THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Doctor of Philosophy (Agricultural Biotechnology)
DEGREE

Agricultural Biotechnology
FIELD

Interdisciplinary Graduate Program
PROGRAM

TITLE: cDNA-AFLP of Flowering Stage in Sugarcane and Colinearity with Sorghum [*Sorghum bicolor* (L.) Moench]

NAME: Miss Pattama Srinamngoen

THIS THESIS HAS BEEN ACCEPTED BY

.....
THESIS ADVISOR
(Associate Professor Sonthichai Chanprame, Ph.D.)

.....
THESIS CO-ADVISOR
(Nongluk Teinseree, Ph.D.)

.....
THESIS CO-ADVISOR
(Professor Ismail Dweikat, Ph.D.)

.....
GRADUATE COMMITTEE CHAIRMAN
(Assistant Professor Sermsiri Chanprame, Ph.D.)

APPROVED BY THE GRADUATE SCHOOL ON .....

.....
DEAN
(Associate Professor Gunjana Theeragool, D.Agr.)

THESIS

cDNA-AFLP of Flowering Stage in Sugarcane and Colinearlity with Sorghum [*Sorghum bicolor* (L.) Moench]

PATTAMA SRINAMNGOEN

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy (Agricultural Biotechnology)
Graduate School, Kasetsart University
2014

Pattama Srinamngoen 2014: cDNA-AFLP of Flowering Stage in Sugarcane and Colinearlity with Sorghum [*Sorghum bicolor* (L.) Moench]. Doctor of Philosophy (Agricultural Biotechnology), Major Field: Agricultural Biotechnology, Interdisciplinary Graduate Program. Thesis Advisor: Associate Professor Sontichai Chanprame, Ph.D 124 pages.

Sugarcane (*Saccharum* spp.) and Sorghum (*Sorghum* spp.) have become an increasingly important crop for bioengery such as biofuel. Sugarcane has an autopolyploid complex genome whereas sorghum has a diploid simple genome. Flowering is one of sugar-related agronomic traits in both species. Here, we obtained cDNA of 0-15 cm long inflorescence of *S. spontaneum* using cDNA-amplified restriction fragment length polymorphism (cDNA-AFLP) to develop flower transcriptome profiling with 26 primer combinations. A total of 183 transcript-derived fragments (TDFs) were screened and 96 TDFs were sequenced. Out of 96, 26 TDFs were selected as flowering putative genes to study colinerity with sorghum genome. For colinearity of flowering putative genes, a genetic map with 169 SSR co-dominant SSR markers and 9 TDFs marker loci were conducted on 14 linkage groups collectively spanning 1077.8 cM that corresponding the 10 sorghum chromosomes. Interestingly, nine TDFs marker loci can be mapped into 5 linkage groups. In this study, we successfully identify the homologous location of sugarcane flowering TDFs in sorghum genome. Moreover, sfw4DS.1X TDF could be a part of gene that related to flowering and showed codominant expression in sorghum RILs population. This means it carries genotypic value of both parents and can be served as a candidate specific marker for breeding selection, while sfw2DS.3E showed dominant expression that good enough for breeding program as well.

_____ _____ ____/____/____

Student's signature Thesis Advisor's signature

ACKNOWLEDGEMENTS

I would like to express my very great appreciation to my supervisor, Associate Professor Dr. Sontichai Chanprame, Professor Dr. Ismail Dweikat and Dr. Nongluk Teinsaree for valuable advice, supervision, kindly suggestions, comments and assistance which were essential for the completion of this thesis.

I would like to thank the Plant Tissue Culture and Transformation Laboratory, Center for Agricultural Biotechnology Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand and Crop Sciecne Laboratory, University of Nebraska-Lincoln, USA for their facilitating the instruments and materials during the research

I would like to thank to the National Science and Technology Development Agency, Ministry of Science and Technology for their financial support granted through doctoral fellowship. I also thank to Center for Agricultural Biotechnology, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom and Center of Excellence on Agricultural Biotechnology, Science and Technology. Postgraduate Education and Research Development Office, Commission on Higher Education, Ministry of Education. (AG-BIO/PERDO-CHE) Thailand, for research fund.

In addition, I would like to thank Associate Professor Dr. Julapark Chunwong for providing the mapping statistical software for data analysis. I also thank to Mrs. Warunee Somanus, Mr. Somphong Chankaew, Mr. John Rajewski, Mr. Nigel La Borde and Mr. Suvatchai Klinsong for helpful me throughout my research.

Finally, I express my special appreciation to my family whose support everything throughout my study.

Pattama Srinamngoen

March 2014

TABLE OF CONTENTS

Page

LIST OF TABLES

APPENDIX TABLES

APPENDIX TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

APPENDIX FIGURES

APPENDIX FIGURES (Continued)

LIST OF ABBREVIATIONS

BLAST	=	Basic local alignment search tool
bp	=	Base pairs
cDNA	=	Complementary DNA
DEPC	=	Diethyl pyrocarbonate
DNA	=	Deoxyribonucleic acid
DNase	=	Deoxyribonuclease
dNTPs	=	Deoxynucleotide triphosphate (s)
EDTA	=	Ethylene diamine tetraacetic acid
NCBI	=	National Center for Biotechnological Information
PCR	=	Polymerase chain reaction
PVPP	=	Polyvinyl polypyrolidone
RNA	=	Ribonucleic acid
SDS	=	Sodium dodecyl sulfate
TF	=	Transcription factor
TDF	=	Transcript-derived fragment
RIL	=	Recombinant inbred line

cDNA-AFLP of Flowering Stage in Sugarcane and Colinearlity with Sorghum [*Sorghum bicolor* (L.) Moench]

INTRODUCTION

Sugarcane, *Saccharum officinarum* is one of the major economic crops in Thailand. The total planted area of sugarcane in Thailand is the 5th biggest in the world from Brazil, India, Cuba and China. The major cultivation areas in Thailand are in Karnjanaburi, Supanburi, Udonthani and Chaiyaphoom province. In Thailand, sugarcane is harvested from November to March. During the 2008/2009 production year, a record of 66.5 million tons of cane was harvested and 7.1 million tons of sugar produced. Thailand was the 2nd exporter of sugarcane; total export value was more than 80,000 million Baht. The main export markets are in Japan, South Korea, Indonesia and Malaysia.

Sorghum (*Sorghum bicolor* L. Moench) is a one of the five most important cereal crops grown worldwide. Traditionally, it is grown for food and fodder by subsistence farmers, based on its ability to adapt to a wide range of environmental conditions and low inputs needed for cultivation. More recently sorghum has been seen as a potential alternative for bio-energy production of fuel-grade ethanol.

In Thailand, most of sugar is produced from sugarcane. This is referring to sucrose content that accumulated in sugarcane stem. The classical problem of sugarcane plantation is the flowering. Because when sugarcane makes the transition turn vegetative stage to the reproductive stage, the assimilated carbon is shunted to flower development and seed production. So, sucrose content that accumulates in its stalks decreased immediately. The study to understand or identify of gene(s) that involved with flowering trait in sugarcane is extremely valuable, and sorghum has a small genome and related to sugarcane. So, sorghum genome will be as an important tool for studying molecular analysis of the complex sugarcane genome (Paterson, 2008).

OBJECTIVES

1. To identify the putative gene(s) affecting in flowering trait in sugarcane by cDNA-AFLP technique.

2. To conduct the genetic linkage groups and QTL mapping that linked to flowering gene(s) in sorghum by simple sequences repeat (SSR) marker.

3. To study colinearity about flowering gene(s) between sorghum and sugarcane.

LITERATURE REVIEW

1. Sugarcane

Sugarcane is an excellent example of crop with a complex genome. The term of "*Saccharum* complex" is used to refer to group of closely related genera including *Saccharum, Narenga, Sclerostachya, Erianthus* and *Miscanthus.*

Saccharum is characterized by a high chromosome number and high level of polyploidy. Saccharum comprises of six species by euploidy and aneuploidy series. *S. spontaneum* and *S. robustum* are considered to represent the basic species as found in the nature and are proposed the ancestor of *S. officinarum*. *S. officinarum* is a modern cultivar, and contains large amounts of sucrose. *S. barberi* and *S. sinensis* are cultivated species found in India and China, respectively. And, *S. edule* is a partial sterile and grown in Melanesia.

Grivet and Arruda (2002) reported that *Saccharum* has about 10 Gbp of genome size and has a giant chromosome. *S. officinarum* has a basic chromosome number of x= 10, indicating that these plants are octaploid and *S. spontaneum* has a basic chromosome number of x=8, indicating that the ploidy level of this species is between 5 and 16 (Grivet and Arruda, 2001; Cuadrado *et al*. 2004). The sugarcane includes 3 main group species;

1. Early cultivar, the noble cane or tropical canes belonging to *S. officinarum* L., the chromosome number is 2n=80, *S. barberi* or the North India cane, thinner and pourer in sugar content, and the chromosome number is 2n=81-124 while *S. sinensis* or Chinese canes and the *Pansahi* group of Indian canes, which are somewhat similar in appearance to the North Indian canes with 2n=116-120. For this group of species, it plays the most important role species in sugarcane breeding program.

2. Wild species, *S. spontaneum* or wild tropical cane with 2n=40-128 chromosome and *S. robustum* or wild cane New Guinea with 2n=60, 80-200 chromosome.

3. Marginal species, *S. edule* which is 2n = 60-122 chromosome number.

Gene control flowering of sugarcane

Under certain photoperiod and soil moisture conditions, sugarcane changes from the vegetative to reproductive stage. This means the growing point ceases forming leaf primordia and starts the production of an inflorescence. The inflorescence, or tassel, of sugarcane is an open-branched panicle. Each tassle consists of several thousand tiny flowers, each capable of producing one seed. The seeds are extremely small and weigh approximately 250 seed per gram or 113,500 seed per pound (http://edis.ifas.ufl.edu/sc034).

Flowering in sugarcane is a complex physiological process related to sucrose content, specific genetic and environmental requirements (Julien, 1972). Coelho *et al*. (2013) presented the *in silico* model analysis by searching the SUCEST database for putative orthologs for flowering times gene of sugarcane under photoperiodic control. A searching found 5 flowering time genes; *GI* (*Gigantea*), *CO* (*Constans*), *EHD1* (*Early heading date1*), *GHD7* (*Grain number*, *Plant Height*, and *Heading Date7*), and *FT* (*Flowering locusT*) and the result of sequence comparison showed significant similarity to flowering time genes of other species. In addition, all 5 flowering time genes, they also need their own mechanisms to make the floral induction signals (Colasanti and Coneva, 2009), so the sugarcane genetic control is still unclear.

2. Sorghum

Sorghum, *Sorghum bicolor* (L.) Moench is the world's fifth most important cereal crop and cultivated for food and sustainable energy. It is a C_4 grass with high photosynthetic efficiency and high productivity. It originated in the northeast quadrant

of and distributed widely throughout tropical and subtropical (Teshome *et al*., 2007). Within the species, *S. bicolor,* which is characterized by a diploid containing ten chromosomes pairs (2n=20) and inter fertile (Curtis, 1968). The small genome of sorghum (730 Mb), therefore, has been attractive model for genetics understanding or representative of C_4 grasses (Zeller, 2000).

Sorghum has many advantages, normally for human consumption and animal feed. Sorghum has been recognized as a particular potential crop, and will be used for supporting the increasing of the world's population (Farrell *et al*., 2006). Moreover, sorghum has become an increasingly important crop for advance biofuel production. Although, most ethanol produced in the U.S.A. is currently from the starch of maize, but sorghum can be produced at less cost than maize (Smith and Buxton, 1993). So, many countries use sorghum as a source of ethanol production such as China, India, Philippines and Australia.

The limitations of sorghum production and its productivity are diseases, insects and environmental constraints. About insect pest, almost 150 species are reported as a pest of sorghum worldwide (Sharma, 1993). Of which, the chinch bug, *Blissus leucopterus* (Say) (Hemiptere: Blissidae) is common and important pest of agronomic crops in the United States such as sorghum corn and rice that can damage at the seedling stage. Subramanian (1995) reported that sorghum chinch bug resistance lines, KS94 and KS95 contain significantly higher levels of total phenolics and tannin when compare to susceptible line, Double Dwarf Yellow Milo (DDYM). So, the development of cultivars resistant to insect pests is a quite important. Moreover, other agronomic traits such as green snap and root lodge are always important for sorghum production and productivity. Stem green snap or brittle snap is a sudden breakage of the stalk by strong winds, most often occurs during periods of rapid vegetative growth. As well as, root lodging can be occurred by either rootworm or strong wind and also by saturated soil. Genetic manipulation by DNA markers and QTL analysis have been investigated to study about insect resistance in sorghum such as restriction fragment length polymorphism (RFLP) (Katsar *et al*., 2002; Deu *et al*., 2005); random amplified polymorprism (RAPD) (Black *et al*., 1992; Aikhionbare *et*

al., 1998; Agrama *et al*., 2002); amplified fragment length polymorphism (AFLP) (Keyan *et al*., 2003); simple sequences repeated (SSR) (Agrama *et al*., 2002; Deu *et al*., 2005; Apotikar *et al*., 2013; Somashekhar *et al*., 2013).

Gene control flowering of sorghum

The inflorescence of a sorghum plant is a panicle with a central rachis from primary branches. The shape and color of the panicle varies from cultivar to cultivar. Each panicle contains between 800 and 3,000 seeds, which are usually partly covered by the plumes. The colors of plumes are black, red, brown or tan. The inflorescences of sorghum open at night or in the early morning**.** Sorghum is self-pollinate, although outcrossing occurs around six percent of the time (Poehlman, 1987; Rooney, 2000).

A genetic photoperiod sensitivity control of flowering in sorghum includes a series of six maturity genes that has been found to alter flowering time: Ma_1, Ma_2, Ma_3, Ma_4, Ma_5 and Ma_6. Of these four maturity genes, Ma_1-Ma_4 inhibit flowering under short day, while Ma_5 and Ma_6 represent in special case because whenever they present in both dominant expression, they extremely inhibit floral initiation regardless of day length (Childs *et al*., 1997).

3. The ABC Model of Flowering

Flowering is known to be one of the major agronomic traits related to sugar-related in both sugarcane and sorghum and comprises of the complex development processes. The initiation of flowering is the change taking place in the shoot apex during the transition turn vegetative to reproductive phase (Swapna and Singh, 2008). Up to now, the genetic basis of genetic control of inflorescence and flower development has not been well understood. The ABC model is now the good model that derived from molecular genetics of *Arabidopsis thaliana* and *Antirrhinum majus* (Coen and Meyerowitz, 1991). This model explains the group of genes encodes transcription factors needed to turn on the genes for organ identities and flower development.

In *Arabidopsis*, a large family of MADS-box genes are concerned in floral development (Coen and Meyerowitz, 1991; Theissen *et al*., 2000; Heinz, *et al*., 2001; Becker and Theissen, 2003). MADS-box genes encode transcription factors in all eukaryote. Plant MADS-box proteins contain a DNA-binding, an intervening (I), a Keratin-like (K). They are dimerisation and C-domain, is a transcriptional activator domain. They control the transition from vegetative to generative growth and determine inflorescence meristem identity. It is also found that MADS-box genes are very conserve among the different plant species. There are many genes required for the initiation and flowers development. For simplicity, they can be divided into 4 classes (Yanofsky and Martin, 1995).

1. The first class is flower time gene, bases on their differential responses to environmental condition, such as day length and vernalization.

2. The second class is flower **meristem identity**, including genes such as *LEAFY* (*LFY*), *APETALA1* (*AP1*), and *CAULIFLOWER* (*CAL*) which specify flower meristem identity, as well as *TERMINAL FLOWER* (*TFL*) which maintains inflorescence meristem identity.

3. A third class includes the **flower organ identity** genes, which determine the fate of organ primordia and are incorporated into the **"ABC" model** of flower development (Coen and Meyerowitz, 1991; Joshua and Meyerowitz, 1995; Hong and Claude, 2000).

4. A fourth class includes **late-acting genes that control ovule development**.

The mechanism of ABC model

Class A or A genes are the genes involved the development of sepals. The genes in this class will produce the LEY protein from *LEY* transcription factor and turn on *AP1* and *AP2* gene. At the same time, LEY protein plus UFO protein turn on *AP3* gene, then class A and B genes, these are *AP1, AP2, AP3* and *PI* are expressed

and controlled the development of patels. After that, LEY protein and some unidentified protein turn on *AG* gene. Then, class B and C genes, these are *AP3*, *PI* and *AG* turn on the developmental program for forming stamen. Finally, expression of C gene alone by AG protein will turn on for carpals development (Figure 1).

