QTL MAPPING FOR REPRODUCTIVE MATURITY

IN LOWLAND SWITCHGRASS POPULATIONS

By

HONGXU DONG

Bachelor of Science in Agronomy

Shandong Agricultural University

Taian, Shandong, China

2012

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 2014

QTL MAPPING FOR REPRODUCTIVE MATURITY

IN LOWLAND SWITCHGRASS POPULATIONS

Thesis Approved:

Dr. Yanqi Wu

Thesis Adviser

Dr. Michael Smith

Dr. Liuling Yan

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my adviser Dr. Yanqi Wu for his patient and detailed instructions. I would also like to thank Dr. Liuling Yan and Dr. Michael Smith for their willingness to serve on my committee.

I acknowledge Dr. Linglong Liu, Dr. Yaling Huang, Mr. Josh Garner for genotyping work and Mr. Gary Williams, Ms. Pu Feng, Mr. Seth Michael Davis, and Mr. Ethan Purkins for field work assistance. Gratitude also goes to Ms. Tilin Fang, Mr. Shiva Makaju, Mr. Laxman Adhikari, Ms. Dan Chang, and Ms. Yuanwen Guo for their help in my thesis defense preparation. This research was sponsored by NSF EPSCoR award EPS-0814361.

The thesis is dedicated to my family. First of all, I would like to thank my parents for their love. Their support, encouragement and love, has always motivated me to overcome obstacles in my life. I love my parents so much. Secondly, I would like to express my gratitude to my older sister, thank you for your sacrifice for our family, I miss the time we had when growing up. Thanks for everything you have done to me and thanks for the generous love you gave me.

Name: HONGXU DONG

Date of Degree: MAY, 2014

Title of Study: QTL MAPPING FOR REPRODUCTIVE MATURITY IN LOWLAND SWITCHGRASS POPULATIONS

Major Field: Plant and Soil Sciences

Abstract:

Switchgrass (*Panicum virgatum* L.) has high potential to be a major cellulosic bioenergy crop. Selection for later flowering plants will extend the growing season, likely resulting in larger biomass yields. However, it is little known of the genetic structure and mechanism for reproductive maturity in switchgrass. Accordingly, the major objective of this study was to identify genomic regions for reproductive development. Two lowland switchgrass populations, a hybrid population consisting of 176 progeny derived from a cross between parents NL94 (♀) × SL93 (♂) and a first-generation self-fertilized population of 265 progeny from NL94, were used in this study. Significant genetic variation for reproductive maturity stages was observed within each of the two populations. A total of 178 simple sequence repeat (SSR) markers were genotyped in the hybrid population for the construction of a linkage map while a pre-existing linkage map of 439 SSR markers was used for quantitative trait loci (QTL) analysis between markers and phenotypic data. QTL analysis revealed that reproductive maturity was a complex trait as controlled by multiple genomic regions. The QTL regions between PVGA-1727/1728 and PVGA-1201/1202 on linkage group (LG) 3b, between nfsg-125 and PVE-781/782 on LG 2b, and between PVCAG-2503/2504 and PVAAG-3253/3254 on LG 7a were identified to be associated with reproductive maturity in both populations. Broad sense heritabilities were 0.08 to 0.66 and 0.03 to 0.48 for the selfed and hybrid populations, respectively. Use of the markers linked to the significant QTLs in the populations could accelerate the development of switchgrass varieties having later flowering time as a means in increasing biomass yield in switchgrass.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Switchgrass (*Panicum virgatum* L.) is a predominant tallgrass species of the North America prairies (Bouton 2008; Wright et al. 2010). Multiple merits of switchgrass make it a highly suitable herbaceous candidate for cellulosic feedstock production, including high biomass yield potential, adaptation to marginal lands, strong stress resistance, minimal requirement of agronomic inputs, stand longevity and ease of management (McLaughlin et al. 1999). On the basis of morphological difference and habitat preference, switchgrass is mainly classified into two distinct ecotypes: lowland and upland (Porter 1966). In the southern Great Plains, lowland switchgrass has much higher biomass yield potential than upland ecotypes (Cassida et al. 2005; Fuentes and Taliaferro, 2002; US Department of Energy 2011; Casler 2012). Therefore, breeding programs in the region have targeted on improving germplasm and developing superior cultivars in lowland switchgrass (Bouton 2007).

Currently, a major goal of breeding efforts in switchgrass as a cellulosic bioenergy crop is focused on improving biomass yield (Casler et al. 2011). Potentially maximum yields of switchgrass are harvested from populations which have longer vegetative growth and later reproductive development. Newell (1968) reported, in a four-year field trial at three locations in Nebraska, that late maturing southern switchgrass collections produced higher yields than early maturing northern strains except in a western site where growing seasons were so short that the southern strains did not reach their full production potential. Talbert et al. (1983) reported that

maturity was negatively correlated with dry matter yield although the correlation was not large (-0.33). Sanderson et al. (1999) performed two field trials of three lowland and six upland ecotypes at five Texas locations and reported the germplasm especially upland ecotypes originated in the Midwest matured much earlier and produced significantly less biomass than the southern lowland accessions, indicating again maturity was related to biomass yield. In a two-year field trial at two Wisconsin locations, Casler (2014) found flowering time was a key factor affecting biomass yield, which explained 67% of the variation among hybrids in biomass yield, and most importantly, biomass yield could be increased by 0.47 Mg ha^{-1} for each day delay in flowering time. The previous results have revealed that the timing of reproductive maturity is an important factor among others determining biomass yield. Consequently, southern germplasm when grown at northern locations have a longer period of vegetative growth, later flowering, and bigger biomass yields than northern early-maturing populations (Newell 1968; Casler et al. 2007). However, growing southern late-maturing populations in northern sites should not be recommended beyond one hardiness zone than their origin due to severe stand losses to winterkill (Newell 1968; Casler et al. 2004, 2007).

Reproductive development is a heritable trait. In the southern United States, as switchgrass plants mature before the end of the growing season, Van Esbroeck et al. (1998) proposed if the duration of vegetative growth of switchgrass was extended through selecting for delayed flowering its biomass yield might increase. The group effectively extended vegetative growth for approximately two weeks by selecting for late flowering plants in 'Alamo' switchgrass, which produced one or two more stem leaves than early flowering plants (Van Esbroeck et al. 1998). In most previous investigations researchers used heading date and/or flowering time to measure earliness versus lateness of reproductive development in switchgrass populations. Talbert et al. (1983) and Van Esbroeck (1998) reported there was large genetic variation for flowering time in lowland switchgrass germplasm. Talbert et al. (1983) reported

high narrow-sense heritabilities (0.91 or above) for switchgrass maturity based on a lowland population of 33 half-sib families. Using 37 lowland half-sib families of one lowland population, Bhandari et al. (2010) reported moderate to high (0.58-0.74) narrow sense heritabilities for heading and flowering time. Bhandari et al. (2011) observed heritability estimates for heading date were larger based on full-sib families than on half-sib families, suggesting dominant gene effect or epistasis likely played an important role. Investigating on upland switchgrass half-sib families, Price and Casler (2014) reported high narrow-sense and realized heritability estimates for flowering time. They recommended flowering time should be used as an effective secondary trait to biomass yield for within-family selection.

Previous experiments have demonstrated that there is substantial variation in reproductive maturity and that selection for delayed flowering time based on field experiments is effective, but the procedures have limitations. Price and Casler (2014) correctly indicated successful field-based selection for late flowering requires large spaced-plant nurseries to assure sufficient variation, and large amounts of time to accurately measure flowering date in the field. If molecular markers are identified to be significantly linked to the genetic variation of reproductive maturity, marker-assisted selection could be used as an alternative in the development of later maturing germplasm. Simple sequence repeat (SSR) markers have been proved to be highly informative due to their high polymorphism and codominant inheritance. Complete and relatively high-density genetic maps have been constructed using SSR markers in switchgrass (Okada et al. 2010; Liu et al. 2012, 2013). To our knowledge, no information is available on association between molecular markers and reproductive maturity in switchgrass. Accordingly, the major objective of this study was to identify genomic regions associated with reproductive maturity in lowland switchgrass using SSR markers.

CHAPTER II

REVIEW OF LITERATURE

Significance of switchgrass for bioenergy production

Modern society's overwhelming dependence on fossil fuels poses great concerns. As petroleum, coal, and natural gas become exhausted, we must find new sources of energy to face global economic development. But ironically, use of fossil fuels lead to global environmental disruptions (Parrish and Fike 2005). Based on knowing that, the United States of America has investigated an array of resources to develop biofuels as alternative to fossil fuels. Ethanol is an environment friendly biofuel source, which can be produced from feedstock resources like sugar, starch, and cellulosic biomass. The use of ethanol for fuel is an effective way to reduce greenhouse gas emissions and bring a boost for energy independence. Many countries have devoted great energy to the development of new technologies used in converting cellulosic biomass to ethanol. The current bioethanol production in the US is almost entirely based on corn (Petrulis et al. 1993). However, many studies indicated that using corn for bioethanol production was not appropriate. First of all, growing corn for ethanol production occupies large areas of cropland and could negatively affect food production (Varvel et al. 2008). Secondly, corn plantation requires lots of energy in field management like irrigation, fertilizer, pesticide and herbicide (Patzek et al. 2005). In 2011, switchgrass was recognized as a promising bioenergy crop in the Billion-Ton Update report (U.S. Department of Energy, 2011). According to a farm-scale study of switchgrass, Schmer et al. (2008) reported that switchgrass produced 540% more

energy than the energy input for feedstock production, and greenhouse gas emissions from converting switchgrass feedstock to ethanol was significantly reduced by 94% compared with that from gasoline. Unlike corn, switchgrass can grow on marginal lands and requires relatively modest levels of chemical fertilizers. Overall, it is considered a resource-efficient, low-input crop for producing bioenergy from farmland. Moreover, The main advantage of using switchgrass over corn as an ethanol feedstock is its cost of production is generally about 1/2 that of grain corn, and more biomass energy per hectare can be captured in the field. Thus, switchgrass cellulosic ethanol should give a higher yield of ethanol per hectare at lower cost.

Biological and agronomical characteristics of switchgrass

Switchgrass is a C4 perennial grass which is native to North America. It is an important member of the tribe Paniceae in the subfamily Panicoideae of the family Poaceae (Wang et al. 2011). The plant grows from 1 to 3 m tall with outstanding stand longevity, once established, it can sustainably produce biomass for more than 10 years (Garland 2010).

Switchgrass can be used for multiple aspects. In addition to the biomass energy production, it is natural habitat for wild life. It can be grown as ground cover for soil conservation, for forage and grazing. Farmers also use it as a substitute for wheat straw in many applications, including livestock bedding, straw bale housing, and as a substrate for growing mushrooms. Additionally, switchgrass can be used as ornamental grass because of its drought-resistant characteristic.

Switchgrass is naturally distributed over large geographical areas spanning from 15 to 55 degree north latitude (Hitchcock 1951), because of its wide distribution, switchgrass is normally classified into two ecotypes: lowland and upland switchgrass (Porter 1966). Further studies recognized four subecotypes, northern upland, south upland, northern lowland and southern lowland, based on their latitudinal adaptation (Casler et al. 2004). Lowland switchgrass grows

well in more moist low areas with warmer temperatures such as the southern USA, plants are taller and coarser than upland plants, while upland switchgrass is mainly distributed in upland sites, plants have a more spreading habit due to longer rhizomes. More importantly, lowland switchgrass has higher biomass yield compared with upland switchgrass (Bouton 2007; Sanderson et al. 1996).

Switchgrass is a largely out-crossing species with self-incompatibility (Talbert et al. 1983; Taliaferro et al. 1999). In field condition, switchgrass plants tend to produce hybrid seeds with wind-facilitated pollination. Although strong self-incompatibility is prevailing in switchgrass populations, successful attempts had been made in identifying self-compatible plants, Liu and Wu (2012) found one lowland plant 'NL94' exhibiting high self-compatibility. With a base chromosome number of nine, switchgrass comprises an array of ploidy levels, from diploid (2n=2x=18) to 12-ploid (2n=12x=108). Lowland switchgrass are predominantly tetraploid (2n=4x=36), while upland switchgrass is largely octoploid (2n=8x=72). Molecular genetics studies revealed that tetraploid switchgrass has a disomic inheritance mode (Liu et al. 2012; Okada et al. 2010).

Breeding switchgrass for bioenergy production

In switchgrass, major efforts are currently being undertaken to improve biomass yield and enhance the traits related to conversion efficiency from cellulose to ethanol and butanol (Bouton 2008). Biomass yield in switchgrass is mainly determined by number of phytomers per tiller and weight per phytomer (Boe et al. 2005). Other factors like frequency of reproductive tiller production, phytomer development rate also play pivotal roles in biomass yield and seed production (Bouton 2007). Three traits have been commonly used as indirect selection criteria for biomass yield: plant height, tiller count, and date of flowering. In upland tetraploid switchgrass, among-and-within-family selection proved moderate heritability (0.41) for plant height, with

greater values for selection of increased height, while heritability for tiller count was generally low (0.06), and flowering date was estimated to have high heritability (0.75) overall in both selection direction (Price et al. 2014). In lowland switchgrass, according to the research of Bhandari et al. (2010), half-sib families were different for biomass yields and other traits, suggesting that the presence of additive gene effect in controlling these traits. Heritabilities were moderate (0.40-0.70) for heading, flowering, and plant spread.

