of the potential A8 recombination point and apomeiosis candidate genes in the ASGR (CONNER *et al.* 2013).

The suppressed recombination within the ASGR of *P. squamulatum* and *C. ciliaris* has been extensively discussed. Only two markers have been identified to recombine with the ASGR in *P. squamulatum*. RFLP marker UGT204 maps at a considerable genetic distance (about 24 cM) from the ASGR (OZIAS-AKINS *et al.* 1998). AFLP marker PQ355 was mapped at a genetic distance of 2 cM and physically mapped outside the "high copy" region of the ASGR and proximal to the centromere through fluorescent *in situ* hybridization (GOEL *et al.* 2006). In the previous chapter we have shown collinearity of the ASGR-carrier chromosome to the sorghum chromosome 2 outside of the ASGR. Based on this observation, sequences with high similarity to sorghum chromosome 2 are potentially located outside of the ASGR. BACs linked to the markers from these sequences can be isolation and physically mapped to the ASGR-carrier chromosome to identify their position relative to the ASGR. Mapping of those SCARs that are linked to ASGR but potentially outside ASGR, to an even larger F1 mapping population might be successful in identifying potential recombinants.

Mapping SCARs for gene expression analysis

The ASGR-linked contig *A7770001.min7_55_Ctg77495* showed similarity to the last exon and 3' UTR of the *G-type lectin S-receptor-like serine/threonine-protein kinase RLK1-like* gene. SCAR primers designed to amplify the predicted exon/exon or exon/3'UTR region from the contig lost specificity to the ASGR but remained specific to the ASGR-carrier chromosome. The gene could be specific to the ASGR but it is certain that multiple copies of the gene are found within the ASGR-carrier chromosome. The ASGR-linked contig

A7770001.min7_55_Ctg76235, showing similarity to a *xylosyltransferase 1-like* gene, lost specificity to the ASGR-carrier chromosome when primers were designed from the predicted exon/exon or exon/3'UTR region from the contig, leading to amplification of all sexual DNA tested. Sequences in ASGR-linked contig *A7770001.min7_55_Ctg232696* showed similarity to a transcriptional corepressor *LEUNIG-like* gene. *LEUNIG* has been shown to be involved in regulation of Arabidopsis floral homeotic gene *AGAMOUS* (CONNER and LIU 2000). As transcription repression is known to play a key regulatory role in cell fate specification, hormone signaling and plant stress responses, this ASGR-linked contig was of particular interest. Mapping of the SCAR primers from the predicted exon/3'UTR region showed amplification of the F_1 apomictic DNA, a few F_1 sexuals, but without any amplification from BC_8 line 58 sexual or IA4X DNA. Specificity of SCAR primers to the ASGR was regained by using the SCAR primer from the predicted intron region paired with a primer from the predicted exon region. The amplification of the F_1 sexual DNA by the SCAR primers designed from the exonic region suggests the presence of another copy of the *LEUNIG* gene or a gene with WD40 repeat domain within the ASGR-carrier chromosome. We couldn't proceed further into expression analysis of these genes because all of our "genic" derived SCAR primers lacked specificity to the ASGR. As the sequence length of the contigs containing these predicted genes are small, the isolation of BACs linked to the ASGR-specific SCARs could be isolated, sequenced, and BAC sequences could be used to compare homology to these genes and identify potential ASGR-specific SCARs for expression analysis of these genes.

In conclusion, this research demonstrates the effectiveness of DSN treatment for enrichment of low copy sequences and the DSN normalized Illumina sequences generated from apomictic and sexual BC_8 and F_1 plants were effective in targeted mapping of SCARs to the

ASGR. Almost 89% of the SCAR primers designed for mapping showed amplification specific to the ASGR. However, our attempt to map the predicted exonic region from the contigs showing homology to proteins resulted in loss of specificity to the ASGR, which might be due to the presence of copies of those genes or motifs outside ASGR. The ASGR SCAR markers identified from this study can be used for isolation of BAC clones, and physically position them in the ASGR-carrier chromosome relative to previously characterized markers. As this technique is found to be effective in targeted mapping to ASGR, potential ASGR-derived contigs that have not been attempted for mapping in this study can be used as resource for mapping to the ASGR in future.

References

AKIYAMA, Y., J. A. CONNER, S. GOEL, D. T. MORISHIGE, J. E. MULLET *et al.*, 2004 High-resolution physical mapping in Pennisetum squamulatum reveals extensive chromosomal heteromorphism of the genomic region associated with apomixis. Plant Physiology **134:** 1733-1741.

AKIYAMA, Y., S. GOEL, J. A. CONNER, W. W. HANNA, H. YAMADA-AKIYAMA *et al.*, 2011 Evolution of the apomixis transmitting chromosome in Pennisetum. BMC Evolutionary Biology **11:** 289-289.

AKIYAMA, Y., P. OZIAS-AKINS and W. W. HANNA, 2005 High-resolution physical mapping reveals that the apospory-specific genomic region (ASGR) in Cenchrus ciliaris is located on a heterochromatic and hemizygous region of a single chromosome. Theoretical and Applied Genetics **111:** 1042-1051.

ALBERTINI, E., G. MARCONI, G. BARCACCIA, L. RAGGI and M. FALCINELLI, 2004 Isolation of candidate genes for apomixis in Poa pratensis L. Plant Molecular Biology **56:** 879-894.

ALBERTINI, E., G. MARCONI, L. REALE, G. BARCACCIA, A. PORCEDDU *et al.*, 2005 SERK and APOSTART. Candidate genes for apomixis in Poa pratensis. Plant Physiology **138:** 2185-2199.

ALBERTINI, E., A. PORCEDDU, F. FERRANTI, L. REALE, G. BARCACCIA *et al.*, 2001 Apospory and parthenogenesis may be uncoupled in Poa pratensis: a cytological investigation. Sexual Plant Reproduction **14:** 213-217.

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS and D. J. LIPMAN, 1990 Basic local alignment search tool. Journal Of Molecular Biology **215:** 403-410.

ANSORGE, W. J., 2009 Next-generation DNA sequencing techniques. New Biotechnology **25:** 195-203.

BARCACCIA, G., and E. ALBERTINI, 2013 Apomixis in plant reproduction: a novel perspective on an old dilemma. Plant Reproduction **26:** 159-179.

BARRELL, P. J., and U. GROSSNIKLAUS, 2005 Confocal microscopy of whole ovules for analysis of reproductive development: the elongate1 mutant affects meiosis II. Plant Journal **43:** 309-320.

BELÓ, A., M. K. BEATTY, D. HONDRED, K. A. FENGLER, B. LI *et al.*, 2010 Allelic genome structural variations in maize detected by array comparative genome hybridization. TAG. Theoretical And Applied Genetics. Theoretische Und Angewandte Genetik **120:** 355-367.

BENNETZEN, J. L., 2007 Patterns in grass genome evolution. Current Opinion in Plant Biology **10:** 176-181.

BENNETZEN, J. L., J. SCHMUTZ, H. WANG, R. PERCIFIELD, J. HAWKINS *et al.*, 2012 Reference genome sequence of the model plant Setaria. Nature Biotechnology **30:** 555-561.

BICKNELL, R. A., 1997 Isolation of a diploid, apomictic plant of Hieracium aurantiacum. Sexual Plant Reproduction **10:** 168-172.

BICKNELL, R. A., and A. M. KOLTUNOW, 2004 Understanding Apomixis: Recent Advances and Remaining Conundrums. Plant Cell **16:** S228-S245.

BOUTILIER, K., R. OFFRINGA, V. K. SHARMA, H. KIEFT, T. OUELLET *et al.*, 2002 Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth, pp. 1737. American Society of Plant Biologists.

CARMAN, J. G., 1997 Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony [electronic resource]. Biological journal of the Linnean Society **61:** 51-94.

CARMAN, J. G., M. JAMISON, E. ELLIOTT, K. K. DWIVEDI and T. N. NAUMOVA, 2011 Apospory appears to accelerate onset of meiosis and sexual embryo sac formation in sorghum ovules. BMC Plant Biology **11:** 9-9.

CATANACH, A. S., S. K. ERASMUSON, E. PODIVINSKY, B. R. JORDAN and R. BICKNELL, 2006 Deletion Mapping of Genetic Regions Associated with Apomixis in Hieracium, pp. 18650. National Academy of Sciences.

CECCATO, L., S. MASIERO, D. SINHA ROY, S. BENCIVENGA, I. ROIG-VILLANOVA *et al.*, 2013 Maternal Control of PIN1 Is Required for Female Gametophyte Development in Arabidopsis. PLoS ONE **8:** 1-7.

CHENG, C.-Y., D. E. MATHEWS, G. E. SCHALLER and J. J. KIEBER, 2013 Cytokinin-dependent specification of the functional megaspore in the Arabidopsis female gametophyte. The Plant Journal: For Cell And Molecular Biology **73:** 929-940.

COMAI, L., A. P. TYAGI, K. WINTER, R. HOLMES-DAVIS, S. H. REYNOLDS *et al.*, 2000 Phenotypic Instability and Rapid Gene Silencing in Newly Formed Arabidopsis Allotetraploids, pp. 1551. American Society of Plant Physiologists.

CONNER, J., and Z. LIU, 2000 LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development. Proceedings of the National Academy of Sciences of the United States of America **97:** 12902-12907.

CONNER, J. A., S. GOEL, G. GUNAWAN, M. M. CORDONNIER-PRATT, V. E. JOHNSON *et al.*, 2008 Sequence analysis of bacterial artificial chromosome clones from the apospory-specific genomic region of Pennisetum and Cenchrus. Plant Physiology **147:** 1396-1411.

CONNER, J. A., G. GUNAWAN and P. OZIAS-AKINS, 2013 Recombination within the apospory specific genomic region leads to the uncoupling of apomixis components in Cenchrus ciliaris. Planta **238:** 51-63.

CORRAL, J. M., H. VOGEL, O. M. ALIYU, G. HENSEL, T. THIEL *et al.*, 2013 - A Conserved Apomixis-Specific Polymorphism Is Correlated with Exclusive Exonuclease Expression in Premeiotic Ovules of Apomictic Boechera Species. **- 163:** - 1672.