4. The Profiling of Gene Expression Methods

There are several methods to profile the expression of the thousands of genes in parallel (Pollock, 2002). They are serial analysis of gene expression (SAGE), massively parallel signature sequencing (MPSS), microarray, and differential display such as subtractive hybridization and cDNA-amplified fragments length polymorphisms (cDNA-AFLP)**.**

cDNA-amplified fragments length polymorphisms (cDNA-AFLP)

cDNA-AFLP is an improvement of traditional differential display technique (Bachem *et al*., 1996). It is a PCR-based on AFLP transcript profiling method in any species without the need for sequence knowledge (Bachem *et al*., 1996; Christain *et al*., 1998; Durrant *et al*., 2000; Qin *et al*., 2000**).** This method starts with cDNA synthesis from total RNA or mRNA using random hexamers as primers, follows by restriction enzyme digestion, ligates with specific adapters, selection transcript cDNA profiling into smaller scale by selective amplification and visualizes by high-resolution polyacrylamind gel (Figure 2). This is a fast, high reproducibility, accuracy and reliability technique for the identification of differentially expressed genes (Vuylsteke *et al*., 2007).

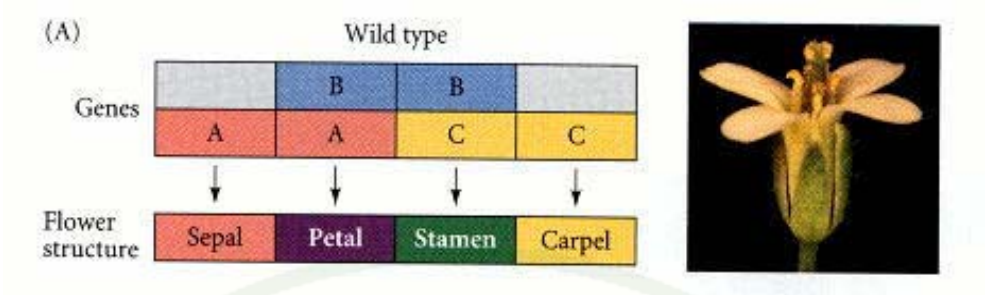


Figure 1 The ABC model for flower development in *Arabidopsis*

Source: Coen and Meyerowitz (1991)

Hsu *et al.* (2008) investigated the cDNA-AFLP technique to study the differential gene expression during tissue culture of flower bud of *Phalaenopsis* Hsiang Fei cv. H.F. The color patterns in orchid have high market value and tissue culture can be a cause of somaclonal variation. The flower bud of wild type *Phalanopsis* have bronze color pattern, while mosaic yellow color occurred in variants. cDNA-AFLP explored 2269 TDFs between those wild type and its somaclonal variation. The result showed that four TDFs showed high homology with genes that known function; casein kinase, isocitrate dehydrogenase, cytochrome P450 and *EMF2.* As well as, Que *et al*. (2011) reported the transcriptome profiling of gene expression during sugarcane *Ustilago scitaminea* infection using cDNA-AFLP. A total of 136 TDFs were found and 28 TDFs showed homology with gene that known function. So, cDNA-AFLP is one of the transcriptome techniques that has powerful to access the genome for gene expression study.

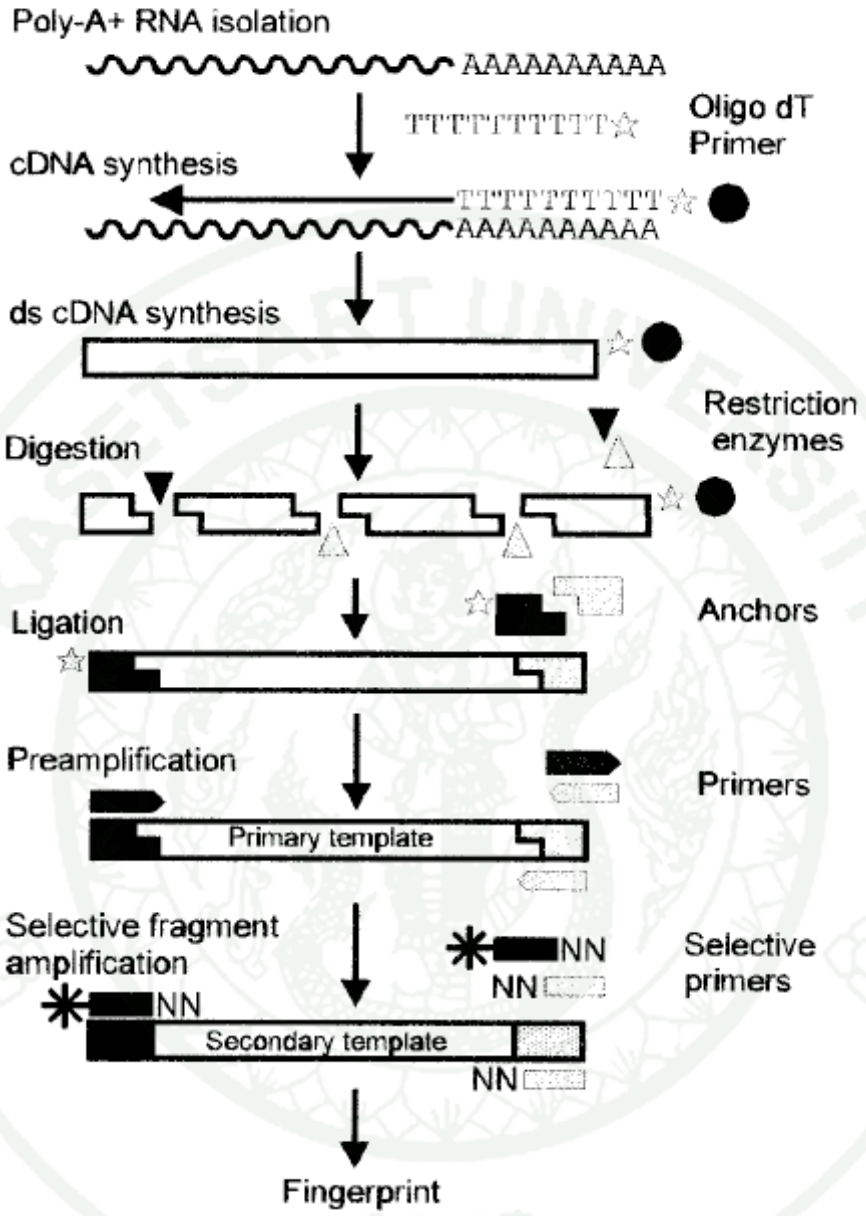
5. Colinearity Genetics Analysis

5.1 Molecular markers

Molecular markers have been developed to be useful tools for genetics study and molecular breeding in crop improvement. The most well known technology has continuously evolved from hybridization-based RFLP (restriction fragment length

polymorphisms) in 1975 to identify DNA sequence polymorprisms for genetic mapping of temperature-sensitive mutation of adeno-virus serotype (Grodzicker *et al*., 1975). The others are random amplified polymorphic DNA (RAPD) (Williams *et al*., 1990); expressed sequence tags (EST) (Adams *et al*., 1991); simple sequence repeats (SSR) (Akkaya *et al*., 1992); single nucleotide polymorprism (SNP) (Jordan and Humphries, 1994); amplified frangment length polymorism (AFL) (Vos *et al*., 1995); diversity array technology (DArT) (Jaccoud *et al*., 2001) and culminating in ultra high-throughput genotyping by sequencing (GBS) at the present.

Microsatellite or short tandem repeats or simple sequence repeats (SSRs) are monotonous repetitions of very short (one to five) nucleotide units, which occur as interspersed repetitive elements in all eukaryotic genome (Tautz and Renz, 1984; Agarwal *et al*., 2008). Dinucleotide repeats like $(CA)_n$ and $(GA)_n$ are the most common repeats in the most eukaryote such as in human $(CA)_n$ repeat occurs once in every 30 kb. SSRs are very polymorphic due to the high mutation rate affecting the number of repeat units. They have locus identity and they are multi-allelic. SSRs markers can be used to detect the variation in the number of short repeat sequences, usually two or three bases repeats as well. The polymorprisms can be easily detected on high resolution gel such as sequencing gel by running PCR amplified fragments product that obtained using unique pair of primers flanking the repeat (Weber and May, 1989). They have wide application for genetic analysis in crop improvement or breeding programs. They are widely used in plants because they are evenly distributed all over the genome, co-dominant, hyper-variability. Little DNA is required, radioactive is not required and suitability for high throughput analysis. Although, the high throughput technologies are available, but SSRs marker still is as an important molecular marker for wide range applications such as genome mapping, marker assisted selection and diversity studies.



Figture 2 Diagram of the cDNA-AFLP method: Step-by-Step procedure

Source: Bachem *et al*. (1996)

5.2 Quantitative trait loci (QTL) analysis in agronomy-related traits in sorghum and sugarcane

QTLs are the short segments of DNA (locus) that have some contribution towards the phenotypic value of quantitative traits. Such locus may carry single or group of genes that are tightly linked and mostly inherited together. Many loci determine the total phenotypic value of the trait such as yield, so each of these loci is called QTLs. Major QTLs are those loci that have major influence on the phenotypic value, whereas minor QTLs have minor influence on the phenotypic value.

QTL mapping is the one of important tools to study in phenotypic variation and plays a role in the success of breeding program. Various studies on sorghum QTL for important agronomy traits base on DNA markers have been reported i.e. flowering (Ray *et al*., 2002; Lekgari, 2010; Yousra *et al*., 2012), plant height (Ray *et al*., 2002; Patrick *et al*., 2008; Lekgari, 2010; Madhusudhana and Patil, 2013), sugar-related agronomy traits (Bian *et al*., 2006; Kimberley *et al*., 2008; Amukelani *et al*., 2010; Lekgari, 2010; Peng *et al*., 2013), grain yield and related agronomy traits (Rami *et al*., 1998; Srinivas *et al*., 2009; Guihua *et al*., 2012; Reddy *et al*., 2013) and disease resistance (Gowda *et al*., 1995; Tao *et al*., 1998; Parth *et al*., 2008; Mohan *et al*., 2010).

While in sugarcane breeding programs, it takes at least 12 years to develop the new commercial cultivars. Molecular markers and QTL analysis provide the opportunity to access the genetic structure of quantitative traits. The two major complicate factors make QTL mapping of sugarcane more difficult than other species are (i) Ploidy level: the cause of complexity genetic by polyploidy and aneuploidy of sugarcane (Heinz and Tew, 1987) (ii) Outbred parents: as sugarcane inbred line is not available, the genetic linkage mapping construction and QTL analysis have been analyzed by high segregation progenies derived from a cross between highly heterozygous outbred parents. So, the copy number of variation or allele dosage has been appearanced by these two factors. The first genetic linkage map in sugarcane was constructed after the development of single-dose markers (SDMs) (Wu *et al*.,

1992; Pastina *et al*., 2012). In a parental cross, an SDM has either a single copy of an allele in one parent only or a single copy of the same allele in both parents, thus segregating in 1:1 (presence : absence) or 3:1 (presence: absence) ratio, respectively

5.3 Colinearily between different plant species

In plant genome, the colinearity studies between species have been reported such as Lagercrantz *et al.* (1996) studied genome colinearity of gene controlling flowering time in *Arabidopsis thaliana* and *Brassica nigra.* The genetic linkage map of *B. nigra* was constructed by 88 F_1 population that derived from a cross between early flowering *B. nigra* (RC) and late-flowering wild Italian accession (Cat). The eleven segments contig surrounding *CO* gene of *A. thaliana* was used as a RFLP probe within *B. nigra* genetic map for colinearty study. The result showed that those *A. thaliana* contig *CO* gene segments homology with *B. nigra* linkage groups 2, 5 and 8 as showed in Figure 3.

The genera *Saccharum* (sugarcane) and sorghum are closely related, and sharing common ancestor about 5 million years ago (Al-Janabi *et al*., 1994). Although, sugarcane has complex polyploid genomes, where as sorghum has a simple diploid genome. Wang *et al*. (2010) reported the microsynteny between sugarcane and sorghum, assessed by comparing 454 pyrosequences of 20 sugarcane bacterial artificial chromosomes (BACs) with sorghum sequences. The result showed that the genetic regions of the sugarcane BACs shared an average of 95.2% sequence identity with sorghum.

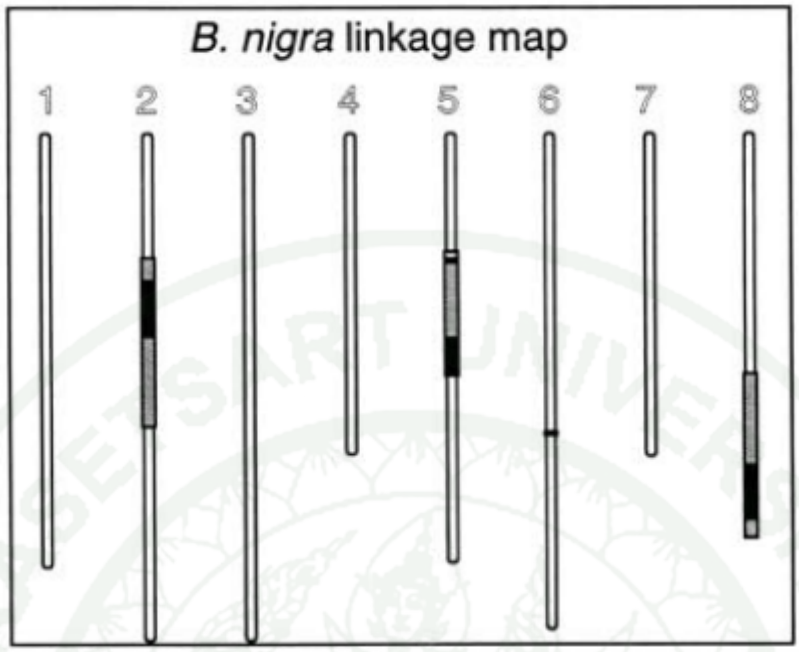


Figure 3 The colinearity between genetic linkage map of *B. nigra* and *A. thaliana* contig *CO* gene segments (solid box)

Source Lagercrantz *et al*. (1996)

Molecular genetic marker cannot be applied to sugarcane because of their high ploidy level (Grivet *et al*., 1994). Linkage between two loci can only determined by a single dose restriction fragment (SDRF) such as a fragment corresponding to an allele present only once per genome, therefore it would be segregated as a monogenic or dominant marker (Wu *et al*., 1992). Comparative mapping using a diploid relative can help to identify the complex polyploid genome species such as sugarcane (Grivet *et al*., 1994).

In 1998, Ming *et al.* conducted comparative genetic map between F_1 progeny of the cross between *S. officinarum* and *S. spontaneun* and sorghum. About 84% of the loci mapped by 242 common probes were homologous between *Saccharum* and *Sorghum* genome. Moreover, Ming *et al*. (2002) have studied quantitative trait loci (QTL) affecting plant height and flowering in segregation population derived from

cross between *S. officinarum* and *S. spontaneum*. The QTL mapping was conducted based on 1,255 single-dose RFLP markers. The results showed that 4 QTLs controlled plant height in sugarcane and corresponded closely to 4 of 6 plant-height QTLs mapped in sorghum. And one QTL controlled flowering in sugarcane and corresponded closely to 1 of 3 flowering QTLs mapped in sorghum.

MATERIALS AND METHODS

Identification of the sugarcane flowering putative genes were performed from inflorescences of *S. spontaneum* line 98-244 which is early flowering line. The sugarcane plants were grown in the field at Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand. For sorghum mapping, the 170 F_5 RIL population that derived from a cross between DDYM, early flowering cultivar and Mapila, none flowering line were used for generating genetic linkage map. The sorghum population was grown in the greenhouse and in the field for phenotypic evaluation at the University of Nebraska-Linclon, Nebraska, USA. QTL analysis and colinearity were investigated to study the floral synteny region between sugarcane and sorghum using JoinMap program and WinChartQTL statistical software as described below.

Part I: Genetic linkage map construction of flowering genes in sorghum

1.1. Plant materials: Genomic DNA extraction

A population of 170 F_5 recombinant inbred line (RILs) of sorghum that derived from a cross between Double Dwarf Yellow Milo (DDYM, early flowering cultivar, photoperiod insensitive) and Mapila (PI 524746, non-flowering line, photoperiod sensitive) was developed through single seed descent.