Being a C4 crop, switchgrass is forty percent more efficient in photosynthetic activity than C3 crops and thus the energy output potential is higher (Samson et al. 2005). A higher photosynthetic efficiency results in more net energy gain during vegetative growth period. However, improvements in photosynthetic efficiency have been limited (Evans 1993), many annual crop breeders have selected for an optimal duration of growth as the avenue to increasing biomass yield (Wallace et al. 1993). For switchgrass, Esbroeck et al. (1998) proposed that an extended duration of vegetative growth in switchgrass by selecting for delayed flowering might be a means to achieve higher biomass yield. Besides, as in the classification of switchgrass ecotypes, a later maturity and more rapid stem elongation rate in lowland switchgrass give rise to a longer retention of photosynthetic tissues and therefore an accompanying higher biomass yield potential compared with upland switchgrass (Casler et al. 2004). Therefore, clarifying the genetic basis for reproductive maturity has considerable importance in switchgrass breeding.

Reproductive maturity and its effect on biomass production in switchgrass

Reproductive development is a heritable trait. Plant development progresses through two distinct phases: vegetative growth, followed by a reproductive phase. During vegetative growth stage, plants generally rapidly increase their photosynthetic capacity and their size and mass. Then, the reproductive maturity phase occurs, during which plants are busy with the production of new individuals or offspring.

Reproductive maturity is a key developmental stage related to biomass yield in switchgrass. Studies on other crops like wheat, maize, rice and cotton, indicated that timing to maturity had significant associations with biomass yield and other related traits (Halloran 1977; Russell and Stuber 1983; Salam and Mackill 1993; Li et al. 2013).

Potentially maximum yields of switchgrass are harvested from populations which have longer vegetative growth and later reproductive development. Newell (1968) reported, in a four-year field trial at three locations in Nebraska, that late maturing southern switchgrass collections produced higher yields than early maturing northern strains except in a western site where growing seasons were so short that the southern strains did not reach their full production potential. Talbert et al. (1983) reported that maturity was negatively correlated with dry matter yield although the correlation was not large (-0.33). Sanderson et al. (1999) performed two field trials of three lowland and six upland ecotypes at five Texas locations, indicating the germplasm especially upland ecotypes originated in the Midwest matured much earlier and produced significantly less biomass than the germplasm from the southern lowland accessions, indicating again maturity was related to biomass yield. The previous results together may have revealed that the timing of reproductive maturity is an important factor among others determining biomass yield. Consequently, southern germplasm when grown at northern locations have a longer period of vegetative growth, later flowering, and bigger biomass yields than northern early-maturing populations (Newell 1968; Casler et al. 2007). However, growing southern late-maturing populations in northern sites should not be recommended beyond one hardiness zone than their origin due to severe stand losses to winterkill (Newell 1968; Casler et al. 2004, 2007).

QTL mapping in plant breeding

Quantitative trait loci (QTLs) are stretches of DNA containing or linked to the genes that underlie a quantitative trait of interest. Mapping regions of the genome that contain genes

involved in specifying a quantitative trait is done using molecular tags such as simple sequence repeat (SSR). This is an early step in identifying and characterizing the actual genes underlying trait variation. QTL analysis is a statistical method that links two types of information—phenotypic data (trait measurements) and genotypic data (usually molecular markers)—in an attempt to explain the genetic basis of variation in complex traits. QTL analysis allows researchers in fields as diverse as agriculture, evolution, and medicine to link certain complex phenotypes to specific regions of chromosomes. The goal of this process is to identify the action, interaction, number, and precise location of these regions.

QTL mapping bring great help to breeders in linking quantitative phenotypic variation to qualitative genotypic marker polymorphism, and speed up the development of improved cultivars (Wang et al. 2011). The first genetic map was constructed with 102 restriction fragment length polymorphism (RFLP) single dosage markers (Missaoui et al. 2005), which were distributed in eight homology groups covering over 400 cM. Simple sequence repeat (SSR) markers, also known as microsatellites, are tandem repeats of 2-6 bp DNA sequence. SSRs are most widely used because of its multiple merits, like high information content, codominant inheritance pattern, easy use, and reproducibility (Kashi et al. 1997). Complete and relatively high-density genetic maps have been constructed using SSR markers in switchgrass (Okada et al. 2010; Liu et al. 2012, 2013).

To our knowledge, no information is available on association between molecular markers and reproductive maturity in switchgrass. Accordingly, the major objective of this study was to identify genomic regions associated with reproductive maturity in lowland switchgrass using SSR markers.

CHAPTER III

METHODOLOGY

Plant materials

Two mapping populations consisted of 441 progeny along with two parental genotypes NL94 and SL93, of which 265 were first-generation inbred lines derived from self-fertilization of NL94, the rest 176 individuals were hybrids from a cross between NL94 (♀) × SL93 (♂). The NL94 parent was chosen in the summer of 2007 from the Oklahoma State University (OSU) northern lowland (NL) breeding population in a low yield environment selection nursery. The SL93 parent was chosen at the same time in 2007 from the OSU southern lowland (SL) breeding population. One cross between NL94 and SL93 parents was made in September to November, 2007. One potted NL94 plant and one SL93 plant were prepared in a greenhouse and just before flowering they were moved into a large growth chamber at the OSU Controlled Environmental Research Laboratory. Seedlings from the seeds harvested on NL94 parent were composed of 279 selfed and 177 crossed progeny as identified by SSR markers (Liu and Wu 2012). Technical details for the parent plants and two progeny populations were described by Liu and Wu (2012). In 2009, individual plants of the two populations and parents were transplanted into a space-plant field nursery on the OSU Agronomy Research Farm. In the summer of 2010, for each member of the two mapping populations and parents, approximately 20 clones were cultivated in individual containers in a greenhouse from ramets or from dormant nodal buds on stems of plants grown in the space-plant nursery (Wu 2014).

Experimental design, establishment and management of field trials

Two field trials of the mapping populations with their parental plants were established in 2011, one on the OSU Agronomy Farm, Stillwater, and the other on Cimarron Valley Research Station (CVRS), Perkins, OK. Soil types were tested to be Kirkland silt loam and Teller fine sandy loam in the Stillwater field and Perkins field, respectively. Plots were arranged in a randomized complete block design at each location, with three replications. Each plot contained three clonal plants of one genotype. The trial in Stillwater was set out in 10 plots per row by 45 rows for each replication, and the trial in Perkins was arranged in 15 plots per row by 30 rows for each replication. Spacing between two neighboring plants in a row and between rows was 1.07 m. Border rows were planted around each field to minimize border effects.

Clonal plants were transplanted into a nursery on the Agronomy Farm on May 16-17 and into a nursery of CVRS on June 1-7, 2011. After transplanting, the two nurseries were immediately sprayed with 1.12 kg Atrazine (6-chloro-N-ethyl-N-isopropyl-1,3,5-triazine-2,4-diamine), 1.12 kg Surflan (Oryzalin: 3,5-dinitro-N^4N^4-dipropylsulfanilamide), and 0.007 kg Escort (Methyl 2-[[[[(4-methoxy-6-methyl—1,3,5-triazin-2-yl) amino] sulfonyl] benzoate) a.i. per ha. Irrigation was applied to provide sufficient soil moisture in the two nurseries for two weeks post transplanting. In the establishment year, no fertilizer was given to the nurseries. In the winter seasons, 2011-2012 and 2012-2013, plants were harvested at 10 cm height above ground surface.

In March, both 2012 and 2013, Atrazine at 2.24 kg, Surflan at 2.24 kg, Roundup (Glyphosate: N-(phosphonomethyl) glycine) at 4.48 kg, a.i./ha were applied before switchgrass plants greened up. In March, to help identify correct row numbers for phenotypic data collection, white posts were installed on the west end every 10 rows. In May, when switchgrass plants

actively grew, urea was applied at a rate of 67.2 kg N/ha. Contaminants and weedy plants occurred in the summer were removed by spot-spraying of Roundup or hand weeding.

Phenotypic evaluation of reproductive maturity

To assure the accuracy of phenotyping, an orange ribbon was tied to a main stem of one representative plant (mostly the middle plant) among three plants per plot before phenotypic data were collected. Phenotypic data were collected based on a numerical system ranging from 1 to 7, in which 1 representing the booting stage to 7 being maximum flowering (Table 1). The phenotypic scale system was developed according to Moore et al. (1991) and Sanderson (1992). Reproductive maturity was evaluated two times in August for the nursery at Stillwater and one time at Perkins in 2012 and two times, for each location in 2013, with an interval of two weeks between two sequential phenotypic data collections.

Table 1 Numerical indexes and corresponding morphological descriptions for scoring reproductive maturity stages

Score	Description of morphological characters
1	Boot stage, inflorescence palpable or visible in flag sheath
2	Initiation of exsertion of inflorescence, tip of inflorescence is visible and without spreading branches
3	Medium exsertion of inflorescence, branches of inflorescence spread out and length of inflorescence reaches about 50%
4	Full exsertion of inflorescence, last branch of inflorescence is out of flag leaf sheath.
5	Initiation of anthesis, florets of less than 30% length of inflorescence are flowering
6	Medium anthesis, florets of 30-70% length of inflorescence are flowering
7	Maximum anthesis, florets of more than 70% of inflorescence are flowering

SSR genotyping and genotypic data collection

DNA isolation from fresh leaf tissues was conducted with a CTAB method (Doyle and Doyle, 1990), with minor modifications made according to Liu and Wu (2012). SSR markers developed by Wang et al. (2011) were used to screen for polymorphism using both parents and six randomly selected hybrid progeny. Then polymorphic SSRs were used to genotype 132 individuals (including two parents) randomly selected from 177 hybrids derived from the cross between NL94 (♀) × SL93 (♂). The reason why 132 individuals were used for genotyping work was determined by the genotype-detecting equipment, LI-COR 4300 DNA Analyzer which allowed 66 samples loaded in each gel. Fluorescence-labeled polymerase chain reaction (PCR) and gel electrophoresis were performed on Biosystems 2720 thermal cyclers (Applied Biosystems, CA) and LI-COR 4300 DNA Analyzer (LI-COR, Lincoln, NE) , respectively. PCR chemical recipe, thermal conditions, and cycle numbers followed the routine procedure outlined by Wu and Huang (2008). At the end of PCR, 5 μl Blue Stop Solution (90% formamide in bromophenol blue) was added to the DNA sample in each well, mixed thoroughly, spin down, and run for extra 3 minutes at 94 ℃ in the thermal cycler. PCR products of one plate labeled with 700-nm fluorescent dye were mixed with the other plate labeled with 800-nm dye. Then 0.5-0.8 μl of each mixed PCR sample was loaded into each well of a 6.5% KB^{plus} gel in 1X Tris borate-EDTA buffer and run at a constant 1500V for 1.5 h on the LI-COR 4300 DNA Analyzer.

For the hybrid population, the type of segregation may vary across SSR markers. Up to four different alleles may be segregating in the progeny population. All the markers were visually scored following the segregation type according to JoinMap 4.0 Manual (Van Ooijen 2006), and genotypic data were recorded into an Excel spreadsheet according to the data file format described in the Manufacturer's manual.

For the selfed population, all the markers were originally recorded as <hkxhk> pattern (locus heterozygous in the parent) by Liu et al. (2012; 2013). SSR-amplified fragments with only one upper band were scored as 'hh', with two bands were scored as 'hk', and with only a lower band were scored as 'kk'. Because selfed population type is not available for analysis in JoinMap and MapQTL, and the outcross (CP) full-sib family population type was also unfeasible because all the markers segregate as <hkxhk> in either coupling {00} or repulsion {11} phases in both parents, resulting in singularity errors for QTL analysis (Van Ooijen 2006), then according to the linkage phase information automatically calculated in JoinMap 4, all the markers were recoded following the format of population type F2: phases {00}: hh>a, hk>h, kk>b; phases {11}: hh>b, hk>h, kk>a.

Data analysis

Linkage analysis was conducted using JoinMap 4.0. The F2 population type was used for the selfed population. Segregation ratios of markers were calculated using chi-square test for goodness-of-fit to the expected ratios. If markers showed severe segregation distortion ($P<0.0001$) they were removed from the analysis. Marker information of the selfed progeny plant "No. 166" was deleted because of its absence in the field. Then all the markers were grouped into linkage groups at a minimum independence test LOD score of 7.0. Maximum-likelihood (ML) mapping algorithm was used to order the loci within each linkage group (LG). Finally, 439 markers were grouped into 18 LGs. After grouping, map distance was calculated using Kosambi mapping function (Stam 1993). For the hybrid population, the outcross pollinated (CP) full-sib family was used. Linkage analysis of the hybrid population followed similar procedure as used in the selfed population. The linkage map of the hybrid population was added with labels of dominant markers and segregation distorted markers using Mapchart 2.2 (Voorrips 2002).

SAS/MEANS was used to calculate mean values and associated standard deviations for phenotypic data collected at each time in each of the two populations (SAS Institute 2003). Microsoft Excel 2010 was used to generate histograms of phenotypic data. SAS/COOR was performed to calculate correlation coefficients among different trials. SAS/MIXED procedure was used to do ANOVA analysis and to obtain the variance components with TYPE3 sum of squares as the estimation method. Broad sense heritabilities (h^2) were calculated using the following formula: $h^2=\hat{\sigma}_g^2 \ / \ [\hat{\sigma}_g^2 + \hat{\sigma}_{g\,e}^2 + (\hat{\sigma}_{error}^2 / \ r)]$, where $\hat{\sigma}_g^2$, $\hat{\sigma}_{g\,e}^2$, $\hat{\sigma}_{error}^2$ are genotypic variance, genotype-by-environmental variance and error variance, respectively, and r is the number of replications per environment.