D'ERFURTH, I., S. JOLIVET, N. FROGER, O. CATRICE, M. NOVATCHKOVA *et al.*, 2009 Turning Meiosis into Mitosis. PLoS Biology **7:** 1-10.

DE MEEÛS, T., F. PRUGNOLLE and P. AGNEW, 2007 Asexual reproduction: genetics and evolutionary aspects. Cellular And Molecular Life Sciences: CMLS **64:** 1355-1372.

DUBCOVSKY, J., R. WUSIRIKA, P. J. SANMIGUEL, C. S. BUSSO, L. YAN *et al.*, 2001 Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. Plant Physiology **125:** 1342-1353.

DUJARDIN, M., and W. HANNA, 1984 Cytogenetics of double cross hybrids between Pennisetum americanum - P. purpureum amphiploids and P. americanum X Pennisetum squamulatum interspecific hybrids. Theoretical and Applied Genetics **69:** 97-100.

DUJARDIN, M., and W. HANNA, 1986 An apomictic polyhaploid obtained from a pearl millet x Pennisetum squamulatum apomictic interspecific hybrid. Theoretical & Applied Genetics **72:** 33.

DUJARDIN, M., and W. W. HANNA, 1989a Crossability of pearl millet with wild Pennisetum species. Crop Science **29:** 77-80.

DUJARDIN, M., and W. W. HANNA, 1989b Developing apomictic pearl millet - characterization of a BC3 plant. Journal of Genetics & Breeding **43:** 145-149.

EBINA, M., H. NAKAGAWA, T. YAMAMOTO, H. ARAYA, S.-I. TSURUTA *et al.*, 2005 Co-segregation of AFLP and RAPD markers to apospory in Guineagrass (Panicum maximumJacq.). Grassland Science **51:** 71-78.

EDGER, P. P., and J. C. PIRES, 2009 Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes [electronic resource]. Chromosome Research **17:** 699-717.

EMBERTON, J., J. X. MA, Y. N. YUAN, P. SANMIGUEL and J. L. BENNETZEN, 2005 Gene enrichment in maize with hypomethylated partial restriction (HMPR) libraries. Genome Research **15:** 1441-1446.

FEUILLET, C., and B. KELLER, 2002 Comparative Genomics in the Grass Family: Molecular Characterization of Grass Genome Structure and Evolution. Annals of Botany **89:** 3-10.

FRISVOLD, G., K. BICKNELL and R. BICKNELL, 2005 A preliminary analysis of the benefits of introducing apomixis into rice, pp. in *AERU Discussion Paper*. Agribusiness & Economics Research Unit, Lincoln University, Canterbury; New Zealand.

FUKUI, K., 1996 Plant chromosome at mitosis, pp. 1-17 in *Plant chromosomes laboratory methods*, edited by K. FUKUI and S. NAGAYAMA. CRC Press, Boca Raton, FL.

GALE, M. D., and K. M. DEVOS, 1998 Comparative genetics in the grasses. Proceedings of the National Academy of Sciences of the United States of America **95:** 1971-1974.

GARCIA-AGUILAR, M., C. MICHAUD, O. LEBLANC and D. GRIMANELLI, 2010 Inactivation of a DNA Methylation Pathway in Maize Reproductive Organs Results in Apomixis-Like Phenotypes, pp. 3249. American Society of Plant Biologists.

GARCÍA, R., M. J. ASÍNS, J. FORNER and E. A. CARBONELL, 1999 Genetic analysis of apomixis in Citrus and Poncirus by molecular markers. Theoretical and Applied Genetics **99:** 511-518.

GAUT, B. S., 2002 Evolutionary dynamics of grass genomes. New Phytologist **154:** 15-28.

GOEL, S., Z. B. CHEN, Y. AKIYAMA, J. A. CONNER, M. BASU *et al.*, 2006 Comparative physical mapping of the apospory-specific genomic region in two apomictic grasses: Pennisetum squamulatum and Cenchrus ciliaris, pp. 389-400 in *Genetics*. Genetics Society of America, Bethesda; USA.

GOEL, S., C. ZHENBANG, J. A. CONNER, Y. AKIYAMA, W. W. HANNA *et al.*, 2003 Delineation by Fluorescence in Situ Hybridization of a Single Hemizygous Chromosomal Region Associated With Aposporous Embryo Sac Formation in Pennisetum squamulatum and Cenchrus ciliaris. Genetics **163:** 1069.

GRIMANELLI, LEBLANC, ESPINOSA, PEROTTI, GONZÁLEZ DE LEÓN *et al.*, 1998 Non-Mendelian transmission of apomixis in maize–Tripsacum hybrids caused by a transmission ratio distortion. Heredity **80:** 40-47.

GRIMANELLI, D., M. HERNÁNDEZ, E. PEROTTI and Y. SAVIDAN, 1997 Dosage effects in the endosperm of diplosporous apomictic Tripsacum (Poaceae). Sexual Plant Reproduction **10:** 279-282.

GRIMANELLI, D., O. LEBLANC, E. PEROTTI and U. GROSSNIKLAUS, 2001 Developmental genetics of gametophytic apomixis. Trends in Genetics **17:** 597-604.

GROSSNIKLAUS, U., 2001 From sexuality to apomixis: Molecular and genetic approaches., pp. 168-211 in *The Flowering of Apomixis: From Mechanisms to Genetic Engineering*, edited by J. G. C. Y. SAVIDAN, AND T. DRESSELHAUS. CIMMYT, IRD, European Commission DG, Mexico, D. F,.

GROSSNIKLAUS, U., and K. SCHNEITZ, 1998 The molecular and genetic basis of ovule and megagametophyte development. Seminars In Cell & Developmental Biology **9:** 227-238.

GUIMARÃES, L. A., D. M. DE A. DUSI, S. MASIERO, F. RESENTINI, A. C. M. M. GOMES *et al.*, 2013 BbrizAGL6 Is Differentially Expressed During Embryo Sac Formation of Apomictic and Sexual Brachiaria brizantha Plants. Plant molecular biology reporter **31:** 1397-1406.

GUO, M., D. DAVIS and J. A. BIRCHLER, 1996 Dosage effects on gene expression in a maize ploidy series. Genetics **142:** 1349-1355.

HAND, M. L., and A. M. G. KOLTUNOW, 2014 The Genetic Control of Apomixis: Asexual Seed Formation. Genetics **197:** 441-450.

HANNA, W. W., 1995 Use of Apomixis in cultivar development. Advances in agronomy **54:** 333-350.

HANNA, W. W., and E. C. BASHAW, 1987 Apomixis: its identification and use in plant breeding. Crop Science **27:** 1136-1139.

HEQIANG, H., J. A. CONNER and P. OZIAS-AKINS, 2009 Genetic mapping of the apospory-specific genomic region in Pennisetum squamulatum using retrotransposon-based molecular markers. Theoretical & Applied Genetics **119:** 199-212.

HESLOP-HARRISON, J. S., A. BRANDES and T. SCHWARZACHER, 2003 Tandemly repeated DNA sequences and centromeric chromosomal regions of Arabidopsis species. Chromosome Research **11:** 241-253.

HÖRANDL, E., and E. M. TEMSCH, 2009 Introgression of apomixis into sexual species is inhibited by mentor effects and ploidy barriers in the Ranunculus auricomus complex. Annals of Botany **104:** 81-89.

INGHAM, L. D., W. W. HANNA, J. W. BAIER and L. C. HANNAH, 1993 Origin of the main class of repetitive DNA within selected Pennisetum species. Molecular and General Genetics **238:** 350-356.

IWATA, A., A. L. TEK, M. M. S. RICHARD, B. ABERNATHY, A. FONSÊCA *et al.*, 2013 Identification and characterization of functional centromeres of the common bean. Plant Journal **76:** 47-60.

KAMM, A., J. S. HESLOP-HARRISON and T. SCHMIDT, 1994 Molecular and physical organization of highly repetitive, undermethylated DNA from Pennisetum glaucum. Molecular & general genetics : MGG **244:** 420-425.

KANTAMA, L., T. F. SHARBEL, M. E. SCHRANZ, T. MITCHELL-OLDS, S. DE VRIES *et al.*, 2007 Diploid Apomicts of the Boechera holboellii Complex Display Large-Scale Chromosome Substitutions and Aberrant Chromosomes, pp. 14026. National Academy of Sciences.

KAUSHAL, P., D. R. MALAVIYA and A. K. ROY, 2004 Prospects for breeding apomictic rice: a reassessment. Current Science **87:** 292-296.

KINDIGER, B., D. BAI and V. SOKOLOV, 1996 Assignment of a gene(s) conferring apomixis in Tripsacum to a chromosome arm: cytological and molecular evidence. Genome **39:** 1133-1141.

KIROV, I., M. DIVASHUK, K. VAN LAERE, A. SOLOVIEV and L. KHRUSTALEVA, 2014 An easy "SteamDrop" method for high quality plant chromosome preparation. Molecular Cytogenetics (17558166) **7:** 1-20.

KOJIMA, A., and Y. NAGATO, 1997 Discovery of highly apomictic and highly amphimictic dihaploids in Allium tuberosum. Sexual Plant Reproduction **10:** 8-12.

KOLTUNOW, A. M., 1993 Apomixis: Embryo Sacs and Embryos Formed without Meiosis or Fertilization in Ovules, pp. 1425. American Society of Plant Physiologists.

KOLTUNOW, A. M., R. A. BICKNELL and A. M. CHAUDHURY, 1995 Apomixis: Molecular Strategies for the Generation of Genetically Identical Seeds without Fertilization, pp. 1345. American Society of Plant Physiologists.

KOLTUNOW, A. M. G., S. D. JOHNSON, J. C. M. RODRIGUES, T. OKADA, Y. HU *et al.*, 2011 Sexual reproduction is the default mode in apomictic Hieracium subgenus Pilosella, in which two dominant loci function to enable apomixis. The Plant Journal: For Cell And Molecular Biology **66:** 890-902.