Fresh leaves of 14-day old plants, grown in a greenhouse, were used for genomic DNA extraction using the cetyltrimethyl ammonium bromide (CTAB) method described by Dweikat (2005). The tissue was grounded by SAP extractor with extraction buffer (Appendix A2.1) in 15 mL centrifuge tube. The tissue mixture was mixed by briefly vortex, and then incubated in water bath at 65 °C for 1 hr. The mixture was laid down for cooling at the room temperature for 10 min, then the equal volume of 24:1 chloroform: isoamyl alcohol (C:I) was added and mixed vigorously. After centrifugation at 3,000 rpm for 10 min, the supernatant was transferred to new clean 15 mL centrifuge tube, and DNA was precipitated with equal volume of cold

isopropanol and 0.2 volume of 5 M NaCl, gently mixed and incubated at -20 °C for 30 min to 2 hr. The mixture was centrifuged at 3,000 rpm for 30 min, and washed twice with 70% cold ethanol. DNA pellet was air dried at room temperature and then re-suspended in TE buffer (Appendix A2.2) with 20 ng RNase A, incubated at 37 °C for 30 min, and then let stand at room temperature for overnight. The equal volume of 24:1 C: I was added into the mixture and centrifuged at 3,000 for 10 min. The supernatant was transferred to clean 15 mL centrifuge tube. The two volume of cold absolute ethanol and 0.2 volume of 8M ammonium acetate were added for DNA precipitation, and then incubated at -20 °C for 30 min to 2 hr. After centrifugation at 3,000 rpm for 10 min, the DNA pellets were air dried at room temperature, and then re-suspended with 100-500 μLTE buffer depending on size of the pellet. DNA concentration was qualified using a fluorophotometer (TKO 100 Fluorophotometer, Hoefer Scientific Instruments, San Francisco, CA).

1.2 Screening for markers

A total of 1,601 primers were used for polymorphic parental screening [A set of 1,375 sorghum SSR primers according to Lekgari (2010) that were collected from Brown *et al*. (1996), Taramino *et al*. (1997), Bhattramakki *et al*. (2000), Kong *et al*. (2000), Schloss *et al*. (2002), Lubbock (unpublished), Burrow *et al*. (2008), Srinivas *et al*. (2008), Srinivas *et al*. (2009) and Li *et al*. (2009), 2 QTLs makers linked to flowering time in wild barley (Ivandic *et al*., 2002), 177 *Zea mays* SSR primers, 25 QTLs flanking markers linked to related traits of flowering and silk in maize (Xie *et al*., 2010; Salvi *et al*., 2011), *FLORICAULA/LEAFY* (*FLO/LFY)* orthologs transcription factor (Kiesten and Doebley, 2005), *APETALA1 (AP1)*-like MADS box gene in wheat (*WAP1*) (Murai *et al*., 2003) and 26 flower-specific sugarcane TDFs that derived by cDNA-AFLP experiment (listed in Appendix Table 2)] were used. A total of 1,601 markers were screened for detection polymorphisms between parental lines. Overall, 348 markers were polymorphic, and only 196 co-dominant markers and 10 dominant sugarcane TDFs were used for linkage map construction.

PCR amplification for simple sequences repeat (SSR) analysis was performed in 25 µL reaction mixture containing 60 ng DNA template, 100 ng primer pair, 125 µM dNTP, 50 mM KCl, 10 mM Tris-HCl, 2 mM $MgCl_2$ and 0.3 unit *Taq* polymerase. The PCR profile was performed as follow:

Step 1	pre-denature	94 °C for 3 min
Step 2	denature	94 °C for 1 min
Step 3	annealing	50°C to 56 °C for 1 min
Step 4	extension	72°C for 1 min (repeated step 2 to 4 for 35 cycles)
Step 5	final extension	72 °C for 5 min.

The amplification products were mixed with DGGE loading buffer (Appendix A2.3), and visualized in 12% non-denatured polyacrylamide gels (Appendix A2.4). The running buffer is 1X DGGE buffer (Appendix A2.5) and using constant voltage of 300V for 150 min. The gel was run in vertical gel electrophoresis in cooling system to control the temperature at 17-20 °C. The gels were stained with 1 µg/mL ethidium bromide solution for 10 min and de-strained in deionized water for 15 min, then images were collected with Gel Doc2000 with 1D imaging software (Bio-Rad, Hercules, CA).

1.3 Phenotypic data collection

170 F_5 recombinant inbred lines were developed from the cross of early flowering cultivar and non-flowering line. The 170 RILs and the parental lines were planted at Lincoln, Nebraska in an augmented incomplete block design. The sorghum plots were planted adjacent to proso millet rows. The millet was utilized as a trap crop to attract natural adult chinch bug population which would allow for more uniform infestation (Rajewski *et al*., 2009).

Flowering data

Flowering data were collected on season 2012. The day to pollen shed was recorded on each individual plant.

Plant height data

From the ground level to the top visible dewlap, was taken on individual plants. The average value of three replications was used for analysis.

Chinch bug resistance

Damage from chinch bug was determined by the mortality percentage assessed at weekly intervals from July to September. The relative resistance was determined by area under the insect damage curve (AUIDC), where:

$$AUIDC = \Sigma\left[\frac{M_n + M_{(n+1)}}{2}\right] * Y_{(n+1)} - Y_n$$

M_n =mortality at time n and Y_n is equal to day at time n.

All traits were analyzed using the statistical software, Minitab, trial version (http://www.minitab.com).

1.4 Genetic linkage map construction and QTLs analysis

A total of 170 F_5 RILs were used to construct linkage groups using Joinmap® 3.0 software (van Ooijen and Voorrips, 2001) with a threshold LOD (logarithm of odds ratio) greater than 3.0, 0.25 recombination frequency (r^2). Kosambi mapping function was used to transform recombination frequency into genetic map distance (cM) (Kosambi, 1944). The linkage groups of SSR marker were assigned to sorghum chromosomes.

The main effect QTLs were detected using the composite interval mapping method (CIM) in Window QTL Cartographer V2.5_010 (Wang *et al*., 2007). A QTL map was drawn using Mapchart 2.0 (Voorrips, 2002).The LOD score for declaring of a putative QTL was determined by a 1000 permutation test. The significant *P* value of 0.05 was used for model selection. The percentage of the phenotypic variation explained by a QTL was estimated as the coefficient of determination (R^2) using single-factor analysis from a general linear models procedure (Wang *et al*., 2007). If two peaks in the same linkage group are detected for single trait, the determination will be depended on the distance between the QTLs. Where the distance between the QTLs is greater than 20 cM, they will be considered as two QTLs. Where it is less than 20 cM, only the higher peak will be considered for QTL position (Ungerer *et al*., 2002: Parth *et al*., 2008).

Part II: Identification of putative flowering genes in sugarcane

2.1 Total RNA preparation

2.1.1 Total RNA isolation

Total RNA from *S. spontaneum* S98-244 line which is an early flowering line was extracted from various size of inflorescences ranges from 0, 0.5, 1, 2, 3, 10, 15 cm (Figure 4) using the Pine Tree method (Chang *et al*., 1993) with minor modifications. The external control was shoot apex of 4 month-old-sugarcane. The hypothesis of the expression of gene during inflorescence development is initial stage at 0.0 cm, on the process of inflorescence development at 0.5-10.0 cm and blooming stage at 15.0 cm. In brief, first day, the samples were grounded into fine power in liquid nitrogen. The 0.5 mL extraction buffer (Appendix A2.6) was added into the frozen sample in 1.5 mL centrifuge tube. The mixture was mixed well by stirring with clean pipette tip. After 5 min, the mixture was added with an equal volume of 24:1 Chloroform: Isoamyl alcohol (C:I), mixed well by vortexing. The mixture was centrifuged at 10,000 rpm for 15 min. The upper aqueous phase was transferred into a

new clean RNase-free tube, and 8 M LiCl was added with the final concentration of 2 M and then mixed by inverting. The mixture was incubated at 4 °C for overnight.

Second day, after incubation, the mixture was centrifuged at 10,000 rpm for 20 min at 4 °C, and then removed the supernatant. The RNA pellets were dissolved in 0.5 mL SSTE buffer (Appendix A2.7) followed with equal volume of 24:1 C:I, and mixed by overtaxing immediately. After centrifugation at 10,000 rpm for 10 min, the supernatant was transferred to a new clean RNase-free tube and added with two equal volume of cold absolute ethanol for RNA precipitation. The mixture was incubated at -20 °C for at least 2 hr, and then centrifuged at 10,000 rpm for 10 min. The RNA pellets were washed twice with 70% cold ethanol, air dried, and then re-suspended with 20-50 µL nuclease free water depending on the size of the pellets. The quality and quantity of total RNA were determined using nanoDrop 8000 UV-Vis spectrophotometer (Thermo scienctific, USA) at 260 nm and 2% denatured agarose gel electrophoresis in 1X NBC running buffer (Appendix A2.8). The RNA was stored at -20 °C for short storage and -70 °C for long storage.

2.1.2 Template preparation

The DNase was applied to remove genomic DNA contaminated. All the samples were treated with DNase I enzyme (#EN0521, Thermo scientific) following to manufacturer's instruction. In brief, one microgram of total RNA was treated with 1 unit of DNase I enzyme. The reaction mixture was incubated at 37 °C for 30 min, and followed with 1 µL of 50 mM EDTA. The reaction mixture was incubated at 65 °C for 10 min for inactivation DNase I enzyme activity, and then placed on ice for 1 min. Alternatively, used 24:1 C:I extraction for purify total RNA. The final total RNA concentration was determined using nanoDrop 8000 UV-Vis spectrophotometer at 260 nm absorbance.

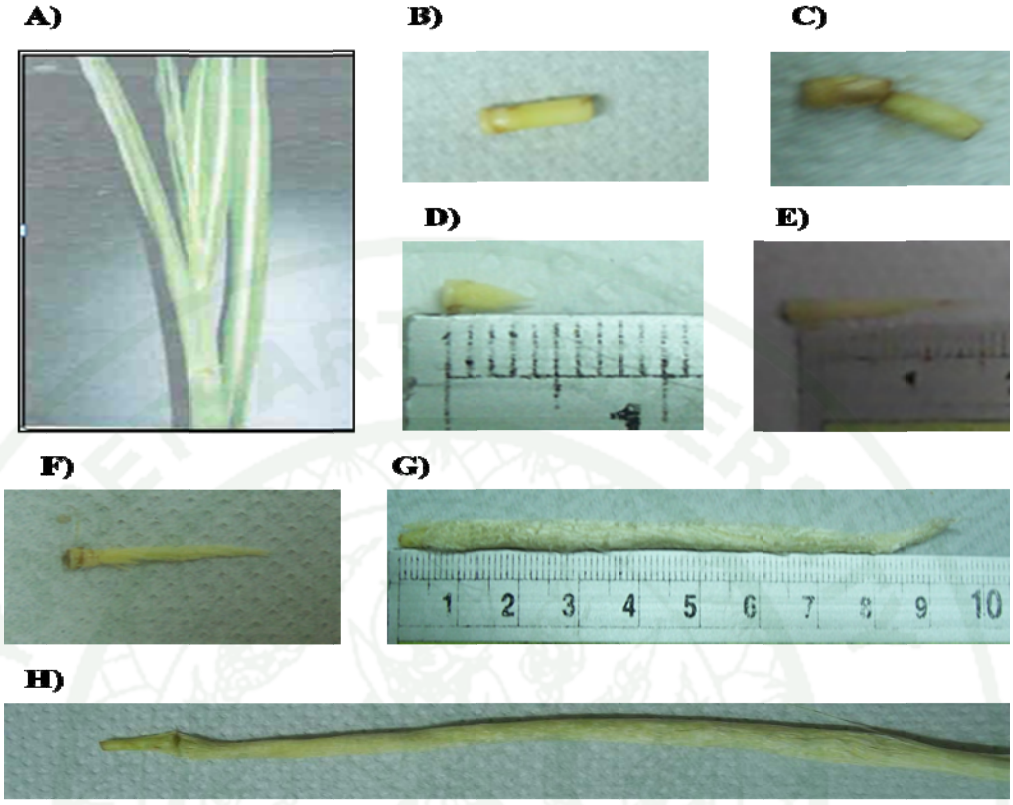


Figure 4 The inflorescence of *S. spontaneoum* line 98-244 for total RNA extraction (A) Top sugarcane stem before remove leave sheath (B) shoot apex at 4 months old (C-H) 0-15.0 cm of the inflorescence length after removed leave sheath

2.2. cDNA-AFLP analysis

The cDNA-AFLP was subjected to standard AFLP technique according to Vos *et al*. (1995) with minor modifications. The cDNA-AFLP transcript profile was determined by selective PCR amplification using 24 different primer combinations.

2.2.1 First-stand cDNA synthesis

First strand was carried out using Superscript™ III First Strand Synthesis System kit (Cat.No: 18080-051, Invitrogen™). About 400 ng of total RNA were mixed with 1 µL of 50 µM Oligo$(dT)_{20}$ primer, 1 µL of 10 mM dNTPs mix, and made

up to 10 µL with DEPC-treated water. The RNA mixture was mixed well and incubated at 65 °C for 5 min, and then place on ice for 1 min. The first strand cDNA synthesis was synthesized with 200 unit of SuperSript™ RT enzyme containing 2 µL of 10X RT buffer, 4 µL of 25 mM $MgCl_2$, 2 µL of 0.1 M DTT and 40 units of RNase OUT™. The reaction mixture was incubated at 50 °C for 50 min, terminated the reaction at 85 °C for 5 min, and then chill on ice. Collect the reaction by briefly centrifuged, 1 unit of RNase H was added, and incubated at 37 °C for 20 min.

2.2.2 Second-stand cDNA synthesis

Before the second strand of cDNA synthesize, the first strand cDNA was pre-treated with 1X DNA polymerase I buffer, 0.2 mM dNTP mix, made up to 140 µL with nuclease free water, and then incubated on ice for 20 min. The reaction mixture was added with 30 units *E. coli* DNA polymerase I (#EP0042,Thermo Scientific) for short fragments of second strand cDNA synthesis, 10 units *E. coli* DNA ligase (Invitrogen™) for ligation those fragments, and then incubated at 16 °C for 2 hr. After that, 5 units T_4 DNA polymerase enzyme (# EP0062, Thermo Scientific) and 0.1 mM dNTPs were added, and incubated at 16 °C for 15 min. The reaction was terminated by adding 10 µL 0.5 M EDTA pH 8.0.The double strand cDNA (ds cDNA) was purified by Wizard® SV Gel and PCR Clean-Up kit (#A9281, Promega). In brief, equal volume of Membrane Binding Solution was added to the ds-cDNA solution, and mixed well by pipetting. The solution mixture was transferred into the column, and incubated for 1 min at room temperature. After centrifugation at 10,000 rpm for 1 min, the membrane that carries cDNA was washed twice with Membrane Wash Solution, and then centrifuged at 10,000 rpm for 5 min. For cDNA elution, 50 µL of nuclease free water was added, incubated 1 min at room temperature, and then centrifuged at 10,000 rpm for 1 min. The cDNA concentration was qualified again using nanoDrop8000 spectrophotometer and ready for further cDNA-AFLP analysis.

2.2.3 Restriction enzyme digestion

The first step of AFLP technique is to generate restriction fragments by using two restriction endonucleases; a rare cutter (*Eco* RI: six-base cutting) and a frequent cutter (*Mse* I or *Tru* 9I, a four-base cutting). About 250 ng of cDNA was digested with 12 units *Eco* RI (#R6011, Promega) and 8 units *Tru* 9I (#R7011, Promega) containing 1X SuRE Cut buffer A (#11417959001, Roche), made up to 20 µL with ddH_2O, and then incubated at 37 °C for 1.5 hr. The reaction mixture was briefly centrifuged for the next step ligation.

2.2.4 Adapter ligation

Digested product was ligated with specific *EcoR*I and *Mse*I adapters (Appendix A1). The double stranded adapters were ligated to the end of the cDNA fragments, to generate the DNA template for the PCR reaction. The ligation mixture was prepared by adding 1 unit T_4 DNA ligase (#M 1801, Promega) containing 1X SuRE Cut buffer A, 2 µL 5 pMol/ µL *Eco* RI adapter, 2.5 µL 50 pMol/ µL *Mse* I adapter, 1.5 µL ATP, made up to 10 µL with ddH_2O, and then the ligation mixture was added into digested product. The digested-ligated reaction mixture was incubated at 16 °C for overnight. To inactivate all enzyme activity, the reaction mixture was incubated at 75 °C for 5 min. All reaction samples were kept at -20 °C until used.