Mean values of the reproductive maturity ratings for each family at different time points were used for QTL mapping analysis. For QTL analysis, interval mapping (IM) and multiple-QTL model (MQM) mapping were performed using MapQTL 6 (Van Ooijen 2009). At a significant p value of 0.05, LOD threshold was calculated by a 1,000 permutation test. After QTL detection in two populations, markers flanking common QTLs were genotyped in the whole selfed population to confirm the results.

CHAPTER IV

FINDINGS

Phenotypic data analysis

Means and associated standard deviations of reproductive maturity ratings in two parents and two mapping populations were given in Table 2. There was substantial variation in reproductive maturity in the two populations (Table 2). The ANOVA analyses indicated that plant genotype consistently had significant effects on the phenotypic variation of reproductive maturity in the two populations while the effects of location and plant by location interaction on the trait varied (Table 3). Variance components are presented by population for each dataset collected in 2012 and 2013 (Table 4). Broad sense heritabilities for reproductive maturity ranged from 0.08 to 0.66 and 0.03 to 0.48 for the hybrid population and selfed population, respectively. Frequency distributions of phenotypic data are shown for the hybrid population (Fig. 1) and selfed population (Fig. 2), separately. The positive correlations among phenotypic values in datasets collected at different time suggested that genetic control of reproductive maturity over time is significantly related (Table 5; Table 6). In both populations, reproductive maturity ratings in 2012 generally had smaller correlations with those in 2013, ranged from 0.21 to 0.49 for the hybrid population, and 0.26 to 0.51 for the selfed population. Reproductive maturity ratings in 2013 generally had medium to high correlations among different time points across two locations. However, in 2012 the reproductive maturity ratings correlation was relatively low between two locations for both populations, about 0.26 for the hybrid population, and 0.49 to 0.58 for the

selfed population.

Linkage analysis for the hybrid population

One hundred seventy eight polymorphic SSR markers were genotyped to generate a linkage map for the hybrid population (Fig. 3), of which four markers were tested to be dominant while all others were codominant (Table 7). The number of loci per linkage group (LG) varied from 3 (LG 7b and 8a) to 22 (LG 3b). The total length of the map was 1080 cM, and the average distance between two adjacent markers was 6.1 cM. Sixteen gaps were found with a distance > 15 cM, which may not be suitable for QTL analysis and marker-assisted application (Beckmann and Soller 1983). The LGs were named according to the previous maps by Okada et al. 2010. However, compared with the high-density linkage map of the selfed population (Liu et al. 2013), LG 4a and 4b merged into one single group in the hybrid population, which resulted in a total of 17 LGs.

QTL detection and their effects

For the selfed population, the IM identified 5 QTLs affecting reproductive maturity, mainly distributed on LGs 2b, 7a (Table 8). Identified QTLs accounted for 14.1-22.1% of the phenotypic variation. The MQM identified 6 QTLs affecting reproductive maturity (Table 9), dispersed on LGs 2b, 3b, 7a and 9a and explained 9.0-22.3% of the phenotypic variance.

Among these QTLs, two located on LG 2b had consistent effects on reproductive maturity. One QTL between markers SWW-583 and PVCA-173/174 was responsible for the maturity ratings of year 2012 (Fig. 4), which accounted for 12.8% of the phenotypic variation in genetic control of reproductive maturity, while the other QTL between PVCA-917/918 and PVE-775/776 showed a significant effect on maturity ratings of year 2013 (Fig. 5), and explained 11.0-22.3% of the variation. Different major QTLs identified for different year may imply an environment-related effect on expression of the QTLs. QTLs between PVGA-1727/1728 and

PVE-987/988 on LG 3b also showed effects on maturity, which explained 8.4-9.0% of the phenotypic variance. QTLs between PVAAG-2503/2505 and PVAAG-3253/3254 on LG 7a, and between PVE-49/50 and SWW-170 on LG 9a were also identified, which accounted for 12.4 and 11.8-12.2 % of the phenotypic variance.

The IM analyses identified 15 QTLs affecting reproductive maturity in the hybrid population (Table 10), on LGs 1a, 2b, 3a, 3b, 4a, 6b, 7a and 8b. Identified QTLs accounted for 10.4-27.4% of the phenotypic variation. MQM analysis identified 12 QTLs occurring on LGs 1a, 2b, 3a, 3b, 7a, 8b, 9a (Table 11), which explained 7.2-18.5% of the phenotypic variance.

Among these QTLs, those located on LG 3b had major effects on reproductive maturity. However multiple regions were identified to have associations with reproductive maturity, one region was identified between marker PVGA-1727/1728 and PVE-987/988, which accounted for 7.6-13.0% of the phenotypic variation, a second region between PVGA-1983/1984 and SWW-2922, which explained 9.9-12.7 % of the phenotypic variation. In addition, genomic region between nfsg-125 and PVE-781/782 identified on LG 2b in the hybrid population was also revealed in the selfed population, which explained 9.9 % of the phenotypic variation. QTL between PVCAG-2503/2504 and PVAAG-3253/3254 on LG 7a also occurred in both populations, which accounted for 10.9% of the phenotypic variation in the hybrid populations. Another QTL region located on LG 8b between PVGA-1275-1276 and nfsg-112 was identified in the hybrid population, accounting for 7.4-8.2% of the phenotypic variance. Other new genomic regions on LG 1a between PVE-1361/1362 and PVGA-2107/2108, and between PVGA-1513/1514 and PVCAG-2517/2518 on LG 9a were identified, each accounting for 8.7% and 18.5% of the phenotypic variation, respectively.

The QTL regions between PVGA-1727/1728 and PVGA-1201/1202 on LG 3b, between nfsg-125 and PVE-781/782 on LG 2b, and between PVCAG-2503/2504 and

PVAAG-3253/3254 on LG 7a were identified to be associated with reproductive maturity in both populations. Markers flanking these major QTL regions were then genotyped in the remaining 127 progeny of the selfed population, resulting in phenotypic and genotypic data of the whole population of 265 progeny used for QTL mapping. These common QTLs were still present in the whole selfed population (Fig. 6).

Table 2 Means and associated standard deviations in reproductive maturity ratings of two parents and two populations

Dataset	NL94 (P1)	SL93(P2)	Hybrid Pop.	Selfed Pop.
STW12-1[a]	2.0±0.0[b]	1.3±0.6	1.9±0.9	3.1±0.9
STW12-2	3.0±0.0	1.7±0.6	2.5±1.2	3.6±1.0
STW13-1	2.0±0.0	1.7±0.6	1.6±0.5	2.6±0.6
STW13-2	3.7±0.6	3.0±0.0	3.0±0.6	4.0±0.7
PKS12	5.3±0.6	4.0±1.0	4.6±1.3	4.0±1.0
PKS13-1	3.0±0.0	2.0±0.0	2.3±0.6	2.8±0.7
PKS13-2	4.7±0.6	3.0±0.0	3.2±0.6	4.0±0.7

[a] STW12-1 stands for the first time reproductive maturity ratings in Stillwater, OK in 2012

[b] Reproductive maturity rating scale given as 1 through 7, being from boot stage to maximum flowering, respectively.

Table 3 P values in the ANOVA analyses for reproductive maturity in two lowland switchgrass populations in two years

Populations	Factor	2012		2013	
		first[d]	second	first	second
Hybrid[c]	Plant ID	0.0009	<.0001	<.0001	<.0001
	Location	<.0001	0.1539	<.0001	<.0001
	ID*Location	0.0026	0.4060	0.1568	<.0001
Self	Plant ID	0.0105	0.0005	<.0001	<.0001
	Location	0.0090	0.0163	0.0008	<.0001
	ID*Location	0.0003	0.9938	<.0001	<.0001

[c] Hybrid stands for the hybrid population

[d] first stands for the first time data collection in each population.

Table 4 Broad sense heritability estimates for reproductive maturity based on components of genetic variance in hybrid and selfed lowland switchgrass populations

Pop.	Dataset	Variance component estimates			Broad sense heritability
		$\hat{\sigma}^2_{g}$	$\hat{\sigma}^2_{ge}$	$\hat{\sigma}^2_{error}$	h^2
Hybrid	STW12-1	1.022641	0.747186	10.575758	0.0828
	STW12-2	1.756017	1.238225	1.643939	0.3786
	STW13-1	0.468561	0.204794	2.160985	0.1653
	STW13-2	0.584665	0.196418	0.960227	0.3358
	PKS12	2.461299	1.288799	3.126894	0.3579
	PKS13-1	0.752803	0.156894	0.876894	0.4214
	PKS13-2	0.563622	0.177031	0.108052	0.6641
Self	STW12-1	1.077239	0.671448	36.792805	0.0280
	STW12-2	1.381548	0.859778	7.572808	0.1408
	STW13-1	0.835126	0.219942	3.231773	0.1948
	STW13-2	0.923278	0.252622	1.151519	0.3967
	PKS12	1.473256	0.810811	8.407596	0.1378
	PKS13-1	0.839928	0.252183	0.647220	0.4829
	PKS13-2	0.852950	0.287930	3.404528	0.1877

Table 5 Phenotypic correlation coefficients and associated probability values among reproductive maturity ratings in a hybrid population of lowland switchgrass

Trails	STW12-2	STW13-1	STW13-2	PKS12	PKS13-1	PKS13-2
STW12-1	0.81	0.49	0.46	0.26	0.37	0.37
	<.0001	<.0001	<.0001	0.0006	<.0001	<.0001
STW12-2		0.44	0.44	0.26	0.26	0.34
		<.0001	<.0001	0.0005	0.0005	<.0001
STW13-1			0.69	0.21	0.65	0.55
			<.0001	0.0054	<.0001	<.0001
STW13-2				0.23	0.59	0.61
				0.0020	<.0001	<.0001
PKS12					0.21	0.34
					0.0054	<.0001
PKS13-1						0.70
						<.0001

Table 6 Phenotypic correlation coefficients and associated probability values among reproductive maturity ratings in a selfed population of lowland switchgrass genotype NL94.

Trails	STW12-2	STW13-1	STW13-2	PKS12	PKS13-1	PKS13-2
STW12-1	0.78	0.47	0.37	0.49	0.40	0.26
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
STW12-2		0.45	0.51	0.58	0.32	0.39
		<.0001	<.0001	<.0001	<.0001	<.0001
STW13-1			0.69	0.30	0.77	0.50
			<.0001	<.0001	<.0001	<.0001
STW13-2				0.36	0.55	0.57
				<.0001	<.0001	<.0001
PKS12					0.44	0.34
					<.0001	<.0001
PKS13-1						0.58
						<.0001

Table 7 Simple sequence repeat marker loci and recombination distances of linkage groups in a full-sib population of NL94 and SL93 lowland switchgrass genotypes.

Linkage group	No. of loci on map	Total length (cM)	Average distance (cM)	Number of gaps > 15 cM	Dominant markers
1a	11	91	8.3	1	1
1b	8	82	10.3	2	0
2a	11	88	8.0	1	0
2b	17	79	4.6	2	0
3a	8	97	12.1	3	0
3b	22	99	4.5	0	0
4a&b	17	66	3.9	1	0
5a	15	70	4.7	0	0
5b	11	83	7.5	2	0
6a	6	49	8.2	1	0
6b	8	64	8.0	2	0
7a	11	43	3.9	0	0
7b	3	9	3.0	0	0
8a	3	7	2.3	0	0
8b	8	56	7.0	0	1
9a	9	52	5.8	0	0
9b	10	45	4.5	1	2
Total	178	1080		16	4
Average	10.5	63.5	6.1	0.9	

Table 8 QTLs identified using Interval Mapping (IM) for reproductive maturity in a selfed population of NL94 lowland switchgrass

Dataset	Linkage group	LOD peak	Position of LOD peak (cM)	Left-Right Locus		% phenotypic variance explained
STW12-1	2b	4.74	68.999	SWW-583	PVCA-173/174	14.6
STW13-1	2b	6.83	31.952	SWW-2501	SWW-1622	20.4
	2b	7.43	49.228	PVCA-917/918	PVE-225/226	22.0
	7a	4.56	37.561	PVAAG-3051/3052	SWW-2532	14.1
STW13-2	2b	5.94	65.348	SWW-583	PVE-1143/1144	18.0
	7a	5.26	34.817	PVAAG-3051/3052	PVAAG-3253/3254	16.1
PKS13-1	2b	6.65	31.952	SWW-2501	SWW-1622	20.3
	2b	7.33	49.228	PVCA-917/918	PVE-225/226	22.1
	7a	4.52	36.817	PVAAG-3051/3052	SWW-2532	14.3
PKS13-2	2b	7.22	50.155	PVCA-917/918	PVE-225/226	21.8

Table 9 QTLs identified using Multiple QTL Mapping (MQM) for reproductive maturity in a selfed population of NL94 lowland switchgrass