KOTANI, Y., S. T. HENDERSON, G. SUZUKI, S. D. JOHNSON, T. OKADA *et al.*, 2014 The LOSS OF APOMEIOSIS (LOA) locus in Hieracium praealtum can function independently of the associated large-scale repetitive chromosomal structure. New Phytologist **201:** 973-981.

LANGMEAD, B., C. TRAPNELL, M. POP and S. L. SALZBERG, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology **10:** R25-R25.

LEBLANC, O., D. GRIMANELLI, D. GONZÁLEZ DE LEÓN and Y. SAVIDAN, 1995 Detection of the apomictic mode of reproduction in maize-Tripsacum hybrids using maize RFLP markers. Theoretical and Applied Genetics **90:** 1198-1203.

LIM, K.-B., T.-J. YANG, Y.-J. HWANG, J. S. KIM, J.-Y. PARK *et al.*, 2007 Characterization of the centromere and peri-centromere retrotransposons in Brassica rapa and their distribution in related Brassica species. The Plant Journal: For Cell And Molecular Biology **49:** 173-183.

LUTZ, K. A., W. WANG, A. ZDEPSKI and T. P. MICHAEL, 2011 Isolation and analysis of high quality nuclear DNA with reduced organellar DNA for plant genome sequencing and resequencing. BMC Biotechnology **11:** 54-54.

MARDIS, E. R., 2008 The impact of next-generation sequencing technology on genetics. Trends in Genetics **24:** 133-141.

MARTINEZ, E. J., J. P. A. ORTIZ, C. L. QUARIN, H. E. HOPP and J. STEIN, 2003 Genetic characterization of apospory in tetraploid Paspalum notatum based on the identification of linked molecular markers. Molecular breeding : new strategies in plant improvement **12:** 319-327.

MARTÍNEZ, E. J., C. L. QUARÍN and M. D. HAYWARD, 1999 Genetic control of apospory in apomictic Paspalum species. Cytologia **64:** 425-433.

MATVIENKO, M., A. KOZIK, L. FROENICKE, D. LAVELLE, B. MARTINEAU *et al.*, 2013 Consequences of Normalizing Transcriptomic and Genomic Libraries of Plant Genomes Using a Duplex-Specific Nuclease and Tetramethylammonium Chloride. PLoS ONE **8:** 1-17.

MATZK, F., A. MEISTER, R. BRUTOVSKÁ and I. SCHUBERT, 2001 Reconstruction of reproductive diversity in Hypericum perforatum L. opens novel strategies to manage apomixis. Plant Journal **26:** 275-282.

MATZK, F., S. PRODANOVIC, H. BÄUMLEIN and I. SCHUBERT, 2005 The inheritance of apomixis in Poa pratensis confirms a five locus model with differences in gene expressivity and penetrance. Plant Cell **17:** 13-24.

MEISTER, G., 2013 Argonaute proteins: functional insights and emerging roles. Nature Reviews Genetics **14:** 447-459.

MORGAN, R. N., P. OZIAS-AKINS and W. W. HANNA, 1998 Seed Set in an Apomictic BC 3 Pearl Millet, pp. 89. The University of Chicago Press.

NAUMOVA, T. N., J. BERGERVOET, K. S. RAMULU, K. BOUTILIER, A. KRAVTCHENKO *et al.*, 2001 Reproductive development in apomictic populations of Arabis holboellii (Brassicaceae) [electronic resource]. Sexual Plant Reproduction **14:** 195-200.

NOGLER, G. A., 1984 Gametophytic apomixis, pp. 474-518 in *Embryology of angiosperms / edited by B.M. Johri*. Berlin : Springer-Verlag, 1984.

NOYES, R. D., and L. H. RIESEBERG, 2000 Two independent loci control agamospermy (apomixis) in the triploid flowering plant Erigeron annuus. Genetics **155:** 379-390.

OGAWA, D., S. JOHNSON, S. HENDERSON and A. KOLTUNOW, 2013 Genetic separation of autonomous endosperm formation (AutE) from the two other components of apomixis in Hieracium. Plant Reproduction **26:** 113-123.

OKADA, T., Y. K. HU, M. R. TUCKER, J. M. TAYLOR, S. D. JOHNSON *et al.*, 2013 - Enlarging Cells Initiating Apomixis in Hieracium praealtum Transition to an Embryo Sac Program prior to Entering Mitosis. - **163:** - 231.

OLMEDO-MONFIL, V., N. DURÁN-FIGUEROA, M. ARTEAGA-VÁZQUEZ, E. DEMESA-ARÉVALO, D. AUTRAN *et al.*, 2010 Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature **464:** 628-632.

OZIAS-AKINS, P., 2006 Apomixis: Developmental Characteristics and Genetics. Critical Reviews in Plant Sciences **25:** 199-214.

OZIAS-AKINS, P., Y. AKIYAMA and W. W. HANNA, 2003 Molecular characterization of the genomic region linked with apomixis in Pennisetum/Cenchrus. Functional & Integrative Genomics **3:** 94-104.

OZIAS-AKINS, P., D. ROCHE and W. W. HANNA, 1998 Tight Clustering and Hemizygosity of Apomixis-Linked Molecular Markers in Pennisetum squamulatum Implies Genetic Control of Apospory by a Divergent Locus that may have no Allelic Form in Sexual Genotypes, pp. 5127. National Academy of Sciences of the United States of America.

OZIAS-AKINS, P., and P. J. VAN DIJK, 2007 Mendelian Genetics of Apomixis in Plants. Annual Review of Genetics **41:** 509-537.

PALMER, L. E., P. D. RABINOWICZ, A. L. O'SHAUGHNESSY, V. S. BALIJA, L. U. NASCIMENTO *et al.*, 2003 Maize genome sequencing by methylation filtration. Science (New York, N.Y.) **302:** 2115-2117.

PATERSON, A. H., J. E. BOWERS, M. D. BUROW, X. DRAYE, C. G. ELSIK *et al.*, 2000 Comparative genomics of plant chromosomes. The Plant Cell **12:** 1523-1540.

PATERSON, A. H., Y.-R. LIN, Z. LI, K. F. SCHERTZ, J. F. DOEBLEY *et al.*, 1995 Convergent Domestication of Cereal Crops by Independent Mutations at Corresponding Genetic Loci, pp. 1714. American Society for the Advancement of Science.

PATERSON, A. H. B. J. E. B. R. I. J. H. G. U. T. A. J. M. H. X. T. A. K. C. J. F. A. U. I. V., 2009 The Sorghum bicolor genome and the diversification of grasses. Nature **457:** 551-556.

PESSINO, S. C., J. P. A. ORITZ, O. LEBLANC, C. B. D. VALLE, C. EVANS *et al.*, 1997 Identification of a maize linkage group related to apomixis in Brachiaria. Theoretical and Applied Genetics **94:** 439-444.

PETERSON, D. G., S. R. SCHULZE, E. B. SCIARA, S. A. LEE, J. E. BOWERS *et al.*, 2002 Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery. Genome Research **12:** 795-807.

PODIO, M., S. FELITTI, L. SIENA, L. DELGADO, M. MANCINI *et al.*, 2014 Characterization and expression analysis of SOMATIC EMBRYOGENESIS RECEPTOR KINASE (SERK) genes in sexual and apomictic Paspalum notatum. Plant Molecular Biology **84:** 479-495.

PUPILLI, F., and G. BARCACCIA, 2012 Cloning plants by seeds: inheritance models and candidate genes to increase fundamental knowledge for engineering apomixis in sexual crops. Journal of Biotechnology **159:** 291-311.

QUARIN, C. L., F. ESPINOZA, E. J. MARTINEZ, S. C. PESSINO and O. A. BOVO, 2001 A rise of ploidy level induces the expression of apomixis in Paspalum notatum. Sexual Plant Reproduction **13:** 243-249.

RAMAKRISHNA, W., J. EMBERTON, M. OGDEN, P. SANMIGUEL and J. L. BENNETZEN, 2002 Structural Analysis of the Maize Rp1 Complex Reveals Numerous Sites and Unexpected Mechanisms of Local Rearrangement. The Plant Cell**:** 3213.

RAVI, M., and S. W. L. CHAN, 2010 Haploid plants produced by centromere-mediated genome elimination. Nature **464:** 615-618.

RAVI, M., M. P. A. MARIMUTHU and I. SIDDIQI, 2008 Gamete formation without meiosis in Arabidopsis. Nature **451:** 1121-1124.

REISER, L., and R. L. FISCHER, 1993 The ovule and the embryo sac. Plant Cell **5:** 1291-1301.

ROCHE, D., P. CONG, Z. CHEN, W. W. HANNA, D. L. GUSTINE *et al.*, 1999 An apospory-specific genomic region is conserved between Buffelgrass (Cenchrus ciliaris L.) and Pennisetum squamulatum Fresen. Plant Journal **19:** 203-208.

ROCHE, D., J. CONNER, M. BUDIMAN, D. FRISCH, R. WING *et al.*, 2002 Construction of BAC libraries from two apomictic grasses to study the microcolinearity of their apospory-specific genomic regions. Theoretical & Applied Genetics **104:** 804-812.

SALSE, J., M. ABROUK, F. MURAT, U. M. QURAISHI and C. FEUILLET, 2009 Improved criteria and comparative genomics tool provide new insights into grass paleogenomics. Briefings in Bioinformatics **10:** 619-630.

SALSE, J., S. BOLOT, M. THROUDE, V. JOUFFE, B. PIEGU *et al.*, 2008 Identification and Characterization of Shared Duplications between Rice and Wheat Provide New Insight into Grass Genome Evolution. The Plant Cell**:** 11.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA Sequencing with Chain-Terminating Inhibitors, pp. 5463. National Academy of Sciences of the United States of America.

SANMIGUEL, P., and J. L. BENNETZEN, 1998 Evidence that a Recent Increase in Maize Genome Size was Caused by the Massive Amplification of Intergene Retrotransposons. Annals of Botany **82:** 37-44.

SANMIGUEL, P., B. S. GAUT, A. TIKHONOV, Y. NAKAJIMA and J. L. BENNETZEN, 1998 The paleontology of intergene retrotransposons of maize. Nature Genetics **20:** 43-45.