2.2.5 Pre-selective amplification

The digested-ligated product was diluted for 10X dilutions, and 4 µL of dilution product was used to be a template in pre-selective amplification. The PCR reaction was performed using 4 µL of dilution product, 1X PCR buffer, 0.25 mM *Eco* RI primer-N, 0.25 mM *Mse* I primer-N (N referred to any selective base such as A and G, respectively), 0.2 mM dNTPs, 3 mM $MgCl_2$, 2 units *Taq* DNA polymerase (#EP0401, Thermo scientific), and then made up the volume to 20 µL with ddH_2O. The PCR profile was performed as follow;

Step 1 pre denature 94 °C for 2 min
Step 2 denature 94 °C for 30 sec
Step 3 annealing 56 °C for 30 sec
Step 4 extension 72 °C for 1 min (repeated step 2-4 for 20 cycles)
Step 5 final extension 72 °C for 5 min

2.2.6 Selective amplification

The pre-selective amplification product was diluted for 10X dilutions. The PCR reaction was performed using 2 µL of dilution product, 1X PCR buffer, 0.25 mM *Eco* RI primer-NNN, 0.25 mM *Mse* I primer-NNN (NNN referred to any selective base such as ACA and GTC, respectively), 0.2 mM dNTPs, 3 mM $MgCl_2$, 1 units *Taq* DNA polymerase and then made up the volume to 10 µL with ddH_2O. The PCR profile was performed as follow;

Step 1 pre denature 94 °C for 2 min
Step 2 denature 94 °C for 30 sec
Step 3 annealing 65 °C for 30 sec
Step 4 extension 72 °C for 1 min (repeated step 2-4 for step-down of annealing 0.7 °C in each step for 12 cycles)
Step 5 denature 94 °C for 30 sec
Step 6 annealing 56 °C for 30 sec
Step 7 extension 72 °C for 1 min (repeated step 5-7 for 25 cycles)
Step 8 final extension 72 °C for 5 min

2.2.7 Visualization using silver staining

The selective PCR products were separated by 5% denatured polyacrylamide gel and visualized by silver nitrate straining (Sanguinetti *et al*., 1994). For casing gel electrophoresis, chamber and glass plate were cleaned 3 times with

95% ethanol, let them dried for 2-3 min. Chamber was treated with 5-6 drops of CLEAR VIEW®. The glass plate was treated with 1 mL glass bound solution (Appendix A2.9) followed 3 times with 95% ethanol. Gel assembly between chamber and glass was composed.

90 mL of 5% denatured polyacylamind gel (Appendix A2.10) was mixed with, 200 µL 10% ammonium persulfate (APS, Appendix A2.11) (#A3678, Sigma-Aldrich) and 20 µL TEMED. The solution mixture was mixed well, poured into gel apparatus, and then left it for polymerization at room temperature for at least 1.5 hr. The amplified products were mixed with 0.4 volume of 1X sequencing loading buffer (Appendix A2.12).The gel was pre-heat or pre-run to 50 °C and the amplified products mixture were loaded at 2 µL per lane. After the running complete at 2 -2.5 hr with 75 W, 300 mA and 3000 V, the gel was visualized using silver staining method.

The gel was fixed in 10% glacial acetic acid (Appendix A2.13) for 20 min with shaking, and then the gel was washed twice for 3 min with ddH_2O. The gel was stained with silver solution (Appendix A2.14) for 30 min, and rinsed with ddH_2O. The gel was developed the band with cool developer solution (Appendix A2.15) until clear pattern appeared. The pattern on the gel was stopped by incubating in 10% glacial acetic acid for 1-3 min, washed the gel with ddH_2O for 30 min, and then air dried at room temperature.

2.3 Classification of specific differentially expression of flowering transcript derived fragments (TDF)

The polymorphic patterns were determined and grouped into many classes, according to the pattern of the DNA banding and flower development stages.

The TDFs from cDNA-AFLP experiment were classified as

Class A1 early-induced which is the gene expresses from the initiation of flowering through blooming

Class A2 late-induced which is the gene expresses at 1.0 cm of inflorescences until blooming stage

Class B up-regulated which is a low level of gene expression at control stage (vegetative stage) is continuously increased until blooming stage

Class C down-regulated which is continuously decreased or down-regulated or switch-off of gene expression under natural environment

Class D early-up and late-down regulated which is a low level of gene expression increases continuously from the control up to stage of 1.0 cm long inflorescence, after that is decreases continuously until the blooming stage.

2.4 TDF recovery from agarose gel electrophoresis

Bands corresponding to those classes were cut out from acylamind gel and transferred to 1.5 mL microcentrifuge tube, 20 µL sterilized H_2O was added and then boiled for 10 min. 2 µL of those products was used to be as template for re-amplification under the condition used for pre-selective amplification (2.2.5). After PCR completed, the PCR products were separated on 1% agarose and 1X TBE running buffer for 40 min (Appendix A2.16). The accurate DNA banding was cut from the gel and eluted using NucleoSpin® Gel and PCR Clean-up kit (#740609,MACHEREY-NAGEL). In brief, an equal volume of NTI buffer per gram weight of agarose gel (containing the DNA) was added, and incubated the DNA sample for 5-10 min at 50 °C. Vortex the DNA sample briefly every 2-3 min until the gel slice is completely dissolved. The DNA solution was loaded into the column, and centrifuged at 10,000 rpm for 30 sec to bind the DNA with the membrane. The membrane was washed twice by adding 700 µL NT3 buffer, and then centrifuged at 10,000 rpm for 30 sec. After drying the membrane and 15-30 µL NE buffer was added to elude the DNA, and then centrifuged at 10,000 rpm for 1 min. The DNA sample was qualified the quality, quantity and size using NanoDrop pectrophotometer and 1% agarose gel electrophoresis, respectively.

2.5 TDF sequencing and homology analysis

TDFs sequences were determined using an automated DNA sequencer; Applied Biosystems highest capacity-based genetic analyzer platforms (1st BASE Pte Ltd, Singapore). After sequenced, the TDFs sequences were compared with nucleotide and protein sequences by The Basic Local Alignment Search Tool (BLAST). Sequence alignments with an *E*-value of 10^{-3} or close to 10^{-3} were adopted to assert a significant match between TDFs sequences and nucleotide sequences database.

Part III: Colinearity of flower-specific sugarcane TDFs and sorghum genome

3.1 Selected TDFs primer design

The significantly TDFs that show high homology with interested gene or protein with known or unknown function, transcription factors and hypothetical protein were selected for colinearity study with sorghum genome. The specific primers were designed with Primer 3 Program free online tool by defaults parameters (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). The primer attribute was determined follow the generally rule such as 18-22 bp primer length, 52-58 °C primer melting temperature (T_m), 40-60% GC content, GC clamp, avoided the primer secondary structure. The most critical of primer design is primer annealing temperature (T_a) which is calculated by

$$T_a = 0.3 \times T_m\,(\text{primer}) + 0.7\ T_m\,(\text{product}) - 14.9$$

where,

T_m(primer) = melting temperature of the primers

T_m(product) = melting temperature of the product

Too high T_a will produce insufficient primer-template hybridization resulting in low PCR product yield. Too low T_a may possibly lead to non-specific products; a high number of base pair mismatches will be occurred**.**

3.2 Polymorphic TDFs with sorghum RILs population

The specific primer TDFs were test for polymorphic in 170 F_5 sorghum RILs population as the same procedures as describe in Part I: 1.2.

3.3 Colinearlity analysis

3.3.1 TDFs map construction and QTL analysis with sorghum mapping.

While sorghum linkage groups was constructed using all co-dominant markers, the significantly TDFs (described above from Part II) were also tested for polymorphic in 170 F_5 sorghum RILs population and added to genetic map construction for colinear study and QTL analysis. For co-dominant TDFs markers, a band was scored either as “A” (DDYM allele) or “B” (Mapila allele) and “H” (heterozygote). For dominant TDFs markers, a resultant band was scored as “D” and missing band as (-). The methodology for TDF mapping and QTLs analysis was done the same with sorghum as described in Part I; 1.4. All sorghum phenotypes; flowering date, plant height and chinch bug resistance were test with TDFs QTLs that derived from sugarcane as well.

3.3.2 Gene annotation

Moreover, for generate the gene prediction, AUGUSTUS version 2.7 (Stanke *et al*., 2008) was used and gene annotation using Gramene Database (http://www.gramene.org/) was investigated to chromosome homology of the flower-specific sugarcane TDFs using DNA and protein database. The hit filtering criteria was at least *E-value* threshold of $10e^{-10}$ and minimum amino acid identity was 40%.

THE EXPERIMENTAL TIME AND PLACES

The experiments were carried out during June 2009-Dec 2012 and the research was conducted at the places as described below:

1. Plant Tissue Culture and Transformation Laboratory, Center for Agricultural Biotechnology, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand.

2. Field Laboratory of Agronomy department at Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand.

3. Crop Science Laboratory, University of Nebraska-Linclon, Nebraska, USA

RESULTS

I. cDNA-AFLP fingerprinting of sugarcane

1.1 RNA extraction

The total RNA was isolated from various stages of inflorescences range from 0, 0.5, 1, 3, 10, 15 cm and sugarcane shoot apex of 4 months years old as an external control with Pine Tree method. The result showed all samples contaminated with gDNA (Figure 5). The problem of RNA isolation from sugarcane is the contamination of phenolics compound and polysaccharides, resulting in brown precipitation. That problem occurrs by the oxidation of phenolic compounds, which can bind to nucleic acid and co- precipitate with RNA. In order to overcome this problem, PVP and β-mercaptonethanol were added into the extraction buffer. Moreover, CTAB was used as the detergent and extract with chloroform instead of phenol to remove protein, for getting colorless RNA (Chang *et al*., 1993).

The high NaCl concentration and CTAB in the extraction buffer can be helped to remove polysaccharide (Fang *et al*., 1992). So, 2 M NaCl was used instead of usual 0.7 M NaCl in the extraction buffer and 1 M NaCl in SSTE buffer to dissolve the RNA pellet. The chloroform extraction is a good step to dissolves the CTAB-RNA complex. The yield and quality of total RNA were good (A_{260}/A_{280} ratio above 1.8) and was good enough for cDNA synthesis (Figure 5, Table 1).

For preparing RNA template for cDNA synthesis, all samples were treated with DNase I enzyme to remove contaminated gDNA. Dnase I, RNase-free is an endonuclease that digests single-and double-stranded DNA or unwanted DNA. The enzyme works by cleaving DNA into 5’ phosphodinucleotide and small oligonucleotide fragment. After treatment, the RNA temples were qualified again and ready for cDNA synthesis.

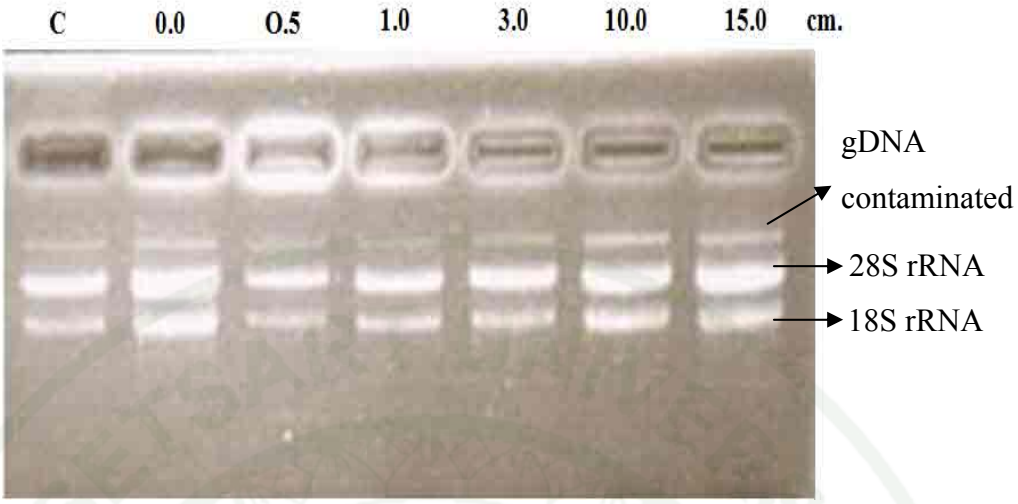


Figure 5 Total RNA analysis by 1.2% denaturing agarose gel electrophoresis. The 28S rRNA, 18s rRNA and gDNA contamination are indicated.

Table 1 The quality and quantity of total RNA extracted from various stage of *S. spontaneum* inflorescences.

Length of inflorescence (cm)	Concentration (ng/µL)	A_{260}/A_{280}	A_{260}/A_{230}
Control	134	2.03	2.17
0.0	334	2.00	1.97
0.5	50	2.14	2.18
1.0	140	2.14	2.40
3.0	102	1.98	1.90
10.0	196	2.1	2.30
15.0	127	2.2	2.15

1.1cDNA-AFLP analysis

In order to identify genes that related to flowering in sugarcane, a RNA fingerprinting technique, cDNA-AFLP, was investigated to display transcript profiling of flowering genes during various stages of flower development. First and second stand cDNA were prepared from total RNA. The first stand cDNA was performed using poly(A)$^+$ with oligo (dT) as a primer. The superscript® III reverse transcriptase is a version of M-MLV RT that was engineered to reduce RNase H activity and provide increased thermal stability. Presenting of RNase H during first stand synthesis can be degraded the mRNA template. Second stand cDNAs were synthesized in a short fragments using first stand cDNA as a template with the activity of *E. coli* DNA polymerase I. After that, short fragments were ligated together using *E. coli* DNA ligase. The 3' end of second stand cDNA was filled up by T_4 DNA polymerase enzyme. For generating the template for AFLP-PCR, cDNAs were double digested with *Eco*RI and *Mse*I restriction enzyme, and then ligated with *Eco*RI and *Mse*I adapter.

The AFLP procedure was performed according to original AFLP protocol with minor modification. Each reaction of digested-ligated cDNA products was diluted to 10 folds and used as a template for pre-amplification with *Eco*RI primer with one selective base such as A vs *Mse*I primer with one selective base such as G. The pre-amplification product was diluted to 10 folds and used as a template for selective amplification with *Eco*RI primer with three selective bases such as ACA vs *Mse*I primer with three selective bases such as GTC as well. The dilution of the template is quite necessary to preventing the background because high amount of templates causes of competitive inhibition between fragments during PCR reaction (Bachem *et al*., 1998)

The differentially expression of cDNA was determined by PCR selective amplification using 26 different primer combinations to identify transcript-derived fragments (TDFs) (Appendix Table A1). Overall, 183 TDFs were detected polymorphic patterns between samples from the various stages of inflorescences as

show in the Figture 6. The result suggested the gene expression change could occur by genes and transcription factors during transition of vegetative to reproductive stage of flower development in natural environment. Among them, 96 TDFs were successfully recovered from denatured polyacrylamide gels for the additional study (Figure 7).

1.3 Identification of differentially flower-specific sugarcane TDFs

In order to characterize the differential expression, the TDFs from cDNA-AFLP experiment were classified as “early-induced” (Class A1), “late-induced” (ClassA2), “up-regulated” (Class B,), “down-regulated” (Class C) and early-up and late-down regulated (Class D) (Table 2, adapted from He *et al*., 2012). Thirty four TDFs (35.42%) belonged to early-induced or Class A1, the gene expresses from the initiation of flowering through blooming, which means those TDFs were induced after the initiation stage by genes or transcription factors-related to promoting flower but were not expressed before. For expression-related Class A2 or late-induced, the gene expresses at 1.0 cm of inflorescences until blooming stage, 2 TDFs (2.08%) were induced. 23 TDFs (23.96%) belonged to Class B or up-regulated expression, a low level of gene expression at control stage was continuously increased until blooming stage. On the other hand, 33 TDFs (34.38%) belonged to Class C, which were continuously decreased or down-regulated and then switch-off of genes expression under natural environment. The remaining 4 TDFs (4.17%) were classified into Class D or early-up and late-down regulated which means that a low level of gene expression increases continuously from the control (vegetative stage) up to stage 1.0 cm-long inflorescence, after that is decreases continuously until the blooming stage.