Dataset	Linkage group	LOD peak	Position of LOD peak (cM)	Left-Right Locus		% phenotypic variance explained
STW12-1	2b	4.11	66.999	SWW-389	PVCA-173/174	12.8
	3b	4.03	40.703	PVGA-1197/1198	PVGA-1201/1202	11.0
STW13-1	2b	8.64	49.228	PVCA-917/918	PVE-775/776	22.3
	3b	4.03	59.323	PVGA-1727/1728	PVE-987/988	8.4
	9a	4.93	94.004	PVE-49/50	SWW-170	11.8
STW13-2	2b	4.14	49.228	PVCA-917/918	PVE-775/776	11.0
	7a	4.67	32.817	PVAAG-2503/2504	PVAAG-3253/3254	12.4
PKS13-1	2b	8.54	49.228	PVCA-917/918	PVE-775/776	22.3
	3b	4.32	59.323	PVGA-1727/1728	PVE-987/988	9.0
	9a	5.00	93.004	PVE-49/50	SWW-170	12.2
PKS13-2	2b	7.22	50.155	PVCA-917/918	PVE-775/776	21.8

Table 10 QTLs identified using Interval Mapping (IM) for reproductive maturity in a hybrid population of NL94 × SL93 lowland switchgrass parents

Dataset	Linkage Group	LOD peak	Position of LOD peak (cM)	Left-Right Locus		% phenotypic variance explained
STW12-1	3b	3.64	34.417	PVGA-1201/1202	PVE-987/988	13.0
STW13-1	1a	5.22	29.253	PVGA-1253/1254	SWW-606	27.4
	2b	3.95	19.056	nfsg-125	SWW-83M	13.3
	2b	3.79	34.473	SWW-2501	nfsg-09	12.8
	3b	3.33	23.334	PVGA-1957/1958	PVGA-1201/1202	13.5
	4a	3.30	18.698	PVCAG-2269/2270	SWW-1795	15.7
	8b	4.97	40.871	PVGA-1275/1276	nfsg-112	16.4
STW13-2	1a	3.18	6.000	PVCAG-2537/2538	SWW-1667	12.6
	1a	3.56	27.253	PVGA-1253/1254	SWW-606	17.4
	3b	4.38	40.052	PVGA-1957/1958	SWW-1643	16.8
	7a	3.86	3.000	PVCAG-2503/2504	PVAAG-3253/3254	16.9
	8b	4.38	40.462	PVGA-1275/1276	nfsg-112	14.3
PKS13-1	3b	3.94	48.639	SWW-1761	SWW-1643	15.1
	6b	3.23	53.838	SWW-1969	PVCA-2147/2148	18.1
	8b	3.17	35.462	PVGA-2005/2006	PVGA-1149/1150	11.4
PKS13-2	2b	3.16	27.373	PVE-1411/1412	PVE-413/414	10.4
	3a	3.33	9.691	PVAAG-3315/3316	PVCA-55/56	10.7
	3b	3.99	48.639	SWW-1761	SWW-1643	15.8

Table 11 QTLs identified using Multiple QTL Mapping (MQM) for reproductive maturity in a hybrid population of NL94 × SL93 lowland switchgrass parents

Dataset	Linkage Group	LOD peak	Position of LOD peak (cM)	Left-Right Locus		% phenotypic variance explained
STW12-1	3b	3.65	34.417	PVGA-1201/1202	PVE-987/988	13.0
STW13-1	1a	3.93	38.496	PVE-1361/1362	PVGA-2107/2108	8.7
	2b	3.31	21.414	nfsg-125	PVE-781/782	7.2
	3b	4.86	52.842	SWW-1761	SWW-2922	9.9
	7a	3.46	35.916	PVGA-2139/2140	SWW-348	11.3
	8b	3.49	40.871	PVGA-1275/1276	nfsg-112	8.2
STW13-2	3b	3.33	37.433	PVGA-1727/1728	PVE-987/988	7.6
	7a	3.64	7.000	PVCAG-2503/2504	PVAAG-3253/3254	10.9
	8b	3.22	40.871	PVGA-1275/1276	nfsg-112	7.4
PKS12	3b	3.00	37.433	PVGA-1727/1728	PVE-987/988	8.7
	9a	3.45	15.96	PVGA-1513/1514	PVCAG-2517/2518	18.5
PKS13-1	2b	3.92	17.056	nfsg-125	PVE-781/782	9.9
	3b	4.86	44.117	PVGA-1983/1984	SWW-1761	11.6
	7a	3.09	38.916	PVGA-2139/2140	SWW-348	7.7
	8b	3.59	35.462	PVGA-2005/2006	PVGA-1149/1150	9.5
PKS13-2	2b	3.59	27.373	PVE-1411/1412	SWW-1622	10.6
	3a	4.58	11.691	PVAAG-3315/3316	PVCA-55/56	13.3
	3b	4.15	45.117	PVGA-1983/1984	SWW-1761	12.7

Fig. 1 Frequency distributions for reproductive maturity ratings in a hybrid population of lowland switchgrass at two locations, Sillwater and Perkins at seven time points over two years

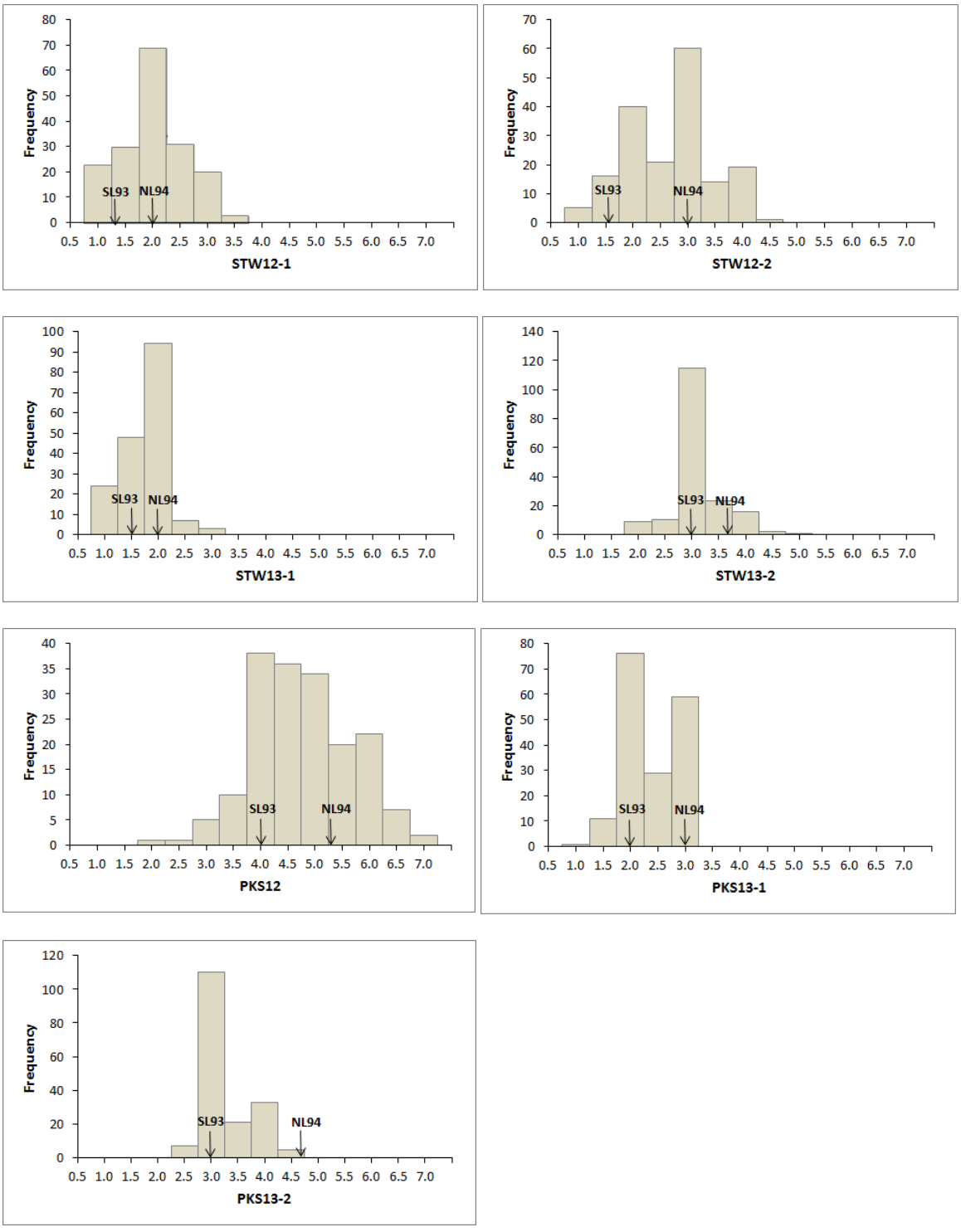


Fig. 2 Frequency distributions for reproductive maturity ratings in a selfed population of lowland switchgrass parent NL94 at two locations, Sillwater and Perkins at seven time points over two years