SAUNDERS, V. A., and A. HOUBEN, 2001 The pericentromeric heterochromatin of the grass Zingeria biebersteiniana (2n = 4) is composed of Zbcen1-type tandem repeats that are intermingled with accumulated dispersedly organized sequences. Genome **44:** 955-961.

SAVIDAN, Y., 2000 Apomixis: genetics and breeding. Plant breeding reviews **18:** 13-86.

SAVIDAN, Y., J. CARMAN, T. DRESSELHAUS and G. H. B. J. C. C. F. N. T. N. V.-C. J. P. S. R. T. G. D. T. J. G.-D.-L. TOENNIESSEN, 2001 The flowering of apomixis: from mechanisms to genetic engineering, pp. in *The flowering of apomixis: from mechanisms to genetic engineering*, edited by Y. SAVIDAN, J. CARMAN and T. DRESSELHAUS. Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT) (International Maize and Wheat Improvement Center), Mexico City; Mexico.

SCHALLAU, A., F. ARZENTON, A. J. JOHNSTON, U. HÄHNEL, D. KOSZEGI *et al.*, 2010 Identification and genetic analysis of the APOSPORY locus in Hypericum perforatum L. Plant Journal **62:** 773-784.

SHAGIN, D. A., D. V. REBRIKOV, V. B. KOZHEMYAKO, I. M. ALTSHULER, A. S. SHCHEGLOV *et al.*, 2002 A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. Genome Research **12:** 1935-1942.

SHAGINA, I., E. BOGDANOVA, I. Z. MAMEDOV, Y. LEBEDEV, S. LUKYANOV *et al.*, 2010 Normalization of genomic DNA using duplex-specific nuclease. BIOTECHNIQUES **48:** 455-459.

SILVEIRA, R. D., M. M. DO CARMO COSTA, M. Ł. ALVES-FERREIRA, V. T. DE CAMPOS CARNEIRO, N. Ł. F. MARTINS *et al.*, 2012 Expressed sequence-tag analysis of ovaries of Brachiaria brizantha reveals genes associated with the early steps of embryo sac differentiation of apomictic plants [electronic resource]. Plant cell reports **31:** 403-416.

SINGH, M., J. A. CONNER, Y. J. ZENG, W. W. HANNA, V. E. JOHNSON *et al.*, 2010 Characterization of apomictic BC7 and BC8 pearl millet: meiotic chromosome behavior and construction of an ASGR-carrier chromosome-specific library. Crop Science **50:** 892-902.

SINGH, M., S. GOEL, R. B. MEELEY, C. DANTEC, H. PARRINELLO *et al.*, 2011 Production of Viable Gametes without Meiosis in Maize Deficient for an ARGONAUTE Protein, pp. 443. American Society of Plant Biologists.

SMITH, D., and R. FLAVELL, 1975 Characterisation of the wheat genome by renaturation kinetics. Chromosoma **50:** 223.

TAS, I. C. Q., and P. J. VAN DIJK, 1999 Crosses between sexual and apomictic dandelions (Taraxacum). I. The inheritance of apomixis. Heredity **83:** 707-714.

TOPP, C. N., and R. K. DAWE, 2006 Reinterpreting pericentromeric heterochromatin. Current Opinion in Plant Biology **9:** 647-653.

TUCKER, M. R., A.-C. G. ARAUJO, N. A. PAECH, V. HECHT, D. L. S. ED *et al.*, 2003 Sexual and Apomictic Reproduction in Hieracium Subgenus Pilosella Are Closely Interrelated Developmental Pathways, pp. 1524. American Society of Plant Biologists.

TUCKER, M. R., T. OKADA, S. D. JOHNSON, F. TAKAIWA and A. M. G. KOLTUNOW, 2012 Sporophytic ovule tissues modulate the initiation and progression of apomixis in Hieracium. Journal of Experimental Botany **63:** 3229-3241.

VAN DIJK, P. J., I. C. Q. TAS, M. FALQUE and T. BAKX-SCHOTMAN, 1999 Crosses between sexual and apomictic dandelions (Taraxacum). II. The breakdown of apomixis. Heredity **83:** 715-721.

VAN DIJK, P. J., and K. VIJVERBERG, 2005 The significance of apomixis in the evolution of the angiosperms: a reappraisal. REGNUM VEGETABILE **143:** 101-116.

VOIGT-ZIELINSKI, M.-L., M. PIWCZYŃSKI and T. SHARBEL, 2012 Differential effects of polyploidy and diploidy on fitness of apomictic Boechera. Sexual Plant Reproduction **25:** 97-109.

WIMPEE, C. F., and J. R. RAWSON, 1979 Characterization of the nuclear genome of pearl millet. Biochimica Et Biophysica Acta **562:** 192-206.

YU, P., C. WANG, Q. XU, Y. FENG, X. YUAN *et al.*, 2011 Detection of copy number variations in rice using array-based comparative genomic hybridization. BMC Genomics **12:** 372-372.

YUAN, Y., P. J. SANMIGUEL and J. L. BENNETZEN, 2002 Methylation-spanning linker libraries link gene-rich regions and identify epigenetic boundaries in Zea mays. Genome Research **12:** 1345-1349.

YUAN, Y., P. J. SANMIGUEL and J. L. BENNETZEN, 2003 High-Cot sequence analysis of the maize genome. The Plant Journal: For Cell And Molecular Biology **34:** 249-255.

ZENG, Y., J. CONNER and P. OZIAS-AKINS, 2011 Identification of ovule transcripts from the Apospory-Specific Genomic Region (ASGR)-carrier chromosome. BMC Genomics **12:** 206-206.

ZERBINO, D. R., 2010 Using the Velvet de novo assembler for short-read sequencing technologies. Current Protocols In Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [Et Al.] **Chapter 11:** 11.15-11.15.

ZHANG, H. B., X. P. ZHAO, X. L. DING, A. H. PATERSON and R. A. WING, 1995 Preparation of megabase-size DNA from plant nuclei. Plant Journal **7:** 175-184.

ZHONG, X., A. V. KAMMEN, J. HANS DE JONG, P. ZABEL, P. F. FRANSZ *et al.*, 1996 High-resolution mapping on pachytene chromosomes and extended DNA fibres by fluorescence in-situ hybridisation. Plant molecular biology reporter **14:** 232-242.

ZHULIDOV, P. A., E. A. BOGDANOVA, A. S. SHCHEGLOV, I. A. SHAGINA, L. L. VAGNER *et al.*, 2005 [A method for the preparation of normalized cDNA libraries enriched with full-length sequences]. Bioorganicheskaia Khimiia **31:** 186-194.

ZHULIDOV, P. A., E. A. BOGDANOVA, A. S. SHCHEGLOV, L. L. VAGNER, G. L. KHASPEKOV *et al.*, 2004 Simple cDNA normalization using kamchatka crab duplex-specific nuclease. Nucleic Acids Research **32:** e37-e37.

Table 3.1. Primers designed for mapping.

Oligo name	Primer #	Sequences (5' - 3')
7_55_Ctg4932F	2891	GTGACAACAGCGCTAGCAGAGCA
7_55_Ctg4932R	2892	TGCATATGCCATGAAGCCGAGCC
7_55_Ctg13423F	2893	GCAGCGCCCAGTGAAGTCGAAG
7_55_Ctg13423R	2894	TGTGTTGGCATCAGCGATGGGG
7_55_Ctg31125F	2895	TTGCCATACCTACCCCGCGTCT
7_55_Ctg31125R	2896	TGTGCCAGCAAAGCATACCTTACC
7_55_Ctg64318F	2897	TGCGGCACGCACATCATTGCTA
7_55_Ctg64318R	2898	GTCACGCATGGATACCGTGGAGC
7_55_Ctg70463F	2899	GTCGGACACCACCGGAGCAAAG
7_55_Ctg70463R	2900	GGCCATGCATACGGAAGGGAGC
7_55_Ctg75992F	2901	TCTCACCGTTCCCTCCACCGAC
7_55_Ctg75992R	2902	CCTCTTGGCCGACGGGGAAAAC
7_55_Ctg95413F	2903	AACGGAGCCACCGCCTAAAACG
7_55_Ctg95413R	2904	ACAACACTCCGCCGCACAAGTT
7_55_Ctg106978F	2905	GCAGCCTAACTGATGGAACCACGC
7_55_Ctg106978R	2906	AGCCCACCATTTTCGTGCAGGT
7_55_Ctg143812F	2907	TGACCAAGATCGTGGCCTGTGC
7_55_Ctg143812R	2908	TGGCATCTTGTATGGTGGCCGAA
7_55_Ctg158353F	2909	CAGTCTCAAGCCCGCTGCAAGG
7_55_Ctg158353R	2910	GCGGCTTTGCCTTCGGGTACAT
7_55_Ctg192225F	2911	AGCTGATGATCTTGATGCGGCGG
7_55_Ctg192225R	2912	CGGTTTGTGCAACCCGGTCGATA
7_55_Ctg194766F	2913	TGACCGAACATGACAAAAGGCA
7_55_Ctg194766R	2914	CCGTAGAATCGAGAAAGCAAGGG
7_55_Ctg232696F	2915	GAGTCTGTCGACCGGAGCGGAA
7_55_Ctg232696R	2916	GGGATCAGGGTGGCACTCCCAA
7_55_Ctg260265F	2917	TTAAGCGTGCCAGACGCTTGGG
7_55_Ctg260265R	2918	GCTCGGTGTCCAGGGTGGGATT
7_55_Ctg293071F	2919	GCAGCAATTGTGGGGAATGTGCC
7_55_Ctg293071R	2920	TGGGAAAGCTGGATCTGCAAGTGT
7_55_Ctg366737F	2921	ATCGTGCCCTGGCTATACGGCT
7_55_Ctg366737R	2922	GGTGAGACTTGCATCGTGGCCG
7_55_Ctg367365F	2923	CGATGGACGTCCTTTGCCCTCA