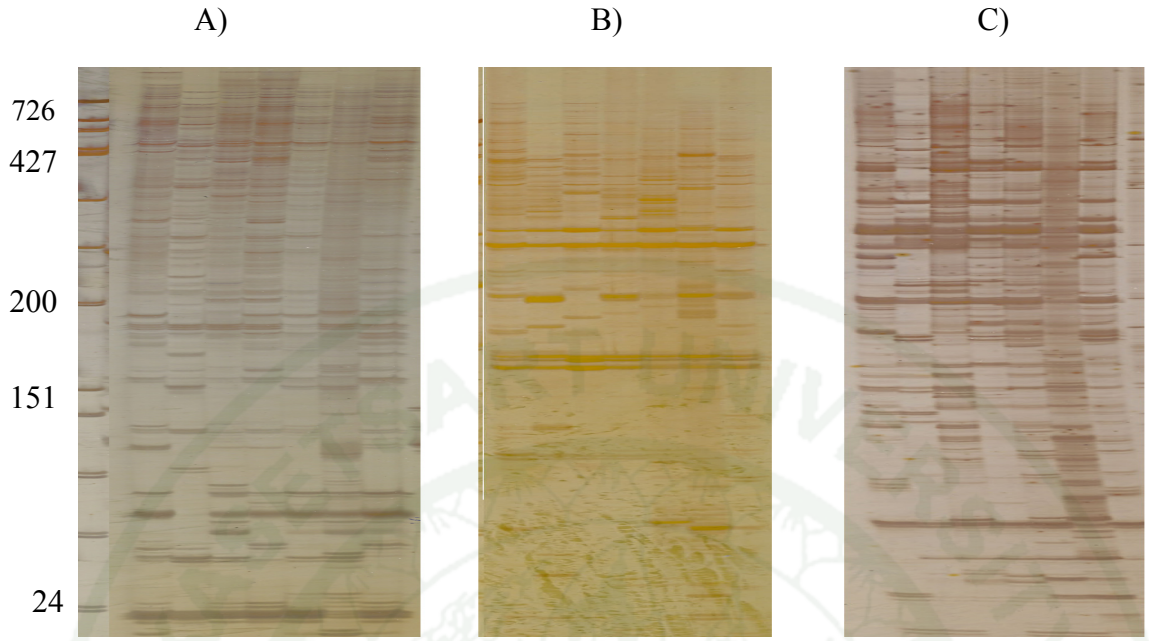


Figure 6 The examples of differentially genes expression profiling using cDNA-AFLP (A) E-GCA/ M-CAA (B) E-TCG/M-AAG (C) E-GCC/M-AAA

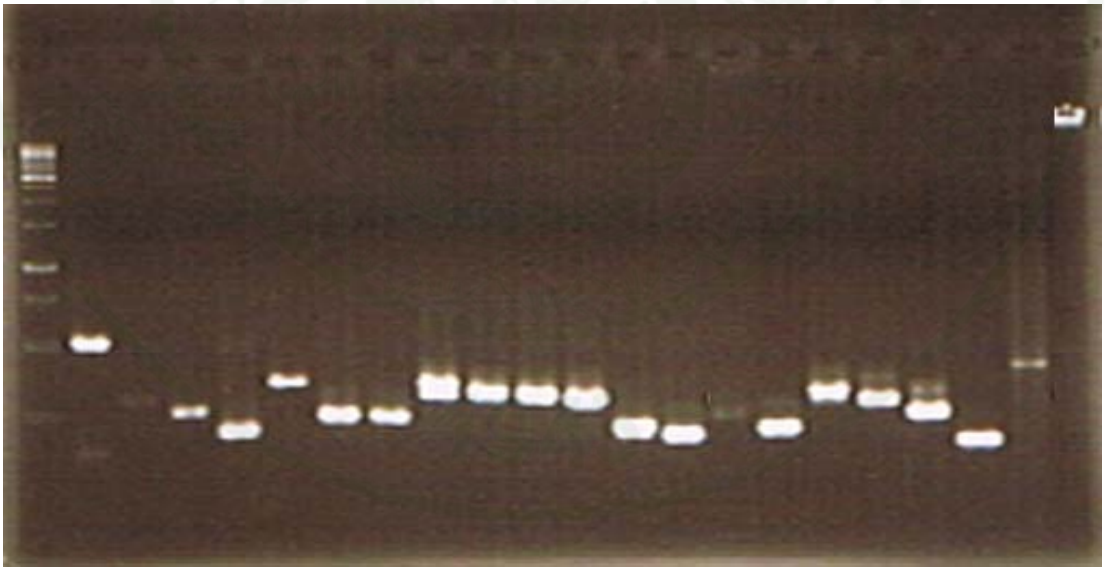
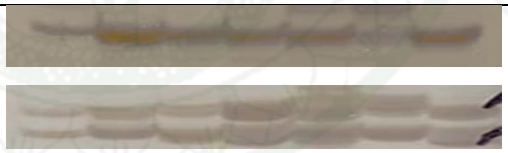
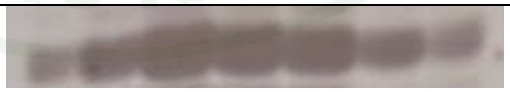


Figure 7 The examples of TDFs after recovery from polyacrylamide gel, re-amplified and eluted. The re-amplified TDFs were separated on 1.2% agarose gel electrophoresis in 1X TBE buffer with 40 V., 40 min.

1.4 TDFs homology analysis

Accordingly from the previous 1.2 section, out of 183, 96 TDFs were successfully excised from polyacrylamide gels, re-amplified with the original primers and then sequenced. The TDFs were ranging in length from 100-600 bp. The homology analysis was performed using BLAST program against with the GeneBank public database. The homology resulting comparison revealed that 61 TDFs showed no homology with any sequences and 35 TDFs showed homology with genes that known functions and unclassified or putative proteins (BLAST expectation values $E < 10^{-3}$).

Table 2 Classification of flower-specific sugarcane TDFs

Class	Expression profile	C 0 0.5 1 3 10 15cm	Number of TDFs
Class A1	Early-induced		34 (35.42%)
Class A2	Late-induced		2 (2.08%)
Class B	Up-regulated		23 (23.96%)
Class C	Down-regulated		33 (34.38%)
Class D	early-up and late-down regulated		4 (4.17%)

The sequences homology analysis was done, and 35 TDFs selected with known function and unclassified or putative protein are shown in the Table 3. For example ClassA1, 4DS_1X showed 94% similar to *IP1* in *S. bicolor*; 5DS_2D showed 93% similar to *epoxide hydrolase 2* in *Z. mays*; Class A2, 5C_2C showed 76% similar to hypothetical protein in *S. bicolor*; Class B, 1.1A_3M showed 93% similar to *CPP* transcription factor in *Z. mays*; Class C, 1.2E_3G showed 97% similar to auxin-independent growth promoter in *Z. mays*; Class D, 3F_1M showed 98% similar to DNA mismatch repair protein (MSH1) in *Z. mays*. Among them, 26 TDFs were selected base on related-flowering trait to test polymorprism in 170 F_5 sorghum RILs and were used to construct sorghum genetic map in the PART I. The specific TDFs primers were designed and showed as in appendix Table 2.

Table 3 Analysis of differentially expressed of flower-specific TDFs sequences homology using BLASTN and BLASTX in sugarcane

TDFs Name	Length (bp)	Accessions	Sequence Similarity	E-value	ID
Class A1					
6S-1E	212	XM002445839.1	*S. bicolor* hypothetical protein, mRNA	1.00E-42	81%
4S_1S	414	NM001111863.1	*Z. mays* MADS box protein (*ZAP1*), mRNA	2.00E-158	94%
3S_1X	281	FL808736.1	*Panicum virgatum* late flowering buds + seed development	3.00E-77	98%
4DS_1X	276	BI140022.1	*S. bicolor* Immature pannicle 1 (*IP1*)	4.00E-101	94%
5S_2B	232	CA299881.1	*Saccharum* hybrid cultivar , mRNA sequence	9.00E-33	98%
1DS_2D	545	BT036187.1	*Z. mays* full-length cDNA mRNA, complete cds	3.00E-165	89%
4DS_2D	316	CA248890.1	*Saccharum* hybrid cultivar FL1, mRNA sequence	9.00E-80	97%
5DS-2D	372	EU965136.1	*Z. mays* epoxide hydrolase 2,mRNA	3.00E-80	93%
4S_3C	150	GH218234.1	*H. vulgare* pre-anthesis spike (white to yellow anther)	4.00E-34	87%
2DS_3E	228	CA203183.1	*S. officinarum* FL1, mRNA sequence	9.00E-46	90%
1S_1C	196	-	No significant but good signal	-	-
3DS_3H	287	-	No significant but good signal	-	-

Table 3 Analysis of differentially expressed of flower-specific TDFs sequences homology using BLASTN and BLASTX in sugarcane (continued)

TDFs Name	Length (bp)	Accessions	Sequence Similarity	E-value	ID
Class A2					
2C-2C	264	AF114171.1	*S. bicolor* BAC clone 25.M18, complete sequence	7.00E-41	86%
5C_2C	137	XP002445018.1	*S. bicolor* hypothetical protein SORBIDRAFT_07g002940	2.00E-09	76%*
Class B					
10A_1E	163	CI160600.1	*O. sativa*, Panicles mixture of 1, 2, 3 weeks after flowering	5.00E-57	100%
1A_3C	337	XM002442182	*S. bicolor* hyprothetical protein, mRNA, similar to mitochondrial import inner membrane translocase	5.00E-148	98%
2A_3C	272	AY436773.1	*Pyrus communis* putative senescence-associated protein mRNA	5.00E-11	91%
2.1A_3C	252	BF656217.1	*S. propinquum* cDNA Floral-Induced Meristem 1 (FM1), mRNA sequence	9.00E-15	94%
3A_3E	225	XM002441945	*S. bicolor* hypothetical protein similar to transposon protein,	3.00E-96	100%
1.1A_3M	320	NM001176095	*Z. mays* CPP transcription factor, mRNA	1.00E-58	93%
1.2A_3M	235	XM002450475	*S. bicolor* hypothetical protein, mRNA	5.00E-54	88%

Table 3 Analysis of differentially expressed of flower-specific TDFs sequences homology using BLASTN and BLASTX in sugarcane (continued)

TDFs Name	Length (bp)	Accessions	Sequence Similarity	E-value	ID
Class C					
2D_1U	212	XM002457048.1	*S. bicolor* short-chain dehydrogenases, hypothetical protein,	1.00E-109	90%
3E_2D	288	GU080320.1	*Saccharum* hybrid cultivar R570	6.00E-36	85%
3E_2R	343	XM002442872.1	*S. bicolor* hyprothetical protein, mRNA	3.00E-148	98%
3E_3C	134	EF115542.1	*Saccharum* hybrid cultivar chloroplast, complete genome	2.00E-44	100%
2.2E_3D	216	NM001112334.1	*Z. mays* protein disulfide isomerase8 (*pdi8*), mRNA	1.00E-51	84%
1.1E_3E	245	CA228083.1	*S. officinarum* FL3, mRNA sequence	5.00E-75	97%
3.2E_3F	252	CF489804.1	*S. bocolor* pollen, mRNA sequence	9.00E-15	77%
1.2E_3G	444	NM001154623	*Z. mays* auxin-independent growth promoter cDNA clone	0	97%
5E_3G	243	XM002462949.1	*S. bicolor* hypothetical protein, mRNA	8.00E-40	88%
2E_3H	280	DAA38636.1	*Z. may* TPA: hypothetical protein ZEAMMB73_143695	3.00E-21	84%*
2E_3I	373	BF421254.1	*S. propinquum* Floral-Induced Meristem 1 (FM1)	5.00E-64	94%
Class D					
3F_1M	418	NM001112428.1	*Z. mays* DNA mismatch repair protein (*MSH1*), mRNA	4.00E-84	98%
6F_1Y	185	XM002460123.1	*S. bicolo*r hyprothetical protein, mRNA	3.00E-97	98%

.*BLASTX

1.5 Validation of differentially TDFs expression using RT-PCR

The 4S_1S sugarcane putative flowering TDFs was selected to analyze the expression using RT-PCR to validate the results of the cDNA-AFLP experiment. The 4S_1S belongs to Class A1, early induced and showed 94% similar to *Z. mays* MADS box protein (*ZAP1*).The RT-PCR result showed that the expression of 4S_1S corresponding with the Class A1 of classification of flower-specific sugarcane (Figure 8)

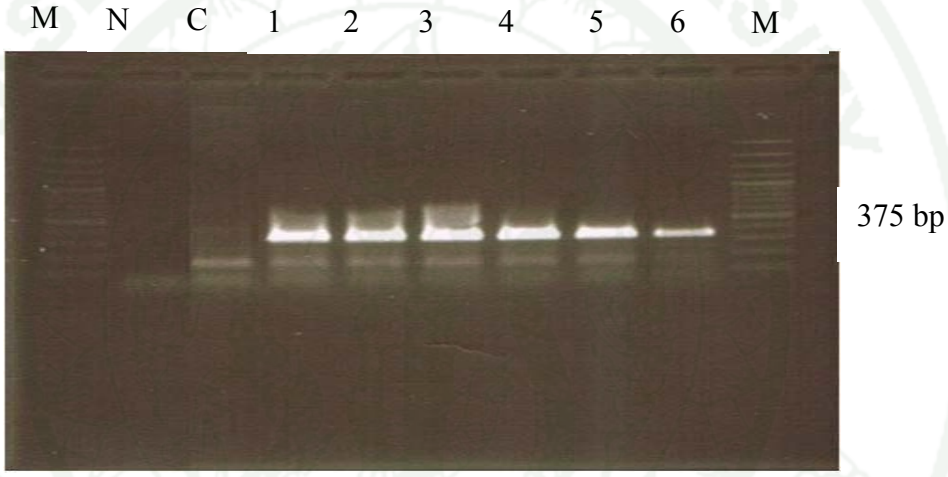


Figure 8 RT-PCR showing the expression of 4S_1S at the different stage of inflorescences development in sugarcane

Where,

M 100 bp ladder plus marker

N Negative control (dH_20)

C Internal control which is 4 months old of shoot apex sugarcane

1 0.0 cm of inflorescence length

2 0.5 cm of inflorescence length

3 1.0 cm of inflorescence length

4 3.0 cm of inflorescence length

5 10.0 cm of inflorescence length

6 15.0 cm of inflorescence length

2. Sorghum genetic linkage map construction and correlation with flower-specific sugarcane TDFs

2.1 Genetic linkage map construction

Of the 1,601 markers were screened using parental lines to identify polymorphic bands with clear polymorphic pattern between the parents. 348 generated polymorphisms between patents. Of the 348 markers, 196 showed co-dominant polymorphic, 10 sugarcane TDFs showed dominant polymorphic (Table 4). Among them, 9 sugarcane TDFs with 12 marker loci showed polymorphism in sorghum population. They were also used for sorghum genetic linkage map construction (Table 5). Finally, the 206 polymorphic markers were used for the construction of the genetic linkage maps and mapping of the QTLs controlling some agronomic traits in the F_5 RILs population.

Table 4 Parental marker screening

Group of Primers	Total of Primers	Co-dominant Polymorphisms	Dominant Polymorphism
Xcup	117	15	5
SAM	414	44	27
Xsbarblk	422	46	40
TX	144	42	14
LBK	80	-	14
Stay Green	48	1	5
Drenhsbm	108	20	16
Xtxtp	38	21	3
Zea mays-SSR	177	4	13
QTL-Flower-Specific SSR and Transcription factor	27	1	5
Sugarcane TDFs	26	2	10
Total	1,601	196	152

Table 5 List of flower-specific sugarcane TDFs for sorghum genetic linkage map construction

No.	TDFs	Marker Loci	Polymorphic pattern
1	sfw3F_1M	sfw3F.1M	dominant
2	sfw4DS_1X	sfw4DS.1X	Co-dominant
3	sfw6F_1Y	sfw6F.1Y	dominant
4	sfw2A_3C	sfw2A.3C	dominant
5	sfw2DS_3E	sfw2DS.3E	dominant
6	sfw3.2E_3F	sfw3.2E.3F	Co-dominant
7	sfw1.2E_3G	sfw1.2E.1.3G	dominant
		sfw1.2E.2.3G	dominant
8	sfw2E_3H	sfw2E.3H	dominant
9	sfw3DS_3H	sfw3DS.1.3H	dominant
		sfw3DS.2.3H	dominant
		sfw3DS.3.3H	dominant

Eight TDFs marker corresponding with 9 loci were mapped into five linkage groups, sfw3.2E_3F on SBI-1a; sfw3DS.1_3H and sfw3DS.2_3H on SBI-1b; sfw4DS_1X and 2DS_3E on SBI-02; sfw3F_1M, sfw2E_3H and sfw2A_3C on SBI-3b and sfw1.2E.1_3G on SBI-05. In addition, three TDFs marker loci could not be mapped to any linkage group due to unlinked by genotypic data and recombination frequency at 0.5. Accordingly, *Flo08* marker, which is *FLORICAULA/LEAFY* (*FLO/LFY)* orthologs transcription factor, was mapped to SBI-4b at 115.7 cM of length.

2.2 QTL analysis

2.2.1 Phenotypic variation

The phenotypic variation varied widely among the RILs population and their parents. The RILs recorded population phenotypic values mean 123.10 days for days to flowering, 325.80 cm for plant height and 78.44% for chinch bug resistance (Table 6). The normality testing indicated that the data was fairish for QTL analysis, and also showed normal frequency distribution for all traits (Figure 9, 10, 11).