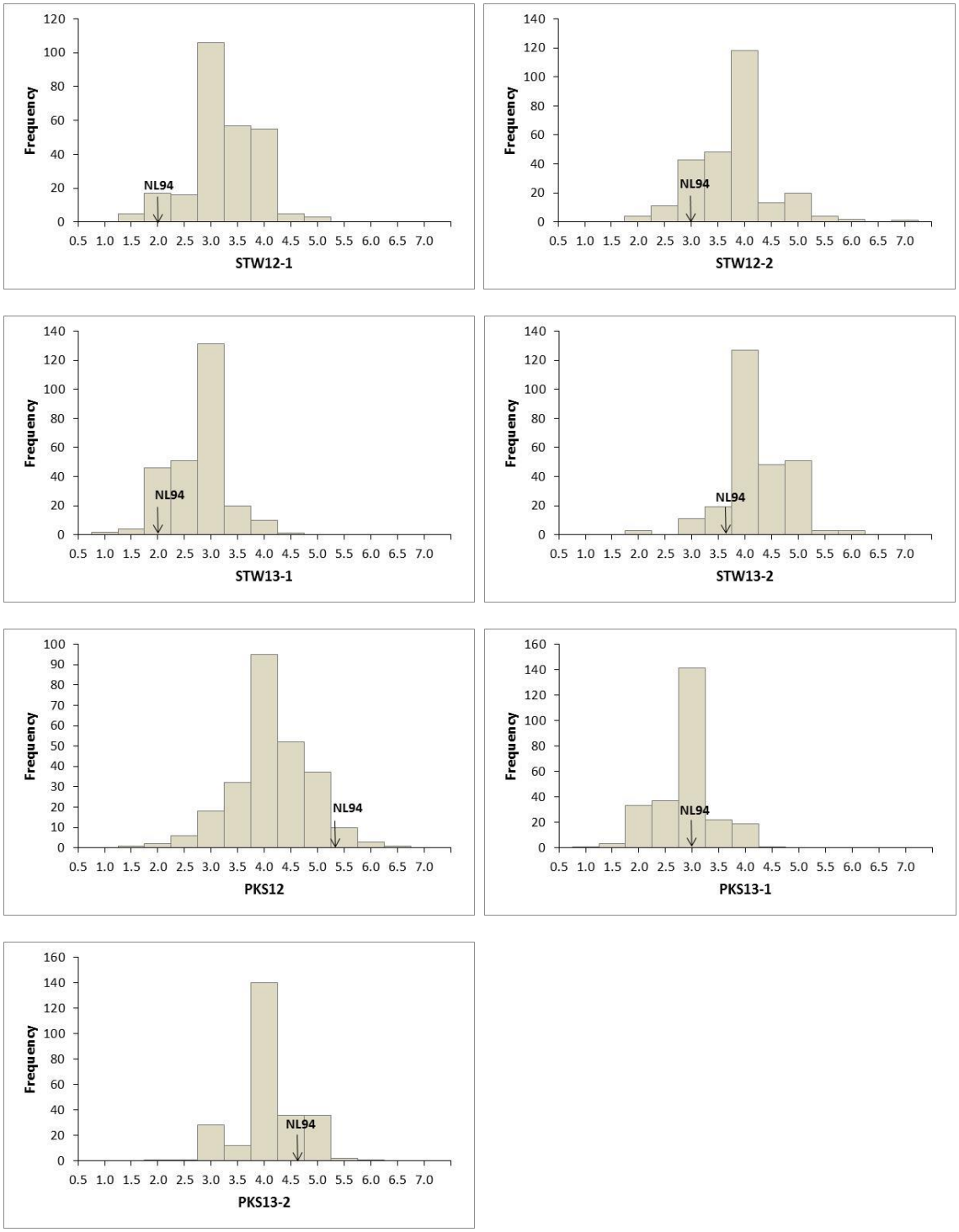
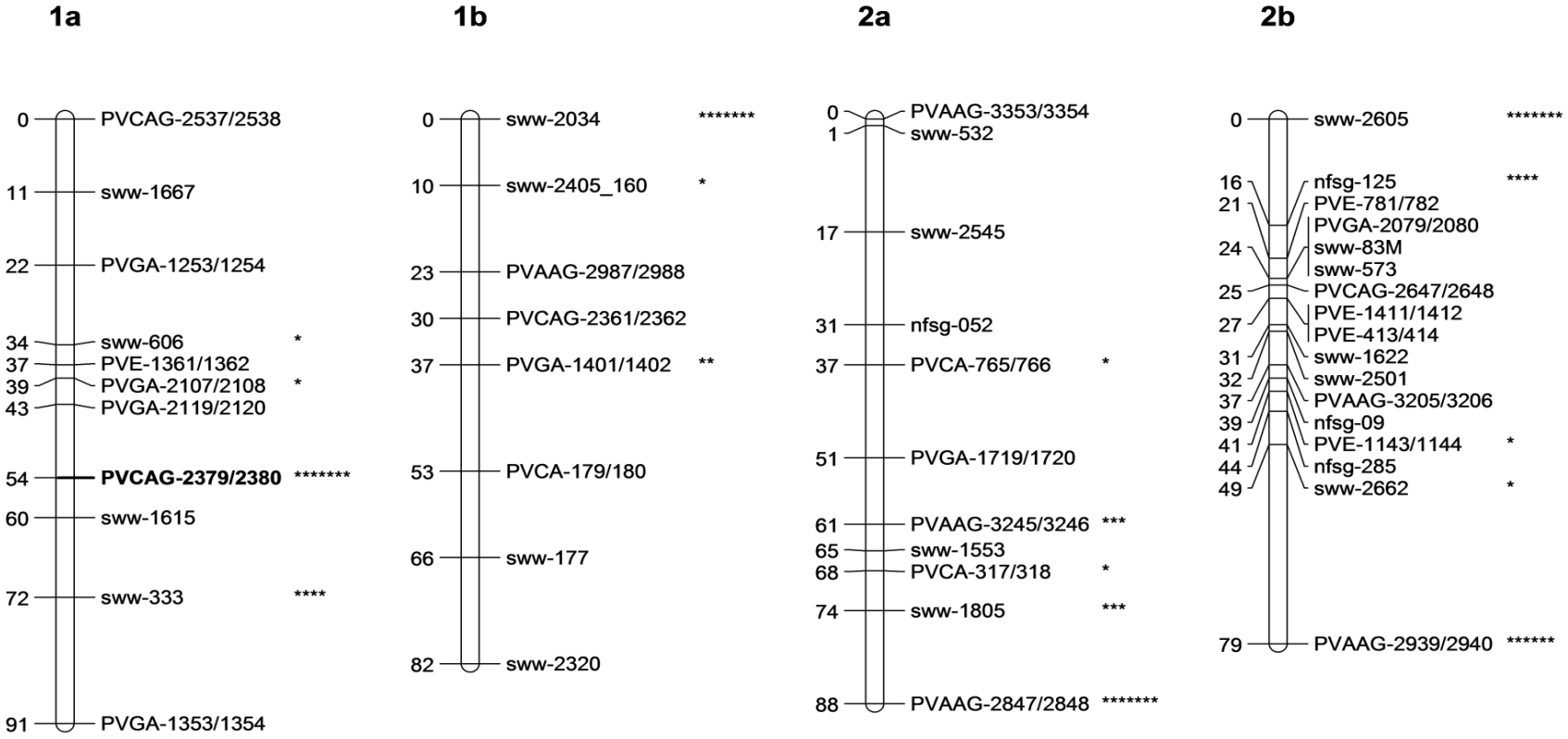
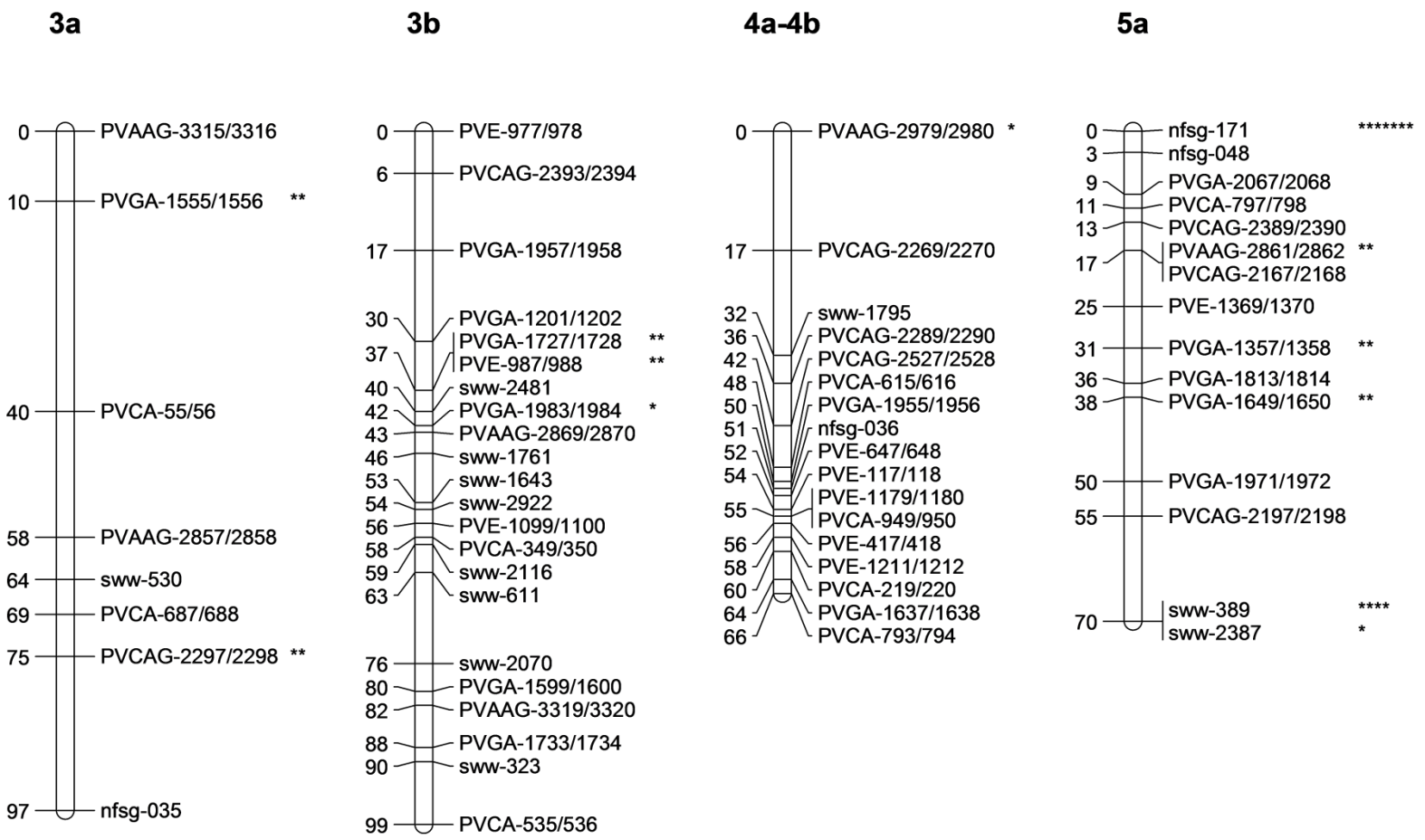
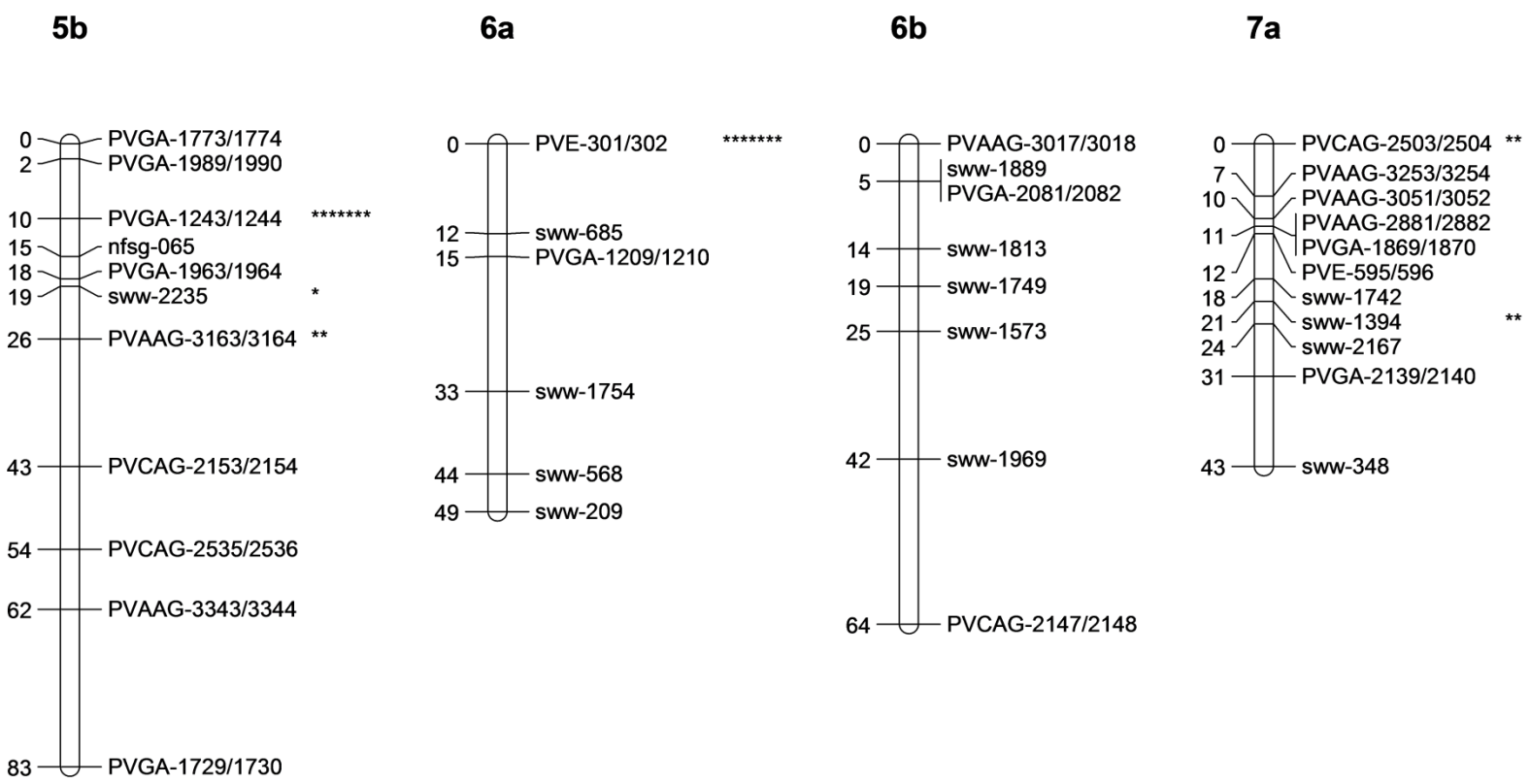


Fig. 3 A linkage map of 136 hybrid progeny derived from the cross between lowland switchgrass parents NL94 and SL93. Map distances in Kosambi map units (cM) of each linkage group are shown on the left, and marker names are shown on the right. Dominant markers are presented in bold. Segregation-distorted loci (SDL) are labeled with different significant levels: **P*<0.1, ***P*<0.05, ****P*<0.01, *****P*<0.005, ******P*<0.001, *******P*<0.0005, ********P*<0.0001




3a
0 PVAAG-3315/3316
10 PVGA-1555/1556 **
40 PVCA-55/56
58 PVAAG-2857/2858
64 sww-530
69 PVCA-687/688
75 PVCAG-2297/2298 **
97 nfsg-035
3b
0 PVE-977/978
6 PVCAG-2393/2394
17 PVGA-1957/1958
30 PVGA-1201/1202
37 PVGA-1727/1728 **
PVE-987/988 **
40 sww-2481
42 PVGA-1983/1984 *
43 PVAAG-2869/2870
46 sww-1761
53 sww-1643
54 sww-2922
56 PVE-1099/1100
58 PVCA-349/350
59 sww-2116
63 sww-611
76 sww-2070
80 PVGA-1599/1600
82 PVAAG-3319/3320
88 PVGA-1733/1734
90 sww-323
99 PVCA-535/536
4a-4b
0 PVAAG-2979/2980 *
17 PVCAG-2269/2270
32 sww-1795
36 PVCAG-2289/2290
42 PVCAG-2527/2528
48 PVCA-615/616
50 PVGA-1955/1956
51 nfsg-036
52 PVE-647/648
54 PVE-117/118
PVE-1179/1180
55 PVCA-949/950
56 PVE-417/418
58 PVE-1211/1212
60 PVCA-219/220
64 PVGA-1637/1638
66 PVCA-793/794
5a
0 nfsg-171 *******
3 nfsg-048
9 PVGA-2067/2068
11 PVCA-797/798
13 PVCAG-2389/2390
17 PVAAG-2861/2862 **
PVCAG-2167/2168
25 PVE-1369/1370
31 PVGA-1357/1358 **
36 PVGA-1813/1814
38 PVGA-1649/1650 **
50 PVGA-1971/1972
55 PVCAG-2197/2198
70 sww-389 ****
sww-2387 *


5b
0
2
10
15
18
19
26
43
54
62
83
PVGA-1773/1774
PVGA-1989/1990
PVGA-1243/1244
nfsg-065
PVGA-1963/1964
sww-2235
PVAAG-3163/3164
PVCAG-2153/2154
PVCAG-2535/2536
PVAAG-3343/3344
PVGA-1729/1730
****
*
**
6a
0
12
15
33
44
49
PVE-301/302
sww-685
PVGA-1209/1210
sww-1754
sww-568
sww-209
****
6b
0
5
14
19
25
42
64
PVAAG-3017/3018
sww-1889
PVGA-2081/2082
sww-1813
sww-1749
sww-1573
sww-1969
PVCAG-2147/2148
7a
0
7
10
11
12
18
21
24
31
43
PVCAG-2503/2504
PVAAG-3253/3254
PVAAG-3051/3052
PVAAG-2881/2882
PVGA-1869/1870
PVE-595/596
sww-1742
sww-1394
sww-2167
PVGA-2139/2140
sww-348
**
**