7_55_Ctg367365R	2924	TCACACCACGCATCGATGTGGTAA
7_55_Ctg429995F	2925	CTGCCGCCGTCCTGAAGAACAT
7_55_Ctg429995R	2926	TGCCGTTTGCAACACCTTCTGTGT
7_55_Ctg608723F	2927	CGCCGTGCGTAACCCAGTCAAT
7_55_Ctg608723R	2928	CCCGTTTCAGTCCACGCTGCAT
7_55_Ctg658018F	2929	GGCAATGCCGCAGTGGCTTTTT
7_55_Ctg658018R	2930	TTGCCGACATATGCTTGCGGCT
7_55_Ctg713348F	2931	GGCCAATCGCCAAGCAGGAGAA
7_55_Ctg713348R	2932	ACCTGAACTTGGCCCCTCACCA
7_55_Ctg793165F	2933	TGCAAGGTCAATCACCTCAGCCT
7_55_Ctg793165R	2934	TCACCAGCATCAACGGTTGGGT
7_63_Ctg423891F	2935	AATGTTGCTGGGCCACGGTCTC
7_63_Ctg423891R	2936	CACGTGCCAGATGCTTCAGATGCT
7_31_Ctg9020F	3026	CCGTGGTATTCCTTTTATTTGC
7_31_Ctg9020R	3027	CAACGATAACAAGCAAAAGCTG
7_31_Ctg305304F	3028	AGAGCAGAGGTCTGGTCAGTTC
7_31_Ctg305304R	3029	CCTTTGATTCCTGTCTCCAGTC
7_31_Ctg2200805F	3030	CGTTTGTGAGGCTGGTTTCT
7_31_Ctg2200805R	3031	CCATCAAGGAACCCACAATG
7_55_Ctg76235F	3032	TATCTTCAGGCTCTCATCAGCA
7_55_Ctg76235R	3033	ATCGGCTATAGTGGTCACTGGT
7_55_Ctg77495F	3034	CACCGTTACGGCAGTACAGTTT
7_55_Ctg77495R	3035	GTAGAGAAGCTCTGCGGACAGT
7_55_Ctg86062F	3036	GGAGAGGTTGTGTCTGCTTCAT
7_55_Ctg86062R	3037	AGCCATCACTGTCTGGTATGTG
7_55_Ctg87055F	3038	TGCAAGGTAGAACCAAGGTACG
7_55_Ctg87055R	3039	CACTGTGTCCATCAATCACCTC
7_55_Ctg93182F	3040	GCAATGAGGTCAGCATGTAGAG
7_55_Ctg93182R	3041	GGGTGCTACAACTCTTGGAAGT
7_55_Ctg143812F2	3042	AACCCGACGATGACCAAGAT
7_55_Ctg143812R2	3043	AAAAGGGGCCCCAGTATGTT
7_55_Ctg223409F	3044	ATCTAGTTCTGCCACCTTGTCG
7_55_Ctg223409R	3045	CACACCAAGGGGATATACACAC
7_55_Ctg221674F	3046	GAAGATTTAGTCGGCCTGATCG
7_55_Ctg221674R	3047	GGGCTCAAATGTTAGGAGTGTC
7_55_Ctg347518F	3048	GCAGGAGAATAGGAGTGCAAGT

7_55_Ctg347518R	3049	GCACAAGTACAGTACCCTGAAGC
7_55_Ctg112355F	3050	GCATTCCATGACCTACCTCACT
7_55_Ctg112355R	3051	AGGAAGCTTTGTCAGTCGTAGC
7_55_Ctg278020F	3052	CCTGTGATTGTGGTAACTCTGG
7_55_Ctg278020R	3053	CGTGGTTACGACCCATTTACTG
7_55_Ctg710134F	3054	AGTGTCGGGTCAAGTTGAGTCT
7_55_Ctg710134R	3055	GCACGATCTAGAGAAAGGGAAC
7_55_Ctg146128F	3056	TCTCTTGAGTAGGGAGCCATTC
7_55_Ctg146128R	3057	GTGCAGCTGTCAGATTTACGAG
7_55_Ctg663335F	3058	CTATATTGGTGCCGAGTCGAG
7_55_Ctg663335R	3059	TATTCTGGGGCAAGAGTCTAGG
7_55_Ctg757126F	3060	ACCCAAGCCCCTCTTTCAAT
7_55_Ctg757126R	3061	GGAGGTGCATTGGACTTATGGT
7_55_Ctg635254F	3062	CTTAGCACCTCCGCTAACACTT
7_55_Ctg635254R	3063	GTGTTGGGCATTTGTTGGTC
7_55_Ctg472609F	3064	CTGCTTCCAGTTTGTGCTACAG
7_55_Ctg472609R	3065	TCTTCAGCCTACGAACAGACAC
7_55_Ctg563223F	3066	GAATCACGGCCTTAAGGTGA
7_55_Ctg563223R	3067	CTGAGGGAGATGTTGAATGG
7_55_Ctg143846F	3068	GGATCGTCGCAAGGAACATT
7_55_Ctg143846R	3069	AGTGCTTGGCTGGTGTAACTG
7_55_Ctg232696F2	3188	GAGAAGGTGCATTTTGACCAAC
7_55_Ctg232696R2	3189	AATCAAGTCACCTCGTTGGATT
7_55_Ctg112355F2	3190	GCTGTTCCACAAGTTTCGGTAT
7_55_Ctg112355R2	3191	TGAATTGGACGATTTTGCATAG
7_55_Ctg76235F2	3192	GCAACCTCTGGTCTCACTGAT
7_55_Ctg76235R2	3193	GTGCTTGAAATGGGGTGACA
7_55_Ctg76235F3	3194	GTGATCACTGCTACCATGCAAT
7_55_Ctg76235R3	3195	GTCTCAGACGGTTCATTTCTGA
7_55_Ctg77495F2	3196	TAGGTTCGGCAGAGTTGATTTT
7_55_Ctg77495R2	3197	GGAAGTGATTGGACAAAGGAAG
7_55_Ctg77495F3	3198	GGGTGATGACGAAGCAATTT
7_55_Ctg77495R3	3199	CCTCTATGCTGGTTCATACACA

Table 3.2. Number of 100 bp Illumina sequence reads generated from the DSN treated libraries. Data provided by Pioneer.

Sample Number	Sequence Filename	Sample Name	Single Read Count	Paired Read Count
1	A7770001	BC_8 Apo_DSN8	22,723,940	268,942,912
2	A7770002	BC_8 Sex_DSN8	109,447,045	161,638,656
3	B4030007	F_1-R5_Apo_DSN	43,253,545	96,779,228
4	B4030008	F_1-R5_Sex_DSN	47,377,584	108,724,480

Table 3.3. Statistics from the Velvet assemblies of BC_8 line 58 apomictic DSN sequences using different K-mer lengths. Data provided by Pioneer. Table header from left to right: Name of the assembly, K-mer length, number of assembled contigs, number of contigs greater than 1 kb in length, minimum contig length in bp, mean contig length in bp, N50 of the assembly, maximum contig length in bp, percentage of reads mapping back to the assembly.

Assembly	K-mers	Contigs	>1K	Min	Mean	N50	Max	% Used
A7770001.min7_31_contigs.fa	31	87,231	13,650	400	728	753	3,901	27
A7770001.min7_39_contigs.fa	39	90,629	16,198	400	750	791	4,695	29
A7770001.min7_47_contigs.fa	47	100,677	18,249	400	751	793	3,709	30
A7770001.min7_55_contigs.fa	55	111,064	20,142	400	751	795	3,953	31
A7770001.min7_63_contigs.fa	63	119,302	21,488	400	748	792	3,927	31

Table 3.4. Contigs selected for SCAR primer design. Table header from left to right: Original Velvet contig name, contig designation for primer name, length of contig in base pair, number of apomictic BC_8 line 58 sequence alignments, number of apomictic F_1 sequence alignments, BlastX top hit against NCBI non-redundant database, BlastN top hit against NCBI sorghum reference genomic sequences.

Contig Name	**Contig-Oligo Name**	**Contig Length**	**BC_8 Apo Hit**	**F_1 Apo Hit**	**BlastX(nr)**	**Sorghum BlastN (e-value)**
A7770001.min7_55_Ctg87055	7_55_Ctg87055	746	890	850	No hit	chr-1 (3e-23)
A7770001.min7_55_Ctg429995	7_55_Ctg429995	861	482	485	No hit	chr-10 (6e-61)
A7770001.min7_31_Ctg305304	7_31_Ctg305304	678	838	112	serine/threonine-protein kinase EDR1-like isoform X2 XP_004955311.1	chr-2 (0.0)
A7770001.min7_55_Ctg713348	7_55_Ctg713348	979	2038	73	No hit	chr-2 (0.0)
A7770001.min7_55_Ctg367365	7_55_Ctg367365	587	323	136	No hit	chr-2 (0.0)
A7770001.min7_55_Ctg112355	7_55_Ctg112355	1130	376	246	No hit	chr-2 (7e-62)
A7770001.min7_55_Ctg366737	7_55_Ctg366737	527	911	207	hypothetical protein EAZ39213.1	chr-2 (4e-57)
A7770001.min7_55_Ctg793165	7_55_Ctg793165	441	139	29	condensin complex subunit 2-like XP_004968877	chr-3 (9e-78)
A7770001.min7_55_Ctg260265	7_55_Ctg260265	1350	707	4332	hypothetical protein ES08426.1	chr-5 (3e-57)

A7770001.min7_55_Ctg663335	7_55_Ctg663335	1010	419	93	uncharacterized protein XP_004974955.1	chr-6 (2e-68)
A7770001.min7_55_Ctg93182	7_55_Ctg93182	1583	2740	10592	probable cleavage and polyadenylation specificity factor subunit 1-like XP_004977785.1	chr-6 (3e-137)
A7770001.min7_55_Ctg194766	7_55_Ctg194766	837	302	136	No hit	chr-6 (1e-117)
A7770001.min7_55_Ctg77495	7_55_Ctg77495	974	394	454	G-type lectin S-receptor-like serine/threonine-protein kinase RLK1-like XP_004975177.1	chr-6 (3e-80)
A7770001.min7_55_Ctg293071	7_55_Ctg293071	769	871	413	No hit	chr-8 (2e-34)
A7770001.min7_55_Ctg76235	7_55_Ctg76235	1027	968	670	xylosyltransferase 1-like XP_004975238.1	No hit
A7770001.min7_55_Ctg158353	7_55_Ctg158353	1092	1023	1556	uncharacterized protein XP_004980471.1	No hit
A7770001.min7_55_Ctg232696	7_55_Ctg232696	971	464	496	Transcriptional corepressor LEUNIG EMS55059.1	No hit
A7770001.min7_31_Ctg2200805	7_31_Ctg2200805	461	116	36	No hit	No hit
A7770001.min7_55_Ctg563223	7_55_Ctg563223	677	203	44	No hit	No hit
A7770001.min7_55_Ctg472609	7_55_Ctg472609	677	1299	50	No hit	No hit