Table 6 Mean phenotypic values of F_5 RILs and their parental cultivars for the studied traits

Traits	Parental lines		RILs Population	
	DDYM	Mapila	Mean	SD
Days to flowering (days)	73	200	123.10	46.21
Plant height (cm)	170	360	325.80	69.03
Chinch bug resistance (%)	40	100	78.44	13.60

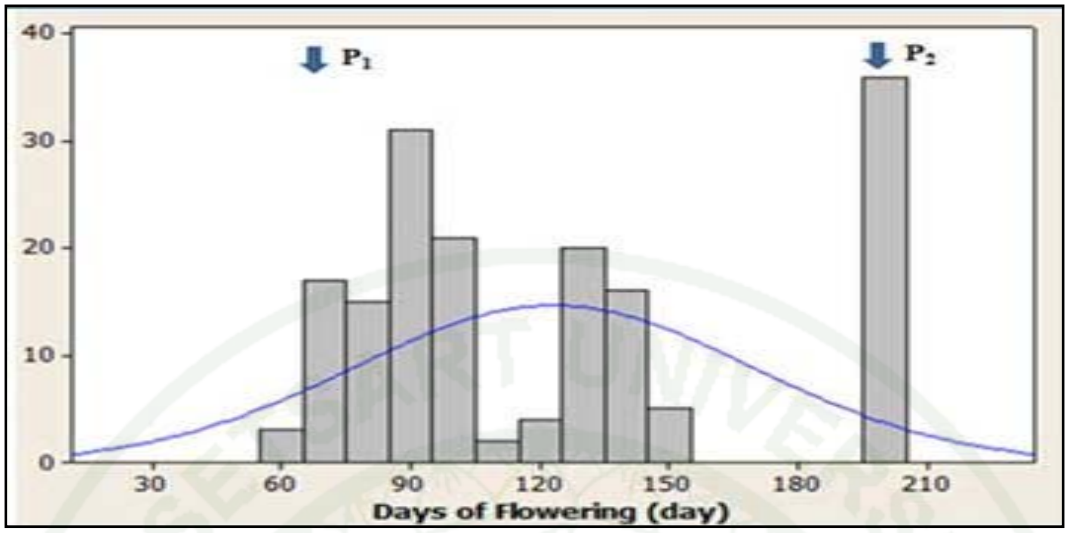


Figure 9 Frequency distribution of F_5 RILs population and their parents for day to flowering (Flo) (P_1: DDYM, P_2: Mapila)

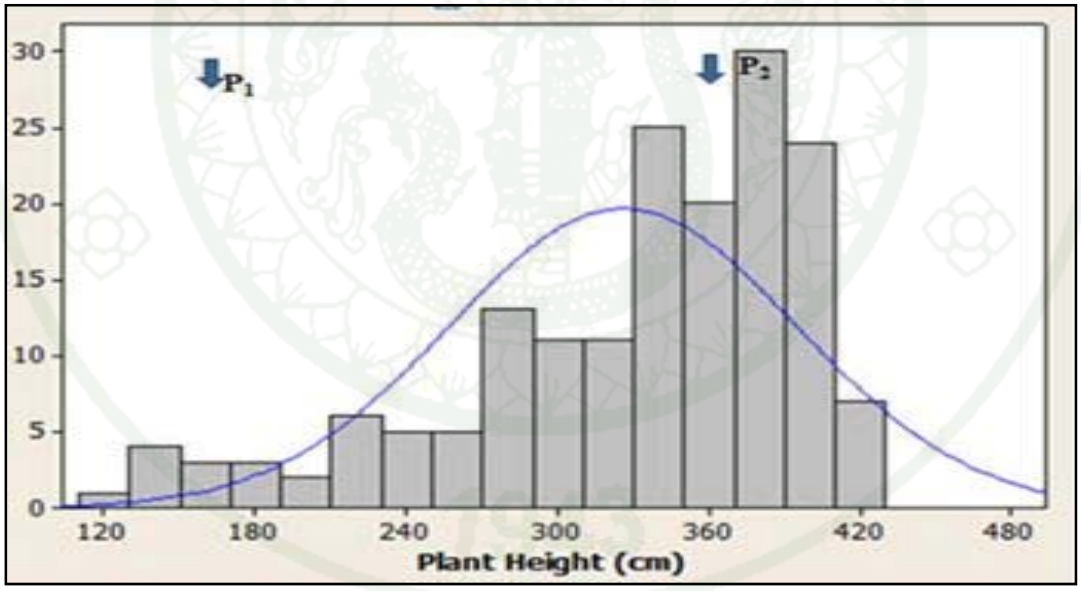


Figure 10 Frequency distribution of F_5 RILs population and their parents for plant height (PH) (P_1: DDYM, P_2: Mapila)

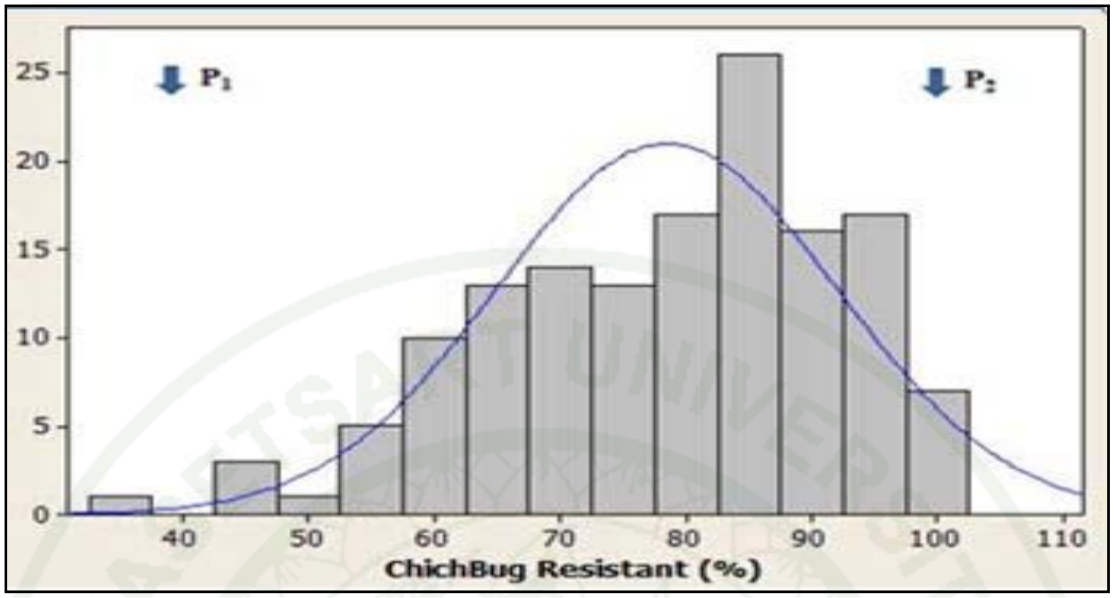


Figure 11 Frequency distribution of F_5 RILs population and their parents for chinch bug resistance (CBR) (P_1: DDYM, P_2: Mapila)

2.2.2 QTL analysis

The effects of QTLs were identified by composite interval mapping (CIM). The LOD threshold ranged from 2.1 to 4.9 was used to declare the QTLs for all traits. The results of significant QTLs analysis are showed in Table 7 and Figure 12.

Days to flowering

CIM detected two QTLs, Flo-1 and Flo-2 on SBI-1b and SBI-4b associated with flowering dates at LOD ≥ 2.5. The phenotypic variation accounted by significant QTLs was 23.92% and their additive effects were 12.92 and -19.27, respectively. The Flo-1 QTL significant maker is SAM18581 at 2.9 LOD score, while Flo-2 QTL associated near TX124 marker at 4.7 LOD score. The major Flo-2 QTL marker showed 16.18% of phenotypic variation, and had negative additive effect which is corresponding with quantitative non-flowering allele with the Mapila male parent.

Plant height

A total of four QTLs collectively accounted for 33.22 % of phenotypic variation in height and detected in 3 Linkage groups associated on SBI-1b, SBI-4b and SBI-9b whereas the additive effects was 18.33 cm for DDYM and ranged from -18.87 to -24.82 cm for Mapila parent. Most of the QTLs had negative additive effect which is corresponding that the parent Mapila had contributed for height alleles.

Chinch bug resistance

Chinch bug is a native North America insect that can destroy grass crop, especially sorghum and corn. 39.88% of 4 significant QTLs, CBR-1 and CBR-2 on SBI-1b, CBR-3 on SBI-02 and CBR-4 on SBI-3a were detected for phenotypic variation and their additive effects were -6.08, 7.42, 6.69 and 4.57 respectively. The CBR-2 on SBI-1b at position 92.12 cM explained 13.60% of phenotypic variation and showed positive additive effect indicating that the increase in resistance was contributed by male parent, Mapila. Interestingly, sfw3DS-1.3H marker linked with CBR-1 QTL at 9.63% of phenotypic variation on SBI-1b as well.

2.3 Gene annotation

For comparative mapping confirmation of TDFs, nine TDFs markers loci were tested for colinearity using chromosome homology with *S. bicolor*, *Z. mays* and *O. sativa* using the Gramene database with BLASTX algorithm. six out of nine TDFs were located corresponding between sorghum genetic linkage mapping and chromosome homology such as sfw4DS_1X located on SBI-1b and showed homology with chromosome 1 of *S. bicolor* by colinearity testing, and also homology with chromosome 2 and 7 of *Z. mays* and *O. sativa*, respectively. While 2 TDFs marker with 3 loci, swf3DS.1.3H, sfw3DS.2.3H and sfw3.2E_3F showed no homology with any plant chromosome (BLASTX expectation value [*E*] < $10e^{-10}$ and 40% of amino acid identity). All data are shown in Table 8.

Table 7 QTLs identification in 170 F_5 RILs of DDYM x Mapila for days to flowering (Flo), plant height (PH), chinch bug resistance (CBR)

Trait	QTL	LG	LOD	Flanking marker	Position (cM)	R^2 (%)	Additive effect	Type of QTL
Day to flower	Flo-1	1b	2.9	SAM18581*-Drenhsdm63	44.81	7.74	12.92	minor
	Flo-2	4b	4.7	Xsbarslbk4.62-Tx124*-Drenhsbm72	83.32	16.18	-19.27	major
					Total	23.92		
Plant height	PH-1	1b	3.1	Xsbarslbk1.39*- Xsbarslbk1.40	4.01	6.74	18.33	minor
	PH-2	4b	3.0	Xsbarslbk4.12*- Xtxtp12	2.01	7.35	-18.87	minor
	PH-3	9b	2.4	Xsbarslbk9.45*- Xsbarslbk9.07	8.01	7.13	-19.05	minor
	PH-4	9b	3.9	Drenhsbm17*-Xsbarslbk9.55	70.91	12.00	-24.82	major
					Total	33.22		
Chinch bug resistance	CBR-1	1b	2.9	Drenhsbm63- sfw3DS.1. 3H*	49.37	9.63	-6.08	minor
	CBR-2	1b	4.2	TX106*-TX157	92.12	13.60	7.42	major
	CBR-3	2	3.4	TX127*-Xtxtp07	36.77	10.40	6.69	major
	CBR-4	3a	2.3	Xtxtp09-SAM16073*	14.76	6.25	4.75	minor
					Total	39.88		

*Nearest marker

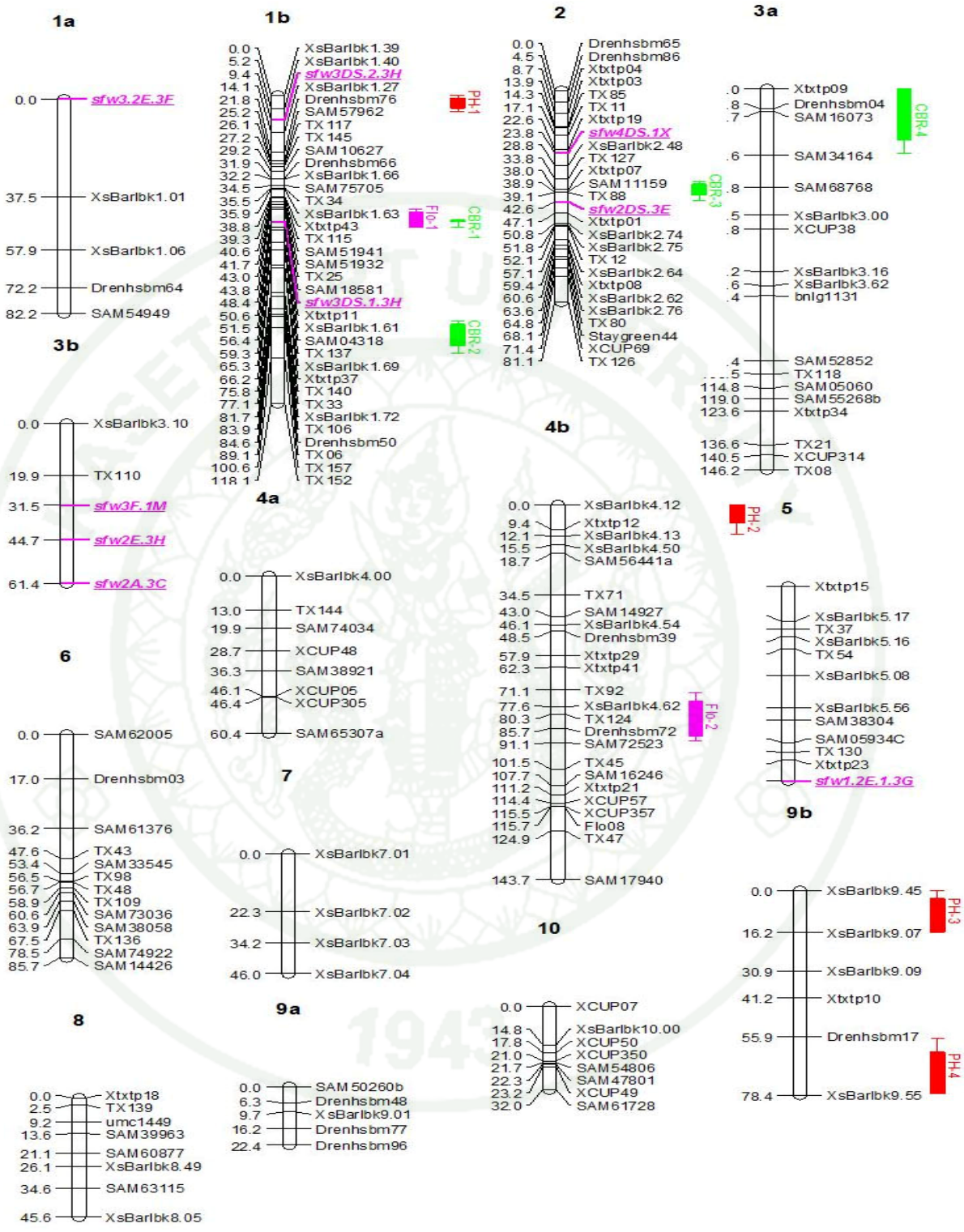


Figure 12 Location of QTLs associated with agronomic traits on a sorghum genetic linkage map based on F_5 RILs population of DDYM x Mapila. Blue box represented day to flowering, red box represented to plant height and green box represented to chinch bug resistance.

Table 8 Colinear chromosomal analysis of flower-specific markers using BLASTX algorithm.

TDFs Name	Plant	Chromosome Homology	*E-Value*	%ID
3F_1M	*S. bicolor*	3	2.20E-28	91.04
	Z. mays	10	2.90E-28	92.54
	O. sativa	4	4.30E-27	89.55
4DS_1X	*S. bicolor*	2	1.80E-28	90.16
	Z. mays	2	1.90E-29	90.16
	O. sativa	7	6.80E-24	81.67
2A_3C	*S. bicolor*	3	1.40E-20	92.31
	Z. mays	9	0.0005	91.91
	O. sativa	6	5.90E-12	65.31
2DS_3E	*S. bicolor*	2	1.30E-15	90
	Z. mays	1	5.50E+00	30
	O. sativa	12	6.90E-11	70
2E_3H	*S. bicolor*	3	1.60E-11	76.92
	Z. mays	2	1.90E-18	84.31
	O. sativa	1	1.50E-09	59.26
3DS_3H	*S. bicolor*	N*	N	N
	Z. mays	N	N	N
	O. sativa	3	8.7	100

N* No similarity

Gene prediction: sfw4DS.1X candidate flowering gene

As the result, the excellent sugarcane TDF, sfw4DS.1X located on SBI-02 or sorghum chromosome 2, was also analyzed the sequence to predict the function. Gene prediction is a preliminary technique to annotate the sequence before downstream can be done. sfw4DS.1X showed 94% high homology with *Immature Pannicle* or *IP1* that derived from *Preanthesis Panicle* of sorghum using BLAST analysis. It's also showed high colinear localized within many plant species such as *Arabidopsis thaliana*, *S. bicolur*, *Zea mays*, *Oryza sativa* and *Glycine max* as showed in Figure 13. For prediction as a gene against *A. thaliana* was done by AUGUSTUS software. The result showed that the sfw4DS.1X is a part of gene that containing a coding region, can be transcribed as a gene and protein translation as well (Figure 14, 15).