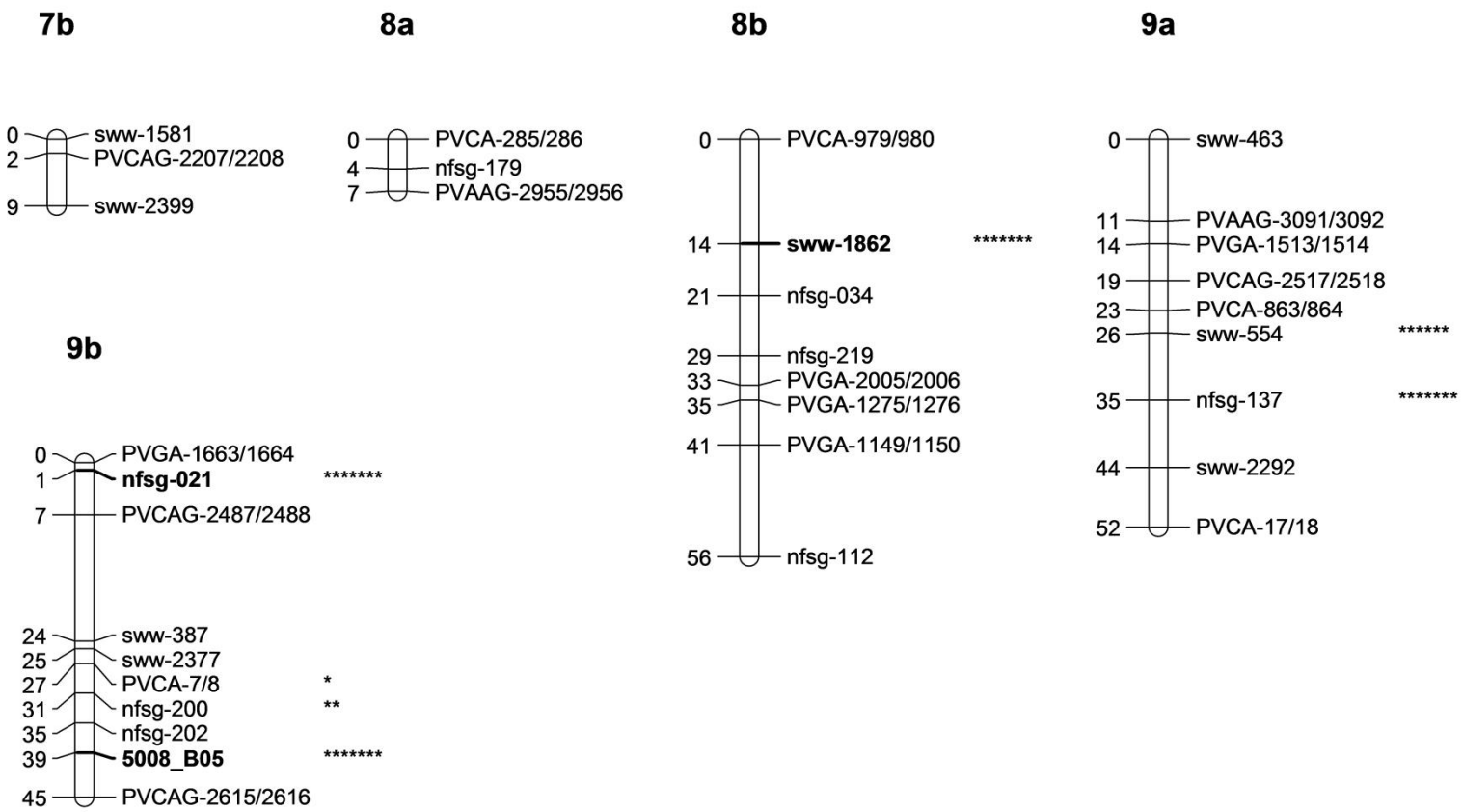
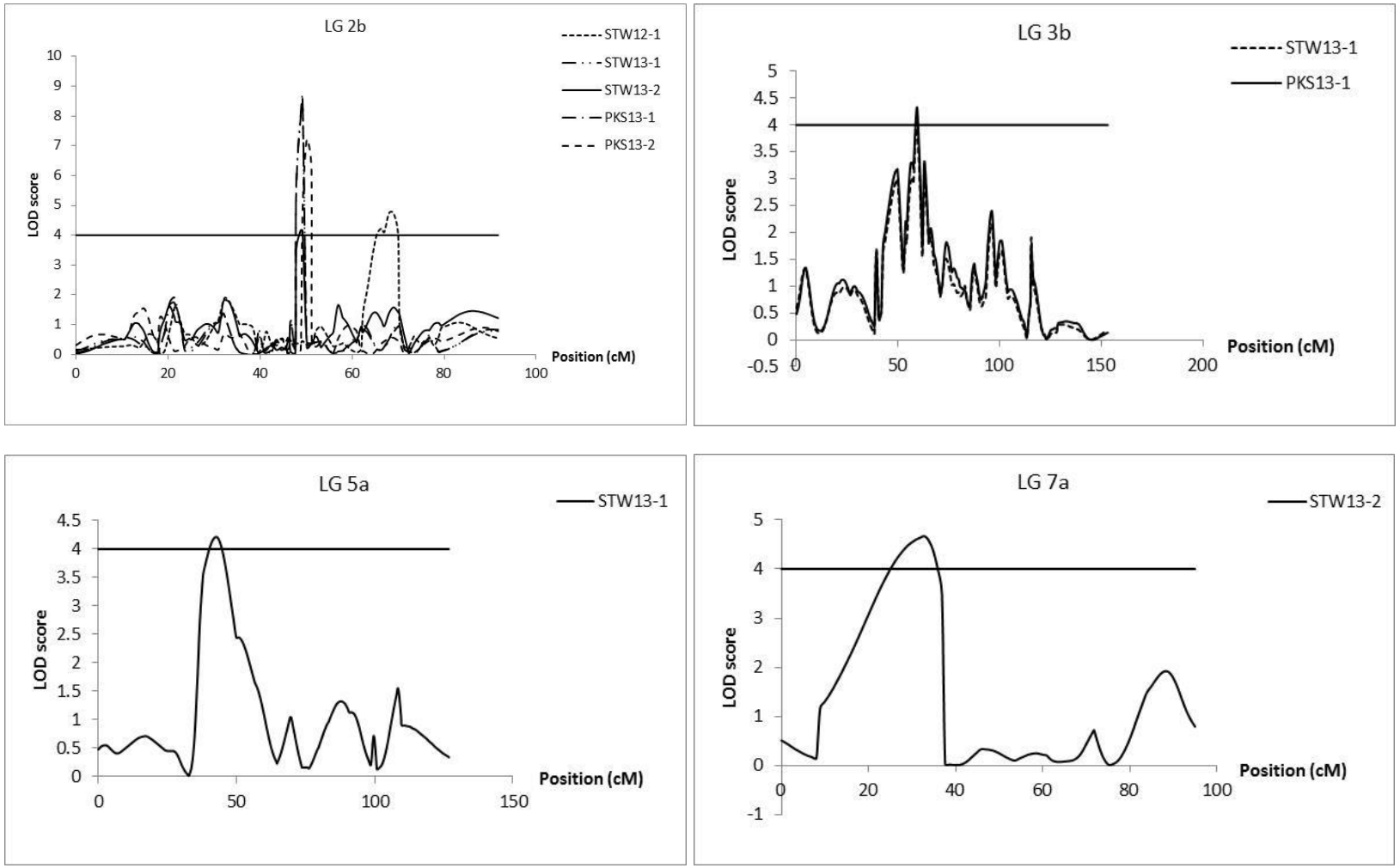

7b
0 sww-1581
2 PVCAG-2207/2208
9 sww-2399
8a
0 PVCA-285/286
4 nfsg-179
7 PVAAG-2955/2956
8b
0 PVCA-979/980
14 sww-1862 *******
21 nfsg-034
29 nfsg-219
33 PVGA-2005/2006
35 PVGA-1275/1276
41 PVGA-1149/1150
56 nfsg-112
9a
0 sww-463
11 PVAAG-3091/3092
14 PVGA-1513/1514
19 PVCAG-2517/2518
23 PVCA-863/864
26 sww-554 ******
35 nfsg-137 *******
44 sww-2292
52 PVCA-17/18
9b
0 PVGA-1663/1664
1 nfsg-021 *******
7 PVCAG-2487/2488
24 sww-387
25 sww-2377
27 PVCA-7/8 *
31 nfsg-200 **
35 nfsg-202
39 5008_B05 *******
45 PVCAG-2615/2616

Fig. 4 MQM QTL graphs for reproductive maturity in the selfed population on linkage group 2b, 3b, 5a, 7a and 9a. Horizontal line indicates the 95% significant threshold value for declaring a QTL.



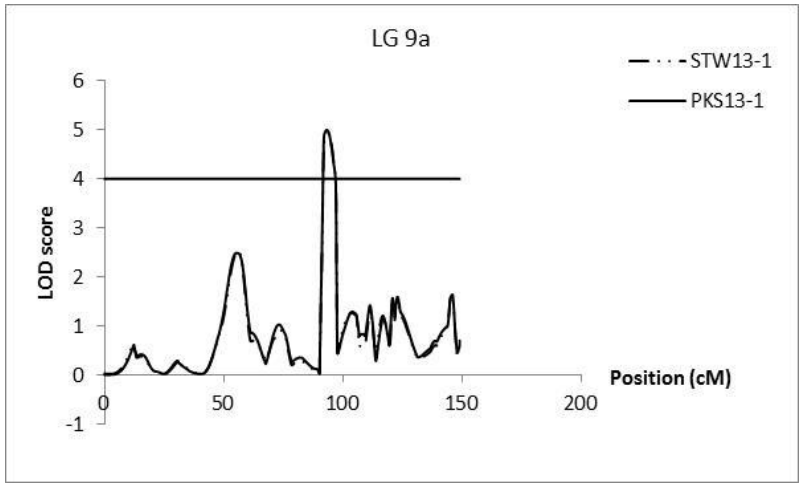
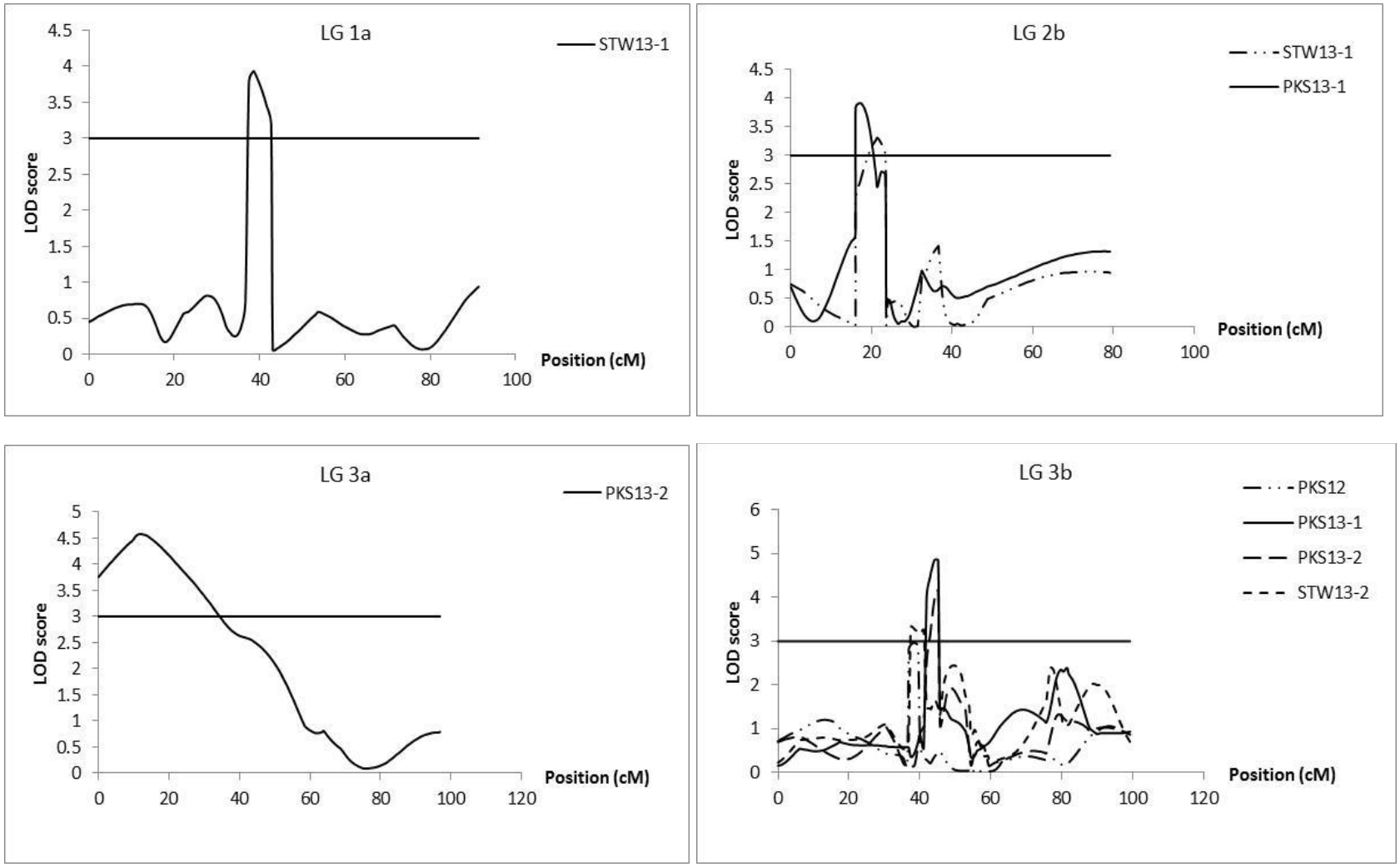
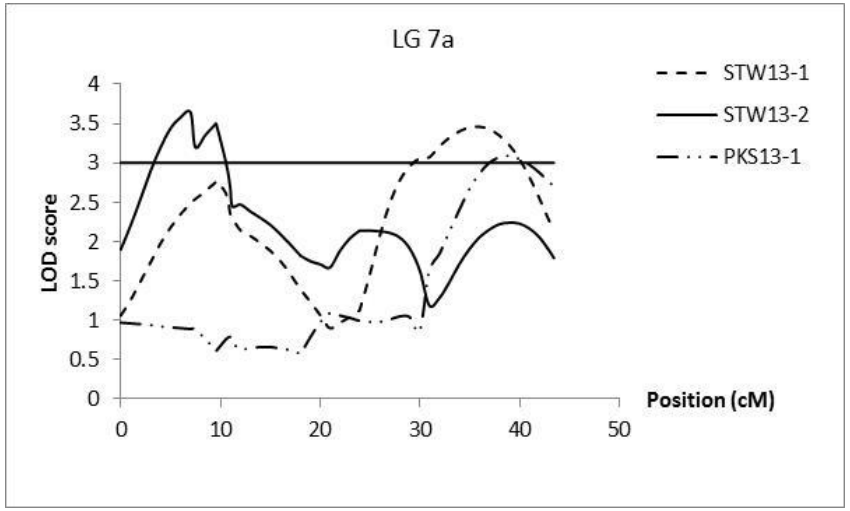
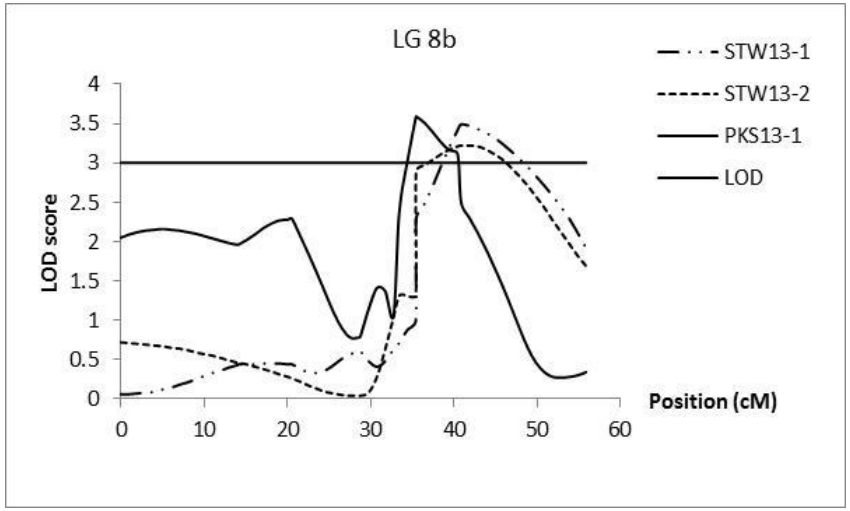
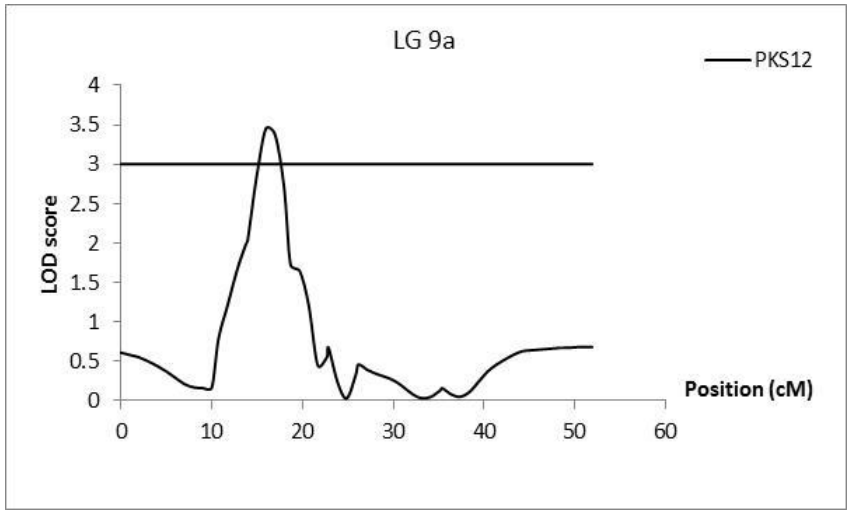