A7770001.min7_55_Ctg757126	7_55_Ctg757126	571	228	87	No hit	No hit
A7770001.min7_55_Ctg146128	7_55_Ctg146128	933	1224	96	No hit	No hit
A7770001.min7_55_Ctg710134	7_55_Ctg710134	614	440	161	No hit	No hit
A7770001.min7_55_Ctg278020	7_55_Ctg278020	758	315	208	No hit	No hit
A7770001.min7_55_Ctg347518	7_55_Ctg347518	611	184	358	No hit	No hit
A7770001.min7_55_Ctg86062	7_55_Ctg86062	476	523	435	No hit	No hit
A7770001.min7_55_Ctg221674	7_55_Ctg221674	676	906	884	No hit	No hit
A7770001.min7_31_Ctg9020	7_31_Ctg9020	604	4096	899	No hit	No hit
A7770001.min7_55_Ctg223409	7_55_Ctg223409	1118	956	937	No hit	No hit
A7770001.min7_55_Ctg4932	7_55_Ctg4932	1390	1554	943	No hit	No hit
A7770001.min7_63_Ctg423891	7_63_Ctg423891	426	690	957	No hit	No hit
A7770001.min7_55_Ctg658018	7_55_Ctg658018	783	933	1104	No hit	No hit
A7770001.min7_55_Ctg95413	7_55_Ctg95413	471	504	1106	No hit	No hit
A7770001.min7_55_Ctg608723	7_55_Ctg608723	476	227	1204	No hit	No hit
A7770001.min7_55_Ctg31125	7_55_Ctg31125	889	1689	1323	No hit	No hit
A7770001.min7_55_Ctg13423	7_55_Ctg13423	1059	1424	1737	No hit	No hit
A7770001.min7_55_Ctg192225	7_55_Ctg192225	1370	3898	2215	No hit	No hit

A7770001.min7_55_Ctg75992	7_55_Ctg75992	1266	2182	2435	No hit	No hit
A7770001.min7_55_Ctg143812	7_55_Ctg143812	2073	1479	2923	No hit	No hit
A7770001.min7_55_Ctg106978	7_55_Ctg106978	820	514	3316	No hit	No hit
A7770001.min7_55_Ctg64318	7_55_Ctg64318	843	276	3700	No hit	No hit
A7770001.min7_55_Ctg143846	7_55_Ctg143846	2024	1715	23	hypothetical protein NP_001131593.1	No hit
A7770001.min7_55_Ctg635254	7_55_Ctg635254	736	972	77	hypothetical protein AAP53925.1	No hit
A7770001.min7_55_Ctg70463	7_55_Ctg70463	995	776	3800	hypothetical protein EAZ18646.1	No hit

Table 3.5. Number of contigs showing ≥ 99% sequence identity against ASGR-BAC and ASGR-recombinant BAC clone sequences. The marker used for BAC isolation and FISH signal location of the BACs within ASGR (***** denotes ASGR-recombinant BAC).

BACs	FISH signal location within ASGR	Marker for BAC Isolation	Number of DSN contig with ≥ 99%BlastN similarity
p303	low copy	A10	6
p002 / p003 / p004	low copy	A14	16
c201	n/a	C4	1
c205	n/a	C4	1
c1000	low copy	HHU27	10
c801	n/a	M02	5
c501 / c522	n/a	O7M	1
p1200	low	pa265	1
p1300*****	outside ASGR	pq355	5
p1000	low copy	py503	12
c004 / c014	low copy	Q8M	6
p102	low copy	Q8M	2
p104	low copy	Q8M	5
c018	n/a	Q8M walk	1
p708	n/a	R13	3
p800 / p801	high copy	U12H	5
c100 / c111	low copy	UGT197	4

c108	low copy	UGT197	3
p201 / p207 / p208	low copy	UGT197	12
p900	high copy	W10M	2

Table 3.6. Summary of mapping results. ***** indicates PCR reaction was with *TaKaRa Ex Taq* HS; Tm: annealing temperature of primers in ºC; +: positive amplification; -: no amplification; N/A: not assayed.

Primer Numbers	**Amplicon Size (bp)**	**Tm (ºC)**	**PCR with Ps26**	**PCR with IA4X**	**PCR with BC_8 apo: sex**	**PCR with F_1 apo: sex**	**PCR with *C. ciliaris* B12-9 : A8 : B-2s**
2891/2892*****	207	65	+	+	+ : +	+ : +	N/A
2927/2928*****	274	65	+	-	+ : -	+ : +	N/A
3030/3031	204	57	+	-	+ : -	+ : +	N/A
3050/3051	211	63	+	-	+ : -	+ : +	N/A
3068/3069	168	62	+	-	+ : -	+ : +	N/A
2893/2894*****	286	66	+	-	+ : -	+ : -	+ : + : +
2895/2896*****	275	65	+	-	+ : -	+ : -	+ : - : -
2897/2898*****	280	68	+	-	+ : -	+ : -	- : - : -
2899/2900	341	69	+	-	+ : -	+ : -	+ : - : -
2901/2902	221	69	+	-	+ : -	+ : -	+ : + : +
2903/2904	200	68	+	-	+ : -	+ : -	+ : + : +
2905/2906*****	209	65	+	-	+ : -	+ : -	+ : + : +
2909/2910	281	68	+	-	+ : -	+ : -	+ : - : -
2911/2912	200	65	+	-	+ : -	+ : -	+ : + : +
2913/2914*****	401	63	+	-	+ : -	+ : -	+ : - : -
2915/2916	289	69	+	-	+ : -	+ : -	+ : + : -
2917/2918*****	249	66	+	-	+ : -	+ : -	+ : - : -
2919/2920*****	275	66	+	-	+ : -	+ : -	+ : + : -
2921/2922	214	68	+	-	+ : -	+ : -	+ : + : +
2923/2924*****	244	69	+	-	+ : -	+ : -	+ : - : -
2925/2926	347	67	+	-	+ : -	+ : -	+ : - : -
2929/2930	243	69	+	-	+ : -	+ : -	+ : + : +
2931/2932	295	68	+	-	+ : -	+ : -	+ : - : -

2933/2934*****	103	65	+	-	+ : -	+ : -	+ : + : +
2935/2936	215	68	+	-	+ : -	+ : -	+ : - : -
3026/3027	332	61	+	-	+ : -	+ : -	+ : - : -
3028/3029	290	63	+	-	+ : -	+ : -	+ : - : -
3032/3033	213	62	+	-	+ : -	+ : -	+ : + : +
3034/3035	207	63	+	-	+ : -	+ : -	+ : + : +
3036/3037	237	66	+	-	+ : -	+ : -	- : - : -
3038/3039	330	63	+	-	+ : -	+ : -	+ : + : +
3040/3041	361	63	+	-	+ : -	+ : -	+ : + : +
3042/3043	121	60	+	-	+ : -	+ : -	+ : - : -
3044/3045	228	63	+	-	+ : -	+ : -	+ : + : +
3046/3047	203	63	+	-	+ : -	+ : -	+ : + : -
3048/3049	270	63	+	-	+ : -	+ : -	+ : - : -
3052/3053	381	65	+	-	+ : -	+ : -	+ : + : +
3054/3055	292	63	+	-	+ : -	+ : -	+ : + : +
3056/3057	235	63	+	-	+ : -	+ : -	+ : - : -
3058/3059	234	66	+	-	+ : -	+ : -	+ : + : +
3060/3061	236	62	+	-	+ : -	+ : -	+ : - : -
3062/3063	361	62	+	-	+ : -	+ : -	+ : + : -
3064/3065	230	65	+	-	+ : -	+ : -	+ : + : +
3066/3067	157	60	+	-	+ : -	+ : -	+ : + : +
3188/3189	109	60	+	-	+ : -	+ : +	+ : + : -
2915/3189	491	60	+	-	+ : -	+ : -	+ : + : -
3190/3191	129	59	+	-	+ : -	+ : +	N/A
3192/3193	97	60	+	+	+ : +	+ : +	N/A
3194/3195	241	60	+	+	+ : +	+ : +	N/A
3196/3197	210	60	+	-	+ : -	+ : +	N/A
3198/3199	235	60	+	-	+ : -	+ : +	N/A

Table 3.7. Contigs chosen and location of SCAR primers for possible gene expression analysis. Region denotes predicted region the primer was derived from based on BlastX alignment of contig to the protein hit; Ex denotes predicted exon region; UTR denotes predicted untranslated region.