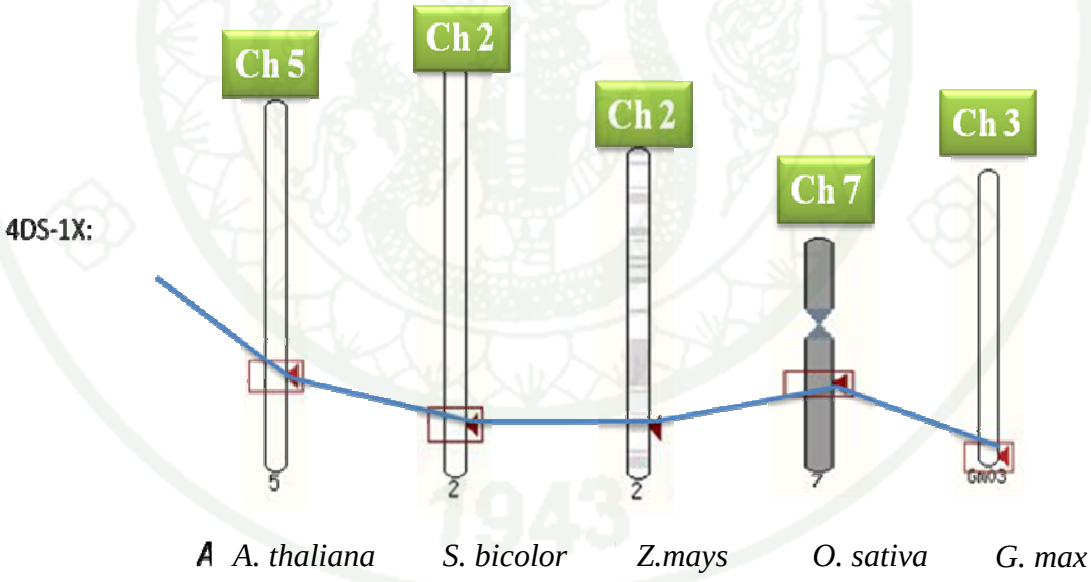


Figure 13 Chromosome homology analysis of sfw4DS.1X using hit filter criteria *E-value* threshold of $10e^{-10}$ and 40% minimum amino acid identity.

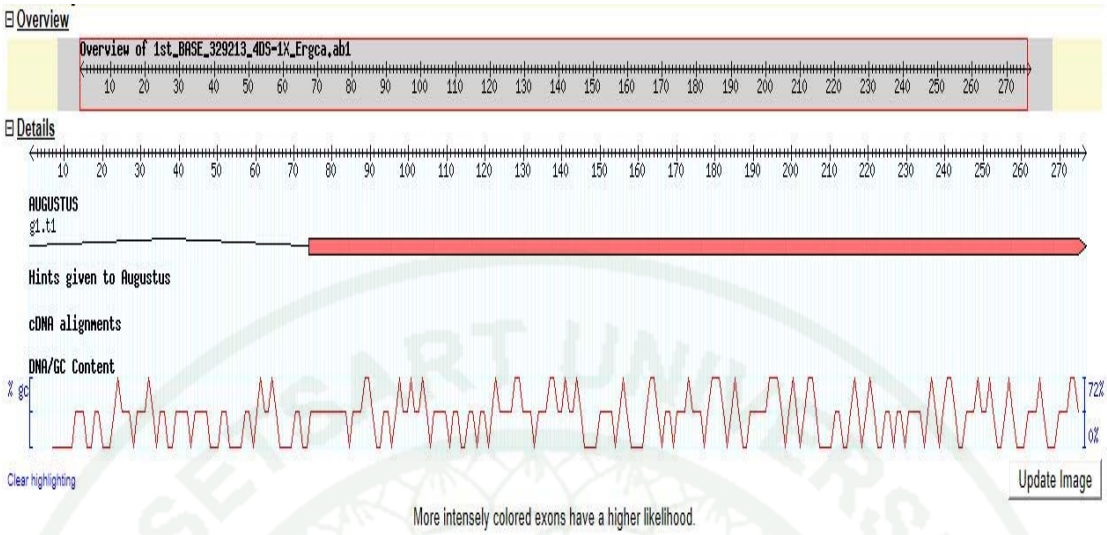


Figure 14 The land mark region of sfw4DS.1X using AUGUSTUS gene prediction program

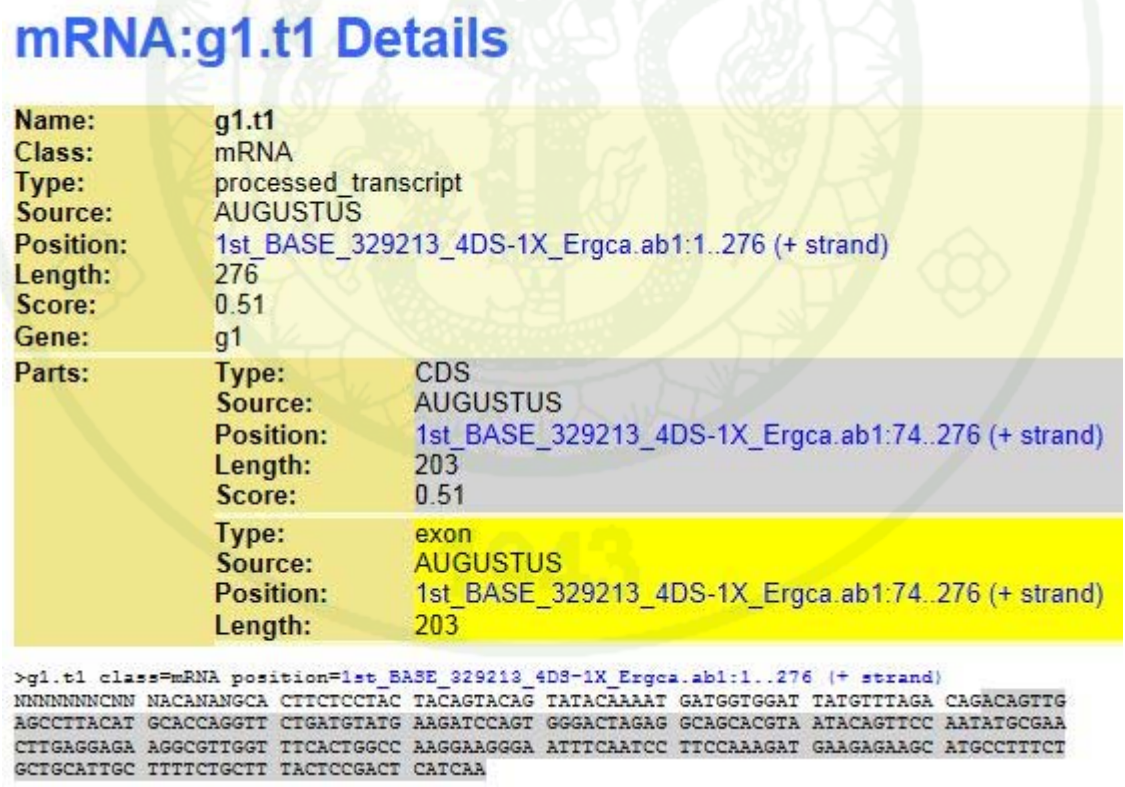


Figure 15 The region of mRNA sfw4DS.1X using AUGUSTUS gene prediction program

DISCUSSION

There are many molecular markers e.g. AFLP, SSR and SNPs that have been used to study genetic regions controlled agronomic traits in plant. Segregating populations such as recombinant inbred lines are needed in order to generate genetic map that linked to quantitative genes. Sugarcane and sorghum belong to the same subtribe, and share common ancestor more than 5 million years ago. While sorghum has a simple diploid genome (2n=20) and sugarcane has complex genome (*S. officinarum*, 2n=80 and *S. spontaneum*, 2n=40-128 [Grivet and Arruda, 2001]), which suggests that sorghum has the potential to be used as a template for genomic study in sugarcane.

In order to develop specific molecular marker, cDNA-AFLP analysis and colinearity using genetic linkage mapping were investigated to identify the differentially genes that linked to flowering in sugarcane using sorghum as a template. In this study, the cDNA-AFLP technique was proven to be rapid and useful tool to identify differentially expressed flowering genes from the inflorescences of wide sugarcane. The expression level of homology transcripts, 4S_1S and 4DS_1X TDFs that belong to ClassA1 showed homology with *ZAP1* in *Z. mays* and *IP1* in *S. bicolor,* respectively, where those regions are well known to be involved with floral architecture and flower development. As well as, the expression of 1.1A_3M TDF which encodes *CPP* transcription factor, *CPP*-like gene plays an important role in reproductive tissue development and control of cell division in plants (Yang *et al*., 2008), belongs to Class B. Moreover, 1.2E_3G TDF encodes auxin-independent growth promoter in *Z. mays*, which has been reported to be involved in flower formation at the boundary of the reproductive shoot apex (Yamaguchi *et al*., 2013). These results suggest that the cDNA-AFLP technique provides a potential genome wide transcript profiling to indentify the differentially genes expression in different mRNA. In our study, 9 out of 96 TDFs were selected as the flowering-specific sugarcane TDFs for mapping construction according to their putative function and homology level of significant.

QTLs of sorghum agronomic traits were investigated in 170 F_5 recombinant inbred lines (RILs) population that derived from a cross between two sorghums, Double Dwarf Yellow Milo (DDYM) and Mapila. A total of 3 traits were evaluated including days to flowering, plant height and chinch bug resistance. A genetic map was constructed with 178 markers loci on 14 linkage groups collectively spanning 1077.8 cM that correspond to the 10 sorghum chromosomes. CIM identified 2 flowering QTLs, Flo-1 and Flo-2 explaining a range of 7.74-16.18% of phenotypic variation on SBI1b and SBI4b, respectively. Other studies have detected several QTLs with RILs population such as SBI-01, SBI-02, SBI-03, SBI-05, SBI-06, SBI-07 and SBI-08 (Srinivas *et al*., 2009), SBI-4a, SBI-06, SBI-07 and SBI-08 (Amukelani *et al*., 2010), SBI-1b (Legrari, 2010), Chromosome 4 (Kong *et al*., 2013), SBI-01, SBI-03, SBI-05, SBI-07, SBI-09 and SBI-10 (Reddy *et al.*, 2013).

As known as, sorghum is a short day plant, yet there exist considerable genotypic variations exist in photoperiodic condition for conversion from the vegetative to reproductive phase. Adaptation to photoperiod insensitivity, in temperate climates, for early maturity requires a minimum of 6 major loci (Quinby, 1974): Ma_1-Ma_6. Ma_1-Ma_4 has shown to promote flowering under shorter photoperiods while inhibiting flowering during longer photoperiods. Ma_5-Ma_6 both strongly inhibit flowering under all photoperiod conditions when both dominant alleles are present (Childs *et al*., 1997). Ma_1 is most responsible for photoperiod insensitivity and was identified as PRR37 (Murphy *et al*., 2011). Ma_3 was shown to encode PhytochromeB (Childs *et al*., 1997).

Plant height plays a critical role in nutrient responsiveness, lodging resistance, and efficiency of harvesting. Plant height correlates significantly with total biomass (Yuan *et al*., 2008), determined by plant height and stem girth. Potential utilization of sorghum as a biofuel crop requires genetic manipulation of sweet sorghum lines for increased biomass. Reduction of plant height is advantageous for machine harvest, wind avoidance and other hazards. Genetic control of plant height is attributed to the effects of four major loci: Dw_1, Dw_2, Dw_3, and Dw_4 (Quinby and Karper, 1954). Previous studies have consistently identified plant height QTL, in various

backgrounds, with major effects on two loci: Dw_1 on SB1-06 and Dw_3 on SB1-07 (Feltus *et al*., 2006; Klein *et al*., 2008; Mace and Jordan, 2011). In this study, four QTLs genomic areas with major effect on plant height were identified. Of these areas, region on SBI-09 was in agreement with the Dw_1 locus, proposed by an earlier study (Brown *et al*., 2008), encoding a major height gene Sb.Ht9.1. The remaining did not align with the Dw_2 and Dw_3 dwarfing genes identified in previous studies. This difference may be attributed to an inability to differing methods of creating maps i.e SSR vs DArT, single marker vs composite interval maker analysis, etc.

Chinch bug, *Blissus leucopterus* (Say) (Heteroptera: Blissidae) is a native North American piercing- sucking insect. Its ability to produce multiple generations in a growing season, makes chinch bug a major source of injury and loss to sorghum stands (Hudson, 1995).The annual cost of loss due to chinch bug is $11million in Nebraska (Rajewski *et al*., 2009). Genetic variation with high inheritance for resistance to chinch bug exists in sorghum and other grasses (Wilde Morgan). This study represents the first mapped QTLs for chinch bug resistance in sorghum. Four QTLs were identified with additive effects for chinch bug resistance, on SBI-1b, SBI-02 and SBI-3a in sorghum. Interestingly, sfw3DS.1.3H sugarcane TDF also showed significantly marker linked to CBR-1 QTL at 9.63% of phenotypic variation. While this is a novel study, several studies were conducted in greenbug for resistance QTLs. Previous QTL mapping efforts in greenbug resistance have shown a major QTL region on SBI-09 (Punnuri *et al*., 2013; Wu and Huang, 2008). Punnuri *et al*., 2013) also found a minor QTL region on SBI-03 for green bug damage resistance.

Eight sugarcane TDFs markers corresponding with 9 loci were mapped into five sorghum linkage groups. The sfw4DS.1X and sfw2DS.3E on SBI-02 linkage group, which is located at 23.8 and 42.6 cM, respectively and also corresponding with sorghum chromosome 2 (calculated using chromosome homology analysis). The sfw4DS.1X showed high homology with flowering trait characterization in sorghum *Immature Panicle* (*IP1*) while sfw2DS.3E showed high homology with FL_1 (flower at 1 cm) of sugarcane. Moreover, sfw4DS.1X could be a part of gene that related to flowering and showed codominant expression in sorghum RILs population. This

means it carries genotypic value of both parents and can be served as a candidate specific marker for breeding selection, while sfw2DS.3E showed dominant expression that good enough for breeding program as well.

In addition, flo08 marker or *FLORI CAULA/LEAFY* (*FLO/LFY)*-like gene that generated from 29 Andropogorreae species in 18 genera, including sugarcane and sorghum and one out-group, *Arundinella hirta* (Kiesten and Doebley, 2005) was mapped to SBI-4b. *FLO* and *LFY* ortholog are one of the key regulation genes in flower development by promoting the reproductive transition in *Antirrhinum majus* (Coen *et al.*, 1990) and *Arabidopsis thaliana* (Weigel *et al.*, 1992), respectively. So, Flo08 marker can be served as a specific marker for flowering trait selection because it showed codominant expression in sorghum RILs that were derived from a cross between early flowering cultivar and none flowering sorghum line. Base on this study, it may indicate that *FLO/LFY*-like gene could be located in SBI-4b or sorghum chromosome 4.

Although, the fine map or complete sequencing of the sugarcane genome still a possibility in near future, the sorghum was successful complete five year ago (Paterson *et al*., 2009). The distributions of sorghum and sugarcane linkage groups were in close agreement (Grivet *et al*., 1994). So, sorghum knowledge becomes a highly valuable resource for genomic study in sugarcane and other C_4 plants. Here, we were successful to study putative flowering genes in sugarcane using genetic colinearity of both sugarcane and sorghum.

CONCLUSIONS

1. The development of flower transcriptome profiling was done using cDNA-AFLP with 26 primer combinations. A total of 183 transcript-derived fragments (TDFs) were screened. 96 TDFs were sequenced. 26 out of 96 TDFs were selected and used to be flowering putative genes for study colinerity with sorghum genome.

2. A sorghum genetic map was constructed with 178 markers loci on 14 linkage groups collectively spanning 1077.8 cM that correspond to the 10 sorghum chromosomes.

3. Eight sugarcane TDFs markers corresponding with 9 loci were mapped into 5 sorghum linkage groups.

4. CIM identified 2 flowering QTLs, Flo-1 and Flo-2 explaining a range of 7.74 – 16.18% of phenotypic variation on SBI-1b and SBI-4b, respectively.