LG 9a
STW13-1
PKS13-1
LOD score
Position (cM)
6
5
4
3
2
1
0
-1
0
50
100
150
200

Fig. 5 MQM QTL graphs for reproductive maturity in the hybrid population on linkage groups 1a, 2b, 3a, 3b, 7a, 8b and 9a. Horizontal line indicates the 95% significant threshold value for declaring a QTL.




LG 7a
STW13-1
STW13-2
PKS13-1
LOD score
Position (cM)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
10
20
30
40
50


LG 8b
STW13-1
STW13-2
PKS13-1
LOD
LOD score
Position (cM)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
10
20
30
40
50
60


LG 9a
PKS12
LOD score
Position (cM)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
10
20
30
40
50
60

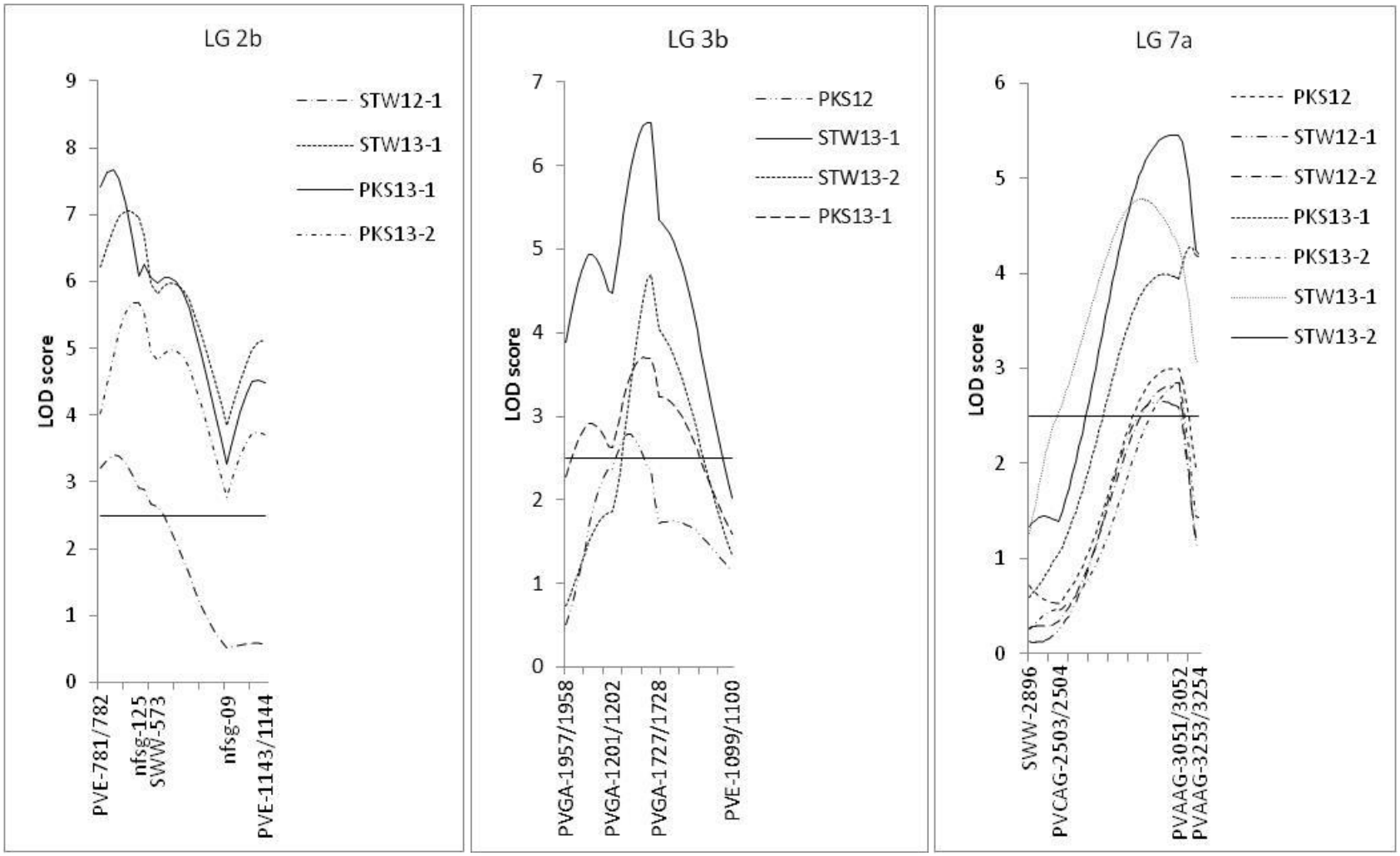


Fig. 6 LOD profiles of three major common QTLs identified on linkage groups 2b, 3b and 7a in the selfed population. Horizontal line indicates the 95% significant threshold value for declaring a QTL.

CHAPTER V

CONCLUSIONS

Reproductive maturity is a key developmental stage related to biomass yield in switchgrass. Studies on other crops like wheat, maize, rice and cotton, indicated that timing to maturity had significant associations with biomass yield and other related traits (Halloran 1977; Russell and Stuber 1983; Salam and Mackill 1993; Li et al. 2013). Mapping the QTLs for reproductive maturity in switchgrass in present study indicated reproductive maturity of lowland switchgrass was controlled by multiple genetic loci and substantially affected by environmental conditions.

Our findings indicated that broad sense heritabilities for the reproductive maturity ranged from 0.08 to 0.66 and 0.03 to 0.48 for the hybrid population and selfed population, respectively. Recent experiments indicated that heritabilities were moderate (0.47-0.70) for heading and flowering time (Bhandari et al. 2010). More recently, Price et al. (2014), using among-and-within-family selection in upland tetraploid switchgrass, revealed relatively high heritability (0.75) for flowering time overall in both selection directions. The discrepancy between previous studies and our research could be due to different genetic background and data collection methods. In Price's and Bhandari's studies, heading date and flowering date were recorded as the number of days after a certain date, while in our research, we focused on the whole reproductive maturity process, and used a numerical scale from 1 to 7 to evaluate the phenotype, which effectively rated different sub-development stages at the same time.

Linkage analysis in our study established 17 linkage groups for the hybrid population, compared with previously published switchgrass maps of selfed population (Liu et al. 2012; Liu et al. 2013), LGs 4a and 4b merged into a single group in our linkage map. A high collinearity between the hybrid population map and the published selfed population maps was observed, excepting some discrepancies described as following: inversions in marker order between our hybrid map and the published selfed population maps. The local rearrangements are common in plant genome mapping (Paterson et al. 1996), and genotyping errors could also be a reason (Johnson and Haydon 2007). In addition, compared with the established linkage map of the selfed population (Liu et al. 2013), seven markers were mapped to their homeologous linkage groups in the hybrid population map, from LG 1b to 1a, 2a to 2b, 3b to 3a, 5b to 5a, 6a to 6b.

Using MQM analysis, our research identified 6 QTLs affecting reproductive maturity in the selfed population, which were dispersed on LGs 2b, 3b, 7a, 9a and explained 9.0-22.3% of the phenotypic variance each QTL. While in the hybrid population, MQM analysis identified 12 QTLs occurring on LGs 1a, 2b, 3a, 3b, 7a, 8b, 9a, which explained 7.2-18.5% of the phenotypic variance.

Among these QTLs identified in the selfed population, two located on LG 2b had consistent effects on reproductive maturity. One QTL between markers SWW-583 and PVCA-173/174 was responsible for the maturity ratings of year 2012, which accounted for 12.8% of the phenotypic variation in genetic control of reproductive maturity, while the other QTL between PVCA-917/918 and PVE-775/776 showed a significant effect on maturity ratings of year 2013, and explained 11.0-22.3% of the variation. Different major QTLs identified for different year may imply an environment-related effect on expression of the QTLs. QTLs between PVGA-1727/1728 and PVE-987/988 on LG 3b also showed effects on maturity, which explained 8.4-9.0% of the phenotypic variance. QTLs between PVAAG-2503/2505 and PVAAG-3253/3254 on LG

7a, and between PVE-49/50 and SWW-170 on LG 9a were also identified, which accounted for 12.4 and 11.8-12.2 % of the phenotypic variance.

Among the QTLs identified in the hybrid population, these on LG 3b had major effects on reproductive maturity. However multiple regions were identified to have associations with reproductive maturity, one region was identified between marker PVGA-1727/1728 and PVE-987/988, which accounted for 7.6-13.0% of the phenotypic variation, a second region between PVGA-1983/1984 and SWW-2922, which explained 9.9-12.7 % of the phenotypic variation. In addition, genomic region between nfsg-125 and PVE-781/782 identified on LG 2b in the hybrid population was also revealed in the selfed population, which explained 9.9 % of the phenotypic variation. QTL between PVCAG-2503/2504 and PVAAG-3253/3254 on LG 7a also occurred in both populations, which accounted for 10.9% of the phenotypic variation in the hybrid populations. A new QTL region located on LG 8b between PVGA-1275-1276 and nfsg-112 was identified in the hybrid population, accounting for 7.4-8.2% of the phenotypic variance. Other genomic regions on LG 1a between PVE-1361/1362 and PVGA-2107/2108, and between PVGA-1513/1514 and PVCAG-2517/2518 on LG 9a were identified, each accounting for 8.7% and 18.5% of the phenotypic variation, respectively.