Contig Name	Primer Number	Region	BlastX Hit (e-value)
A7770001.min7_55_Ctg232696	3188/3189	Ex/Ex	Transcriptional corepressor LEUNIG (e-11)
A7770001.min7_55_Ctg76235	3192/3193	Ex/Ex	xylosyltransferase 1-like (e-12)
A7770001.min7_55_Ctg76235	3194/3195	3'UTR/Ex	
A7770001.min7_55_Ctg77495	3196/3197	Ex/Ex	G-type lectin S-receptor-like serine/threonine-protein kinase RLK1-like (e-21)
A7770001.min7_55_Ctg77495	3198/3199	Ex/3'UTR	

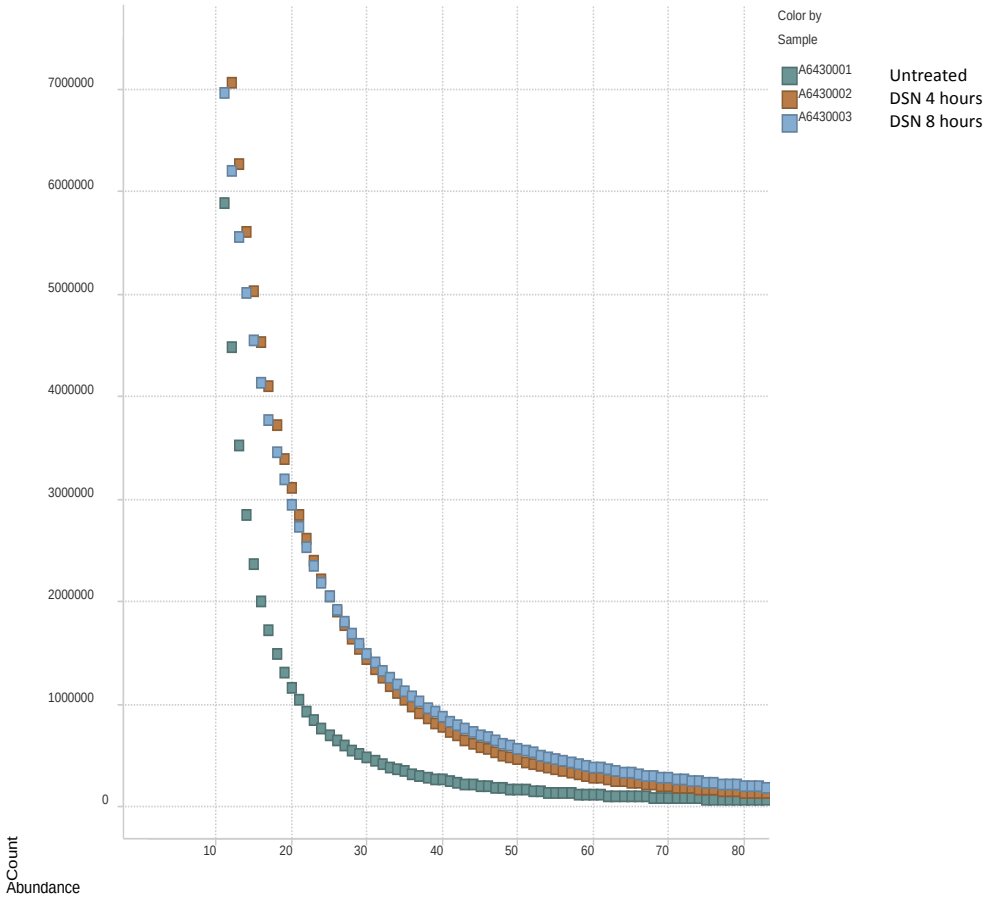

Figure 3.1. Count vs. abundance of 31 K-mers for BC_8 line 58 apomictic control and DSN treated sequences. Data provided by Pioneer. Both 4 and 8 hour DSN treated samples are enriched in counts for low abundance K-mers as compared to the untreated sample.

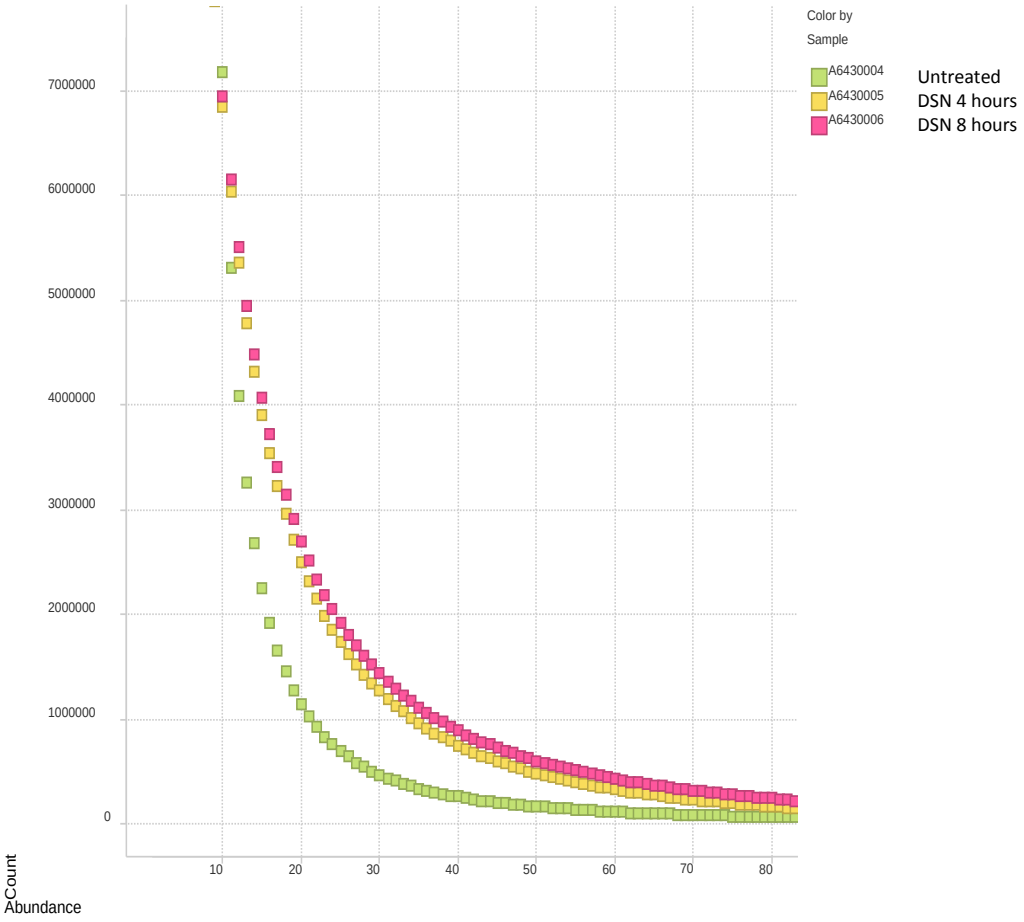

Figure 3.2. Count vs. abundance of 31 K-mers for BC_8 line 58 sexual control and DSN treated sequences. Data provided by Pioneer. Both 4 and 8 hour DSN treated samples are enriched in counts for low abundance K-mers as compared to the untreated sample.

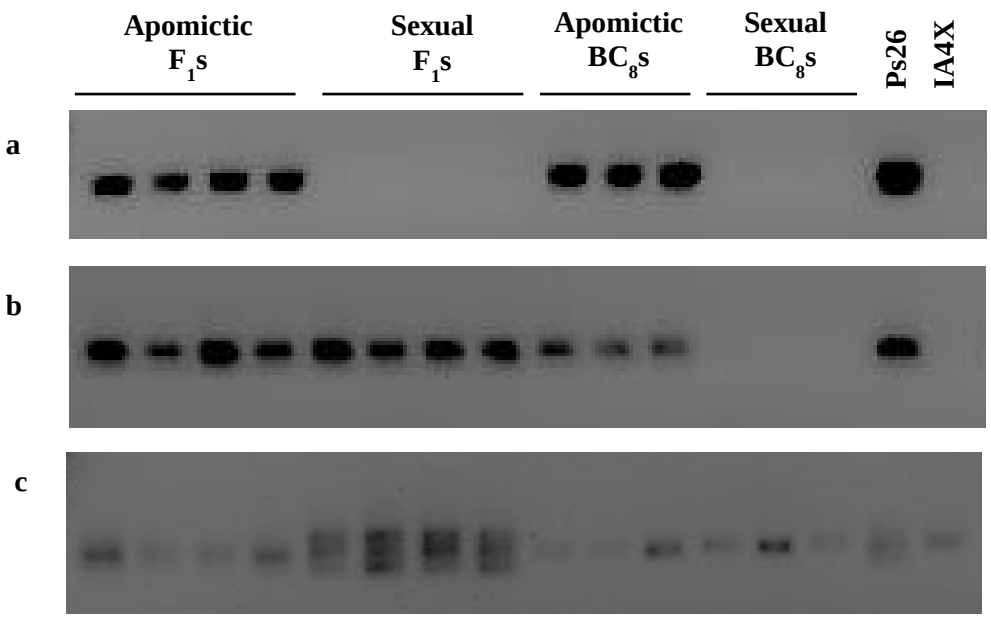

Figure 3.3. Examples for pre-screening of SCAR primers for mapping to the ASGR and the ASGR-carrier chromosome. **a**, ASGR-linked SCAR primers 3060/3061. Amplification is identified in apomictic $F_{1,}$ BC_8 and Ps26 DNA but not in sexual $F_{1,}$ BC_8 or IA4X DNA. **b**, ASGR-carrier chromosome specific SCAR primers 3050/3051. Amplification is identified in apomictic and sexual F_1, apomictic BC_8 and PS26 DNA, but not in sexual BC_8 or IA4X DNA. **c**, Nonspecific amplification of SCAR primers 2891/2892. Amplification in all samples is identified (Apomictic and sexual F_1 and BC_8, Ps26 and IA4X DNA).