To our knowledge, no QTL mapping work has been reported in switchgrass, thus no previous information is available about how the reproductive maturity is genetically regulated in switchgrass. Recent research in foxtail millet [*Setaria italica* (L.) Beauv.], a panicoid grass closely related to switchgrass, revealed that multiple genomic regions were involved in the control of flowering time (Mauro-Herrera et al. 2013). Sixteen QTLs conditioning flowering time in foxtail millet were dispersed on LGs II, III, IV, V, VII, and VIII. The percentage of phenotypic variance explained by individual QTLs ranged from 2.5% to 41.9%. Compared with our results, common QTLs controlling reproductive maturity in both populations were identified on LGs 2b, 3b, and 7a (Mauro-Herrera et al. 2013).

The results of this study indicated that common QTL regions between PVGA-1727/1728 and PVE-987/988 on LG 3b, between nfsg-125 and PVE-781/782 on LG 2b, and between PVCAG-2503/2504 and PVAAG-3253/3254 on LG 7a identified in both populations were associated with reproductive maturity in switchgrass. Compared with the QTLs revealed in the selfed population, new QTL regions located on LG 8b between PVGA-1275-1276 and nfsg-112 was identified in the hybrid population, accounting for 7.4-8.2% of the phenotypic variance. Other new genomic regions on LG 1a between PVE-1361/1362 and PVGA-2107/2108, and between PVGA-1513/1514 and PVCAG-2517/2518 on LG 9a were identified, each accounting for 8.7% and 18.5% of the phenotypic variation, respectively. It was speculated that the extra genomic regions identified in the hybrid population could be due to parental genetic basis difference, since SL93 (♂) in the hybrid population provided one more source of genetic variation compared with that of the selfed population, which derived from one single parent NL94 (♀). These newly identified SSR markers and their chromosomal locations would facilitate further isolation of genes controlling reproductive maturity through a map-based cloning approach, and could eventually expedite the application of marker assisted selection (MAS) in switchgrass breeding programs.

REFERENCES

Beckmann JS, Soller M (1983) Restriction fragment length polymorphism in genetic improvement: methodologies, mapping and costs. Theor Appl Genet 67:35-43

Bhandari HS, Saha MC, Mascia PN, Fasoula VA, Bouton JH (2010) Variation among half-sib families and heritability for biomass yield and other traits in lowland switchgrass (*Panicum virgatum* L.). Crop Sci 50:2355-2363

Bhandari HS, Saha MC, Fasoula VA, Bouton JH (2011) Estimation of genetic parameters for biomass yield in lowland switchgrass (*Panicum virgatum* L.) Crop Sci 51:1525-1533

Boe A, Casler MD (2005) Hierarchical analysis of switchgrass morphology. Crop Sci 45:2465-2472

Bouton JH (2007) Molecular breeding of switchgrass for use as a biofuel crop. Curr Opin Genet Dev 17:553-558

Bouton JH (2008) Improvement of switchgrass as a bioenergy crop. In: Vermerris W (ed) Genetic improvement of bioenergy crops. Springer, New York, pp 309-345

Casler MD, Vogel KP, Taliaferro CM, Wynia RL (2004) Latitudinal adaptation of switchgrass populations. Crop Sci 44:293-303

Casler MD, Vogel KP, Taliaferro CM, Ehlke NJ, Berdahl JD, Brummer EC, Kallenbach RL, West CP, Mitchell RB (2007) Latitudinal and longitudinal adaptation of switchgrass populations. Crop Sci 47:2249-2260

Casler MD, Christian MT, Shawn MK, Buell CR, Wang ZY, Cao PJ, Schmutz J, Ronald P (2011) The switchgrass genome: tools and strategies. Plant Genome 4:273-282

Casler MD (2012) Switchgrass breeding, genetics, and genomics. In: A. Monti (ed) Switchgrass. Springer-Verlag, London, pp 29-53

Casler MD (2014) Heterosis and reciprocal-cross effects in tetraploid switchgrass. Crop Sci. doi: 10.2135/cropsci2013.12.0821

Cassida KA, Muir JP, Hussey MA, Read JC, Venuto BC, Ocumpaugh WR (2005) Biomass yield and stand characteristics of switchgrass in south central U.S. environments. Crop Sci 45: 673-681

Coistich DE, Friebe B, Sheehan MJ, Casler MD, Buckler ES (2010) Genome-size variation in switchgrass (*Panicum virgatum*): flow cytometry and cytology reveal rampant aneuploidy. Plant Genome 3:130-141

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissues. Focus 12:13-15

Esbroeck GA, Hussey MA, Sanderson MA (1998) Selection response and developmental basis for early and late panicle emergence in Alamo switchgrass. Crop Sci 38:342-346

Evans LT (1993) Crop evolution, adaptation and yield. Cambridge Univ. Press, Cambridge, UK

Fuentes RG, Taliaferro CM (2002) Biomass yield stability of switchgrass cultivars. In: J. Janick and A. Whipkey (ed) Trends in new crops and new uses. ASHS Press, Alexandria, VA, pp 276-282

Garland CD (2010) Growing and harvesting switchgrass for ethanol production in Tennessee. University of Tennessee Extension. Available at http://utextension.tennessee.edu/publication/Documnets/SP701-A.pdf

Halloran GM (1977) Development basis of maturity differences in spring wheat. Agron J 69:899-902

Hitchcock AS, Chase A (1951) Manual of the grasses of the United States. United States Department of Agriculture. Miscellaneous Publication No. 200. U.S. Government Printing Office, Washington

Johnson PCD, Haydon DT (2007) Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. Genetics 175:827-842

Jones MD, Brown JG (1951) Pollination cycles of some grasses in Oklahoma. Agron J 43:218-222

Li CQ, Wang XY, Dong N, Zhao HH, Xia Z, Wang R, Converse RL, Wang QL (2013) QTL analysis for early-maturing traits in cotton using two upland cotton (*Gossypium hirsutum* L.) crosses. Breed Sci 63:154-163

Liu LL, Huang YL, Punnuri S, Samuels T, Wu YY, Mahalingam R (2013) Development and introgression of EST-SSR markers into an established linkage map in switchgrass. Mol Breed 32:923-931

Liu LL, Wu YQ (2012) Identification of a selfing compatible genotype and mode of inheritance in switchgrass. Bioenergy Res 5:662-668

Liu LL, Wu YQ, Wang YW, Samuels T (2012) A high-density simple sequence repeat-based genetic linkage map of switchgrass. G3 2:357-370

Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, Buckler ES, Costich DE (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery. PLoS Genetics 9:1-14

Mart ńez O, Curnow RN (1992) Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor Appl Genet 85:480-488

Martinez-Reyna JM, Vogel KP (2002) Incompatibility systems in switchgrass. Crop Sci 42:1800-1805

Mauro-Herrera M, Wang XW, Barbier H, Brutnell TP, Devos KM, Doust AN (2013) Genetic control and comparative genomic analysis of flowering time in Setaria (Poaceae). G3 3:283-295

McLaughlin SB, Bouton J, Branshy D, Conger BV, Ocumpaugh WR, Parrish DJ, Taliaferro C, Vogel KP, Wullschleger SD (1999) Developing switchgrass as a bioenergy crop. In: Janick J (ed) Perspectives on new crops and new uses. ASHS Press, Alexandria, VA, pp 282-299

Moore KJ, Moser LE, Vogel KP, Waller SS, Johnson BE, Pedersen JF (1991) Describing and quantifying growth stages of perennial forage grasses. Agron J 83:1073-1077

Newell LC (1968) Effects of strain source and management practice on forage yields of two warm-season prairie grasses. Crop Sci 8:205-210

Okada M, Lanzatella C, Saha MC, Bouton J, Wu RL, Tobias CM (2010) Complete switchgrass genetic maps reveal subgenome collinearity, preferential pairing and multilocus interactions. Genetics 185:745-760

Parrish DJ, Fike JH, Bransby DI, Samson R (2008) Establishing and managing switchgrass as an energy crop. Forage and Grazinglands doi:10.1094/FG-2008-0220-01-RV.

Paterson AH, Lan TH, Reischmann KP, Chang C, Lin YR (1996) Toward a unified genetic map of higher plants, transcending the monocot-dicot divergence. Nat Genet 14:380-382

Patzek TW, Anti SM, Campos R, Ha K, Lee J, Li B, Padnick J, Yee SA (2005) Ethanol from corn: clean renewable fuel for the future, or drain on our resources and pockets? Environment, Development and Sustainability 7:319-336

Petrulis M, Hines F, Sommer J (1993) Ethanol production and employment. U.S. Dept. of Agriculture, Economic Research Service. Agric. Info. Bull. 678.

Porter C (1966) An analysis of variation between upland and lowland switchgrass, *Panicum virgatum* L. Ecology 47:980-992

Price DL, Casler MD (2014) Inheritance of secondary morphological traits for among-and-within-family selection in upland tetraploid switchgrass. Crop Sci 54:1-8

Russell WK, Stuber CW (1983) Effects of temperature and photoperiod on the duration of vegetative growth in maize. Crop Sci 23:847-850

Salam MA, Mackill DJ (1993) Effects of a photoperiod sensitivity gene on morphological characters in rice. Field Crop Res 33:261-268

Samson R, Mani S, Robert B, Shahab S, Diego Q, Segundo Q, Veronica R, Claudia HL (2005) The potential of C4 perennial grasses for developing a global BIOHEAT industry. Plant Sci 24:461-495

Sanderson MA (1992) Morphological development of switchgrass and kleingrass. Agron J 84:415-419

Sanderson MA, Reed RL, Ocumpaugh WR, Hussey MA, Van Esbroeck G, Read JC, Tischler CR, Hons FM (1999) Switchgrass cultivars and germplasm for biomass feedstock production in Texas. Bioresource Tech 67:209-219

Schmer MR, Mitchell RB, Perrin RK (2008) Net energy of cellulosic ethanol from switchgrass. Proc. Nat. Acad. Sci. USA 1050:464-469

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3:739-744

Talbert LE, Timothy DH, Burns JC, Rawlings JO, Moll RH (1983) Estimates of genetic parameters in switchgrass. Crop Sci 23:725-728

Taliaferro CM, Vogel KP, Bouton JH, Mclaughlin SB, Tuskan GA (1999) Reproductive characteristics and breeding improvement potential of switchgrass. In: Overend R and Chornet E (ed). Proceedings of the 4th biomass conference of the America biomass, a growth opportunity in green energy and value-added products. Elsevier Sciences, Oxford, UK. pp147-153

U.S. Department of Energy (2011) *U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry*. R.D. Perlack and B.J. Stokes (Leads), ORNL/TM-2011/224. Oak Ridge National Laboratory, Oak Ridge, TN. 227p

Van Esbroeck GA, Hussey MA, Sanderson MA (1998) Selection response and developmental basis for early and late panicle emergence in Alamo switchgrass. Crop Sci 38:342-346

Van Ooijen JW (2006) JoinMap 4: software for the calculation of genetic linkage maps in experimental populations. Kyazma B. V, Wageningen

Varvel GE, Vogel KP, Mitchell RB, Follett R, Kimble J (2008) Comparison of corn and switchgrass on marginal soils for bioenergy. Biomass and Bioenergy 32:18-21

Voorrips RE (2002) Mapchart: Software for the graphical presentation of linkage maps and QTLs. J Hered 93(1):77-78

Wallace DH, Baudoin JP, Beaver J, Coyne DP, Halseth DE, Masaya PN, Munger HM, Myers JR, Silbernagel M, Yourstone KS, Zobel RW (1993) Improving efficiency of breeding for higher crop yield. Theor Appl Genet 86:27-40

Wang YW, Samuels TD, Wu YQ (2011) Development of 10,30 genomic SSR markers in switchgrass. Theor Appl Genet 122:677-686

Wright L, Turhollow A (2010) Switchgrass selection as a 'model' bioenergy crop: A history of the process. Biomass Bioenergy 34:851-868

Wu YQ (2014) Classic genetics and breeding for bioenergy traits. In: Luo, H. and Y.Q. Wu (ed) Compendium of bioenergy plants: switchgrass. Science publishers, Inc. jointly with CRC Press of Taylor and Francis Group. pp 170-188.

Wu YQ, Huang YH (2008) QTL mapping of sorghum resistance to greenbugs by molecular markers. Theor Appl Genet 117:117-124

VITA

HONGXU DONG

Candidate for the Degree of

Master of Science

Thesis: QTL MAPPING FOR REPRODUCTIVE MATURITY IN LOWLAND SWICTHGRASS POPULATIONS

Major Field: Plant and Soil Sciences

Biographical:

Education:

Completed the requirements for the Bachelor of Science in Agronomy at Shandong Agricultural University, Taian, Shandong, China in 2012.

Experience:
Employed by Oklahoma State University, Department of Plant and Soil Sciences as a Graduate Research Assistant (August 2012 to present)

Professional Memberships:
American Society of Agronomy, since 2012
Crop Science Society of America, since 2012
Soil Science Society of America, since 2012