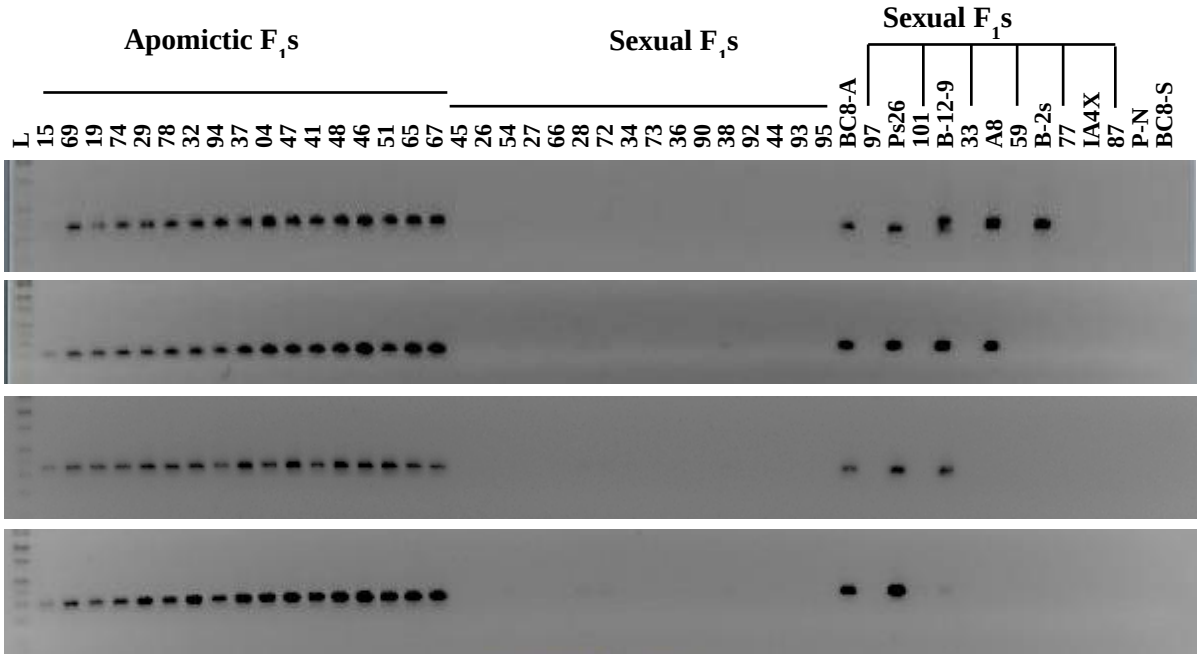

Figure 3.4. Example of mapping ASGR-linked SCAR primers to a larger R5-F_1 population. (L: DNA Hi-Lo marker, BC8-A: BC_8 line 58 apomictic, BC8-S: BC_8 line 58 sexual, P-N: PCR negative control (without template DNA). In figures **a** through **d**, all apomictic F_1 DNA amplified while sexual F_1 DNA did not. **a**, all three *Cenchrus ciliaris* genotypes (B-12-9, B-2s and A8) amplified (p3040/3041). **b**, B-12-9 and A8 DNA amplified but not sexual *C. ciliaris* B-2s (p3046/3047). **c**, only apomictic *C. ciliaris* B-12-9 amplified (p2899/2900). **d**, none of the *C. ciliaris* genotype amplified (p2931/2932).

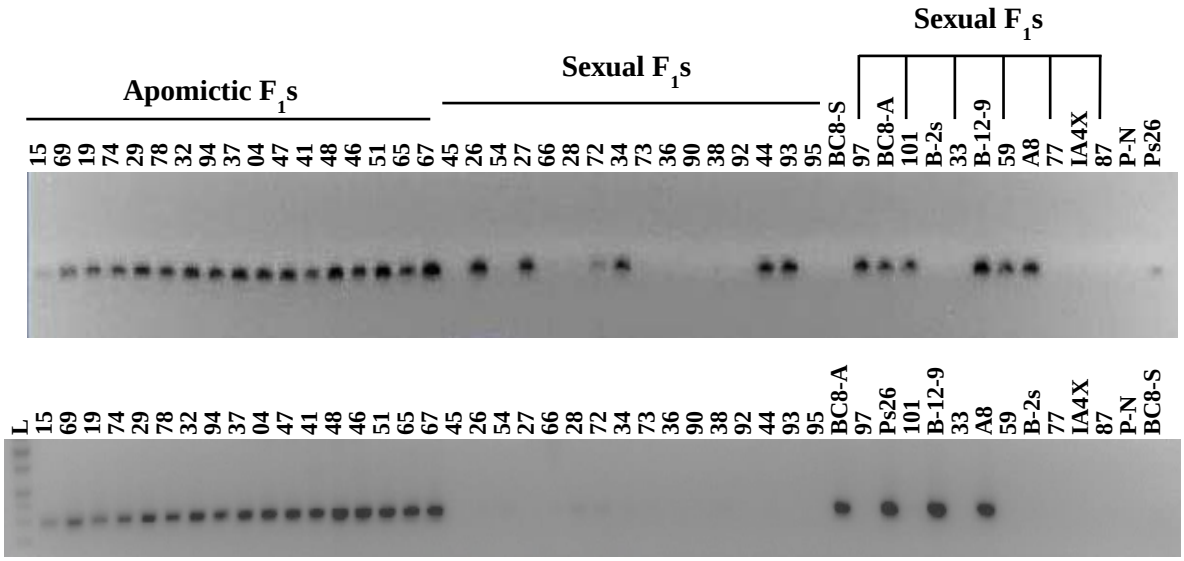

Figure 3.5. Mapping of SCAR primers from contig *A7770001.min7_55_Ctg232696* in F_1 population (L: DNA Hi-Lo marker, BC8-A: BC_8 line 58 apomictic, BC8-S: BC_8 line 58 sexual, P-N: PCR negative control without template DNA). **a**, SCAR primers 3188/3189 amplified F_1 apomicts and nine F_1 sexuals (not specific to ASGR). **b**, SCAR primers 2915/3189 amplified F_1 apomicts but not F_1 sexual (ASGR specific).

CHAPTER 4

SUMMARY

Apospory is a form of asexual reproduction through seed in which embryos develop from unreduced embryo sacs derived from nucellar cells of the ovule bypassing meiosis and fertilization of gametes. Apospory in *Pennisetum squamulatum* and *Cenchrus ciliaris* is controlled by a physically large, hemizygous and heterochromatic chromosomal block called the apospory-specific genomic region (ASGR). The ASGR in *P. squamulatum* is characterized as a recombinationally suppressed region with blocks of an abundant Opie-2 like retrotransposon and located on the telomeric region of the short arm on the ASGR-carrier chromosome. An apomictic backcross 8 (BC_8) line 58 (*P. glaucum*) was identified that contained only the ASGR-carrier chromosome from *P. squamulatum*. To understand more about the genetics and evolution of the ASGR and the ASGR-carrier chromosome, two studies were conducted focusing on physical and genetic mapping of the ASGR and the ASGR-carrier chromosome.

In the first study, we physically mapped eight bacterial artificial chromosome (BAC) clones carrying SCAR markers developed from ovule transcripts derived from the ASGR-carrier chromosome but not mapped specifically to the ASGR due to the lack of marker polymorphisms. A comparative genomic hybridization experiment between sexual and apomictic BC_8 line 58 plants showed enrichment for chromosome 2 of sorghum in apomictic sample. Additionally, *in silico* analysis identified a collinear distribution of orthologous gene for more than 80% of the ASGR-carrier chromosome transcripts on chromosome 2 of sorghum and foxtail millet. However, genic portions of the ASGR did not show large collinearity to any sorghum or foxtail

millet chromosome. Based on these results, we hypothesized that the ASGR-carrier chromosome could be collinear to chromosome 2 of sorghum and foxtail millet outside the ASGR. BAC clones containing the ASGR-carrier chromosome transcripts SCAR markers were identified by screening the polyhaploid BAC library from *Pennisetum*. Eight BAC clones carrying the ASGR-carrier chromosome SCAR markers were physically mapped to the ASGR-carrier chromosome through fluorescence *in situ* hybridization onto metaphase chromosomes from root tips of apomictic BC_8 line 58 plants. Physical mapping revealed large scale collinearity of the ASGR-carrier chromosome to chromosome 2 of sorghum and foxtail millet outside of the ASGR. This suggests that the ASGR was derived by introgression of chromatin into the chromosome evolutionarily similar to chromosome 2 of sorghum and foxtail millet in *P. squamulatum*. This study can be a substantial resource on investigating deeper into the evolution of the ASGR and the ASGR-carrier chromosome in *Pennisetum* and *Cenchrus* species. In addition, the lack of synteny of the ASGR to chromosome 2 of sorghum and foxtail millet can help narrow down the search for putative apomixis candidate genes in *P. squamulatum* based on sequences derived from apomictic BC_8 line 58 plants. Additionally, we now have a physical order of markers on the ASGR-carrier chromosome which can be used to evaluate potential rearrangements of the ASGR-carrier chromosome in gamma irradiated derived plants.

The second study was conducted to determine if sequence characterized amplified region (SCAR) markers identified from duplex specific nuclease normalized Illumina libraries would genetically map to the ASGR at a high frequency. Duplex specific nuclease (DSN) is a thermostable enzyme that displays strong preference for cleaving double stranded DNA and DNA in RNA-DNA hybrid duplexes compared to single stranded DNA and RNA. Since the ASGR is abundant in an Opie-2 like retrotransposon and sequence analysis of the *ASGR-BBM-*

like gene suggests a relatively recent origin of the ASGR, we hypothesized that DSN normalization might be an effective strategy for depletion of repetitive sequences and enrichment of unique low copy sequences from the ASGR for genetic. Enrichment of sequences with lower abundance was observed for DSN treated BC_8 line 58 DNA compared to the untreated control, when sequences 31 bp in length were counted and graphed. Forty-four potential ASGR-specific contig sequences were designed for SCAR markers. Thirty-nine mapped specifically to the apomicts in an F_1 mapping population derived from the cross between apomictic *P. squamulatum* and sexual *P. glaucum*. No recombination was detected for any of the mapped SCARs. Eighteen of the SCARs mapping specifically to apomicts in *Pennisetum* also showed amplification specificity to apomicts in a limited study of buffelgrass (*C. ciliaris*). Four markers mapped to the A8 plant, a recombinant buffelgrass plant that is capable of producing unreduced aposporous embryo sacs but has lost the ability to undergo parthenogenesis. Genetic mapping of the SCAR markers from putative exon regions of the ASGR-specific contigs with significant homology to proteins lost specificity to the ASGR. We show that the DSN normalization of genomic DNA libraries can be an efficient strategy on targeted mapping of a physically large. The ASGR-specific markers identified from the study can be used for isolation of BAC clones and physical mapping to the ASGR relative to already known markers or for continuing sequence analysis to identify putative candidate apomixis genes residing at the ASGR. Since the first study of this thesis has shown that the chromosome 2 of sorghum and foxtail millet have a large scale collinearity with the ASGR-carrier chromosome outside the ASGR, markers from contigs showing homology to sorghum and foxtail millet chromosome 2 can be screened in larger F_1 population to identify new recombination breakpoints of the ASGR.