5. This study represents the first mapped QTLs for chinch bug resistance in sorghum. Four QTLs were identified with additive effects for chinch bug resistance, on SBI-1b, SBI-2 and SBI-3a in sorghum. Interestingly, sfw3DS.1.3H sugarcane TDF also showed significantly marker linked to CBR-1 QTL at 9.63% of phenotypic variation.

6. The excellent sugarcane TDF, sfw4DS.1X located on SBI-02 or sorghum chromosome 2, showed 94% high homology with *IMMATURE PANNICLE* or *IP1* that derived from preanthesis panicle of sorghum using BLAST analysis. It's also showed high collinear localized within many plant species such as *A. thaliana*, *S.bicolor*, *Z. mays*, *O.sativa* and *G. max*. The result indicated that sfw4DS.1X is a part of gene that containing a coding region, can be transcribed as a gene and protein translation as well.

LITERATURE CITED

Adams, M.D., J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, C.R. Merril, A. Wu, B. Olde and R.F. Moreno. 1991. Complementary DNA sequencing:Expressed sequence tags and human genome project. **Science** 252: 1651–1656.

Agarwal, M., N. Shrivastava and P. Harish. 2008. Advances in molecular marker techniques and their applications in plant sciences. **Plant Cell Rep**. 27: 617–631.

Agrama, H.A., G.E. Widle, J.C. Reese, L.R. Campbell and M.R. Tuinstra. 2002. Genetic mapping of QTLs associated with greenbug resistance and tolerance in *Sorghum bicolor*. **Theor. Appl. Genet** 104: 1373-1378.

Aikhionbare, F.O., K.P. Pruess and Z.B. Mayo. 1998. Greenbug (Homoptera: Aphididae) biotypes characterized using random amplified polymorphic DNA. **Genet. Analysis-Biomolec Engineer.** 14: 105–108.

Akkaya, M.S., A.A. Bhagwat and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. **Genetics** 132: 1131-1139.

Al-Janabi, S.M., M. McClelland, C. Petersen and B.W. Sobral, 1994. Phylogenetic analysis of organellar DNA sequences in the Andropogoneae: *Saccharine.* **Theor. Appl. Genet**. **88:** 933–944.

Amukelani, L.S., M. Frisch and W. Friedt. 2010. Genetic mapping of QTLs for sugar-related traits in a RIL population of *Sorghum bicolor* L. Moench. **Theor. Appl. Genet.** 121: 323-336.

Apotikar, D.B., D. Venkateswarlu, R.B. Ghorade, R.M. Wadaskar, J.V. Patil and P.L. Kulwal. 2013. Mapping of shoot fly tolerance loci in sorghum using SSR markers. **J. Genet.** 90: 59-66.

Bachem, C.W.B., R.J.F.J., Oomen, and R.G.F. Visser. 1998. Transcript imaging with cDNA-AFLP: a step-by-step protocol**. Plant Mol. Biol. Rep**. 16: 157-173.

Bachem, C.W.B., van der R.S. Hoeven, de S.M. Bruijn, D. Vreugdenhil, M. Zabeau and R.G.F. Visser. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. **Plant J**. 9:745–753.

Becker, A. and G. Theissen. 2003. The major clades of MADS-box genes and their role in the development and evolution in flowering plants. **Mol. Phylogenet. Evol**. 29(3): 464-89.

Bhattramakki, D., J. Dong, A. K. Chhabra, and G. E. Hart. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. **Genome** 43:988 - 1002.

Bian, Y.L., Y. Seiji, I. Maiko and C. Hong-Wei. 2006. QTLs for sugar content of the stalk in sweet sorghum (*Sorghum bicolor* L. Moehch.). **Agric. Sci. China.** 5(10): 736-744.

Black, W.C.I.V., N.M. DuTeau, G.J. Puterka, J.R. Nechols and J.M. Pettorini. 1992. Use of the random amplified polymorphic DNA polymerase chain-reaction (RAPD-PCR) to detect DNA polymorphisms in aphids (Homoptera, Aphididae). **Bull. Entomol. Res**. 82: 151–159.

Brown, P.J., W.L. Rooney, C. Franks and S. Kresovich. 2008. Efficient mapping of plant height quantitative trait loci in a sorghum association population with introgressed dwarfing genes. **Genetics** 180: 629–637.

Brown, S.M., M. S. Hopkins, S. E. Mitchell, M. L. Senior, T.Y. Wang, R.R. Duncan, F. Gonzalez-Candelas and S. Kresovich. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. **Theor. Appl. Genet.** 93: 190-198.

Burrow, G.B., C.D. Franks, V. Acosta-Martinez and Z. Xin. 2008. Molecular mapping and characterization of BLMC, a locus for profuse wax (bloom) and enhanced cuticular features of sorghum (*Sorghum bicolor* (L.) Moench.), **Theor. Appl. Genet**. 118: 423- 431.

Chang, S., J. Puryear, and J. Cairney. 1993. A simple and efficient method for isolating RNA from pine trees. **Plant Mol. Biol. Rep**. 11: 113-116.

Childs, K.L., F.R. Miller, M.M. Cordonnier-Pratt, L.H. Pratt, P.W. Morgan and J.E. Mullet. 1997. The sorghum photoperiod sensitivity gene, encodes a phytochome B1. **Plant Physiol**. 113: 611–619.

Coelho, C.P., A.P. Costa Netto, J. Colasanti and A. Chalfun-Junior. 2013. A proposed model for the flowering signaling pathway of sugarcane under photoperiodic control. **Genet. Mol. Res**. 12(2): 1347-1359.

Coen, E.S. and E.M. Meyerowitz. 1991. The war of the whorls: Genetic interactions controlling flower development. **Nature** 353: 31-37.

Coen, E.S., J. M. Romero, S. Doyle, R. Elliott, G. Murphy and R. Carpenter. 1990. Foricaula: a homeotic gene required for flower development in *Antirrhinum majus*. **Cell** 63: 1311–1322.

Colasanti, J. and V. Coneva. 2009. Mechanisms of floral induction in grasses: something borrowed, something new. **Plant Physiol**. 149: 56-62.

Cuadrado, A., R. Acevedo, S. Moreno Diaz de la Espina, N. Jouve and C. de la Rorre. 2004. Genome remodelling in three modern *S. officinarum* x *S. spontaneum* sugarcane cultivars**. J. Expt. Bot**. 55 (398): 847-854.

Curtis, D.L. 1968. The relationship between the date of heading of Nigderian sorghums and the duration of growing season. **J. Applied Ecol**. 5: 215-226.

Deu, M., A. Ratnadass, M.A. Hamada, J.L. Noyer, M. Diabate and J. Chantereau. 2005. Quantitative trait loci for head-bug resistance in sorghum. **Afr. J. Biotechnol**. 4 (3): 247-250.

Durrant, W.E., O. Rowland, P. Piedras, K.E. Hammond-Kosack and J.D.G. Jones. 2000. cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. **Plant Cell** 12: 963–977.

Dweikat, I. 2005. A Diploid, interspecific, fertile hybrid from cultivated sorghum, *Sorghum bicolor*, and the common johnsongrass weed *Sorghum halepense*. **Mol. Breeding.** 16: 93-101.

Fang, G., S. Hammar and R. Grumel. 1992. A quick and inexpensive method for removing polysaccharides from plant genomic DNA. **Biotechnique** 13: 52-55.

Farrell, A.E., R.J. Plevin, B.T. Turner, A.D. Jones, M. O'Hare and D.M. Kammen. 2006. Ethanol can contribute to energy and environmental goals. **Science** 3(11): 506-508.

Feltus, F.A., G.E. Hart, K.F. Schertz, A.M. Casa, S. Kresovich, S. Abraham, P.E. Klein, P.J. Brown and A.H. Paterson. 2006. Alignment of genetic maps and QTLs between inter-andintra-specific sorghum populations. **Theor. Appl. Genet**. 112: 1295–1305.

Gowda, P.S.B., R.A. Frederiksen, C.W. Magill and G.W. Xu. 1995. DNA markers for downy mildew resistance genes in sorghum. **Genome** 38: 823-826.

Grivet, L. and P. Arruda. 2002. Sugarcane genomics: depicting the complex genome of an important tropical crop. **Curr. Opin. Plant Biol**. 5: 122-127.

Grivet, L., A. D'Hont, P. Dufour, P. Hamon, D. Roques and J.C. Glaszmann. 1994. Comparative genome mapping of sugarcane with other species within the Andropogoneae tribe. **Heredity** 73: 500-508.

Grodzicker, T., J. Williams, P. Sharp and J. Sambrook. 1975. Physical mapping of temperature sensitive *mutants of adenovirus*. **Cold Spring Harbor Symp. Quant. Biol.** 39: 439-446.

Guihua, Z., G. Zhai, Q. Feng, S. Yan, A. Wang, Q. Zhao, J. Shao, Z. Zhang, J. Zou, B. Han and Y. Tao. 2012. Identification of QTLs for eight agronomically important traits using an ultra-high-density map based on SNPs generated from high-throughput sequencing in sorghum under contrasting photoperiods. **J. Exp. Bot**. 63 (15): 5451-5462.

He, H., N. Yajing, C. Huawen, T. Xingjiao, X. Xinli, Y. Weilun and D. Silan. 2012. cDNA-AFLP analysis of salt-inducible genes expression in *Chrysanyhemum lavandulifolium* under salt treatment**. J. Plant Physiol**. 169: 410-420.

Heinz, D.J. and T.L. Tew. 1987. Hybridization procedures, pp 313–342. *In:* D.J. Heinz, ed. **Sugarcane Improvement through Breeding**. Elsevier, Amsterdam.

Heinz, S., A. Becker, K-U. Winter, C. Kircher and G. TheiBen. 2001. MADS-box gene are involved in floral development and evolution. **Acta Bioc. Polonica**. 48 (2): 351-358.

Hong, M. and de Pamphilis Claude. 2000. The ABCs of floral evolution. **Cell** 101: 5-8.

Hudson, R. 1995. Insects of pearl millet and their control**,** pp. 72–74. *In* I. D. Teare, ed. **Proc. of the First Nat'l Grain Pearl Millet Symp**., Tifton, GA, 17–18 January 1995. Spec. Pub. of USDA and Univ. of Georgia, Tifton, GA.

Hsu, T-W., W-C. Tsai. D-P. Wangc, S. Lin, Y-Y. Hsiao, W-H. Chen and H-H. Chen. 2008. Differential gene expression analysis by cDNA-AFLP between flower buds of *Phalaenopsis* Hsiang Fei cv. H.F. and its somaclonal variant. **Plant Sci.** 175: 415-422.

Ivandic, V., C.A. Hackett , E. Nevo, R. Keith, W.T.B. Thomasand and B.P Forster. 2002. Analysis of simple sequence repeats (SSRs) in wild barley from the fertile crescent: associations with ecology, geography and flowering time. **Plant Mol. Biol**. 48: 511-527.

Jaccoud, D., K. Peng, D. Feinstein and A. Kilian. 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. **Nucl. Acids Res**. 29: e25.

Joshua, L. and E. M. Meyerowitz**.** 1995**.** *UFO:* An Arabidopsis gene involved in both floral meristem and floral organ development. **Plant Cell** 7: 529-548.

Jordan, S.A. and P. Humphries. 1994. Single nucleotide polymorphism in exon 2 of the BCP gene on 7q31-q35**. Hum. Mol. Genet**. 3: 1915.

Julien, R. 1972. The photoperiodic control of flowering in *Saccharum*. *In* **Proceedings of the International Society of Sugarcane Technologists**. 14: 323-333.

Katsar, C., A.H. Paterson, G.L. Teetes and G.C. Peterson. 2002. Molecular analysis of sorghum resistance to the Greenbug (Homoptera: Aphididae). **J. Econ. Entomol.** 95(2): 448-457.

Keyan, Z.S., L. Haiwen, E.K. Patricia, L. G. Roberto and R.A. Salzman. 2003. Using high-throughput amplified fragment length polymorphism to distinguish sorghum greenbug (*Homoptera*: *Aphididae*) biotypes. **Agric. Forest Entomol**. 5: 311-315.

Kiesten, B., and J.F. Doebley. 2005. Molecular evolution of *FLORICAULA/LEAFY* orthologs in the Andropogoneae(Poaceae). **Mol. Biol. Evol**. 22(4): 1082-1094.

Kimberley, B.R., D.R. Jordan, S.C. Chapman, I.D. Godwin, E.S. Mace and C.L. McIntyre. 2008. Identification of QTL for sugar-related traits in a sweet x grain sorghum (*Sorghum bicolor* L. Moench) recombinant inbred population. **Mol. Breeding.** 22: 367-384.

Klein, R.R., J.E. Mullet, D.R. Jordan, F.R. Miller, W.L. Roone, M.A. Menz, C.D. Franks and P.E. Klein. 2008. The effect of tropical sorghum conversion and inbred development on genome diversity as revealed by high resolution genotyping. **Crop Sci.** 48: S12–S26.

Kong, L., J. Dong and G.E. Harrt. 2000. Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple sequence repeats (SSRs). **Theor. Appl. Genet**. 101: 438 – 448.

Kosambi, D.D. 1944. The estimation of map distances from recombination values. **Ann. Eugen.** 12: 172-175.

Lagercrantz, U.L.F., J. Putterill, G. Coupland and D. Lydiate. 1996. Comparative mapping in *Arabidopsis* and *Brassica*, fine scle genome colinearity and congruence of genes controlling flowering time. **Plant J**. 9 (1): 13-20.

Lekgari, A.L. 2010. **Genetic mapping of quantitative trait loci associated with bioenergy traits, and the assessment of genetic variability in sweet sorghum (*Sorghum bicolor* (L.). Moench)**. Ph.D. dissertation, Agronomy and Horticulture. University of Nebraska, Lincoln, NE.

Li, M., N. Yuyama, L. Luo, M. Hirata and H. Cai. 2009. *In silico* mapping of 1758 new SSR markers developed from public genomic sequences for sorghum. **Mol. Breeding.** 24: 41-47.

Mace, E.S. and D.R. Jordan. 2011. Integrating sorghum whole genome sequence information with a compendium of sorghum QTLstudies reveals uneven distribution of QTL and of gene-rich regions with significant implications for crop improvement. **Theor. Appl. Genet**. 123: 169–191.

Madhusudhana, R. and J.V. Patil. 2013. A major QTL for plant height is linked with bloom locus in Sorghum (*Sorghum bicolor* (L.) Moench)**. Euphytica** 191: 259-268.

Ming, R., S.-C. Liu, Y.-R. Lin, J. da Silva, W. Wilson, D. Braga, A. van Deynze, T.F. Wenslaff, K.K. Wu, P.H. Moore, W. Burnquist, M.E. Sorrells, J.E. Irvine and A.H. Paterson. 1998. Detailed alignment of *Saccharum* and *Sorghum* chromosome: comparative organization of closely related diploid and polyploidy genomes. **Genetics** 150: 1663-1682.

Ming, R., A. Terrye, D. Mante, E. Hernandez, P.H. Moore, E. James, E. Irvine and A.H. Perterson. 2002. Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploid genome. **Genome** 45: 794-803.

Mohan, S.M., R. Madhusudhana, K. Mathur, D.V.N. Chakravarthi, S. Rathore, R.N. Reddy, K. Satish, G. Srinivas, N.S. Mani and N. Seetharama. 2010. Identification of quantitative trait loci associated with resistance to foliar diseases in sorghum [*Sorghum bicolor* (L.) Moench]. **Euphytica** 176: 199-211.

Murai, K., M. Miyamae, H. Kato, S. Takumi and Y. Ogihara. 2003. *WAP1*, a wheat *APETALA1* homolog, plays a central role in the phase transition from vegetative to reproductive growth. **Plant Cell Physiol**. 44 (12): 1255-1265.

Murphy, R.L., R.R. Klein, D.T. Morishige, J.A. Brady, W.L. Rooney, F.R. Miller, D.V. Dugas, P.E. Klein and J.E. Mullet. 2011. Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. **Proc. Natl. Acad. Sci**. **USA.** 108: 16469–16474.

Parth, D.K., D.R. Jordan, E.A.B. Aitken, E.S. Mace, P. Jun-ai, C.L. McIntyre and I.D. Godwin. 2008. QTL analysis of ergot resistance in sorghum. **Theor. Appl. Genet**. 117: 369-382.

Pastina, M.M., R. Gazaffi, M. Malosetti, G.R.A. Margarido, K.M. Oliveira, L.R. Pinto, A.P. Souza, F.A. van Eeuwijk and A.A.F. Garcia. 2012. A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. **Theor. Appl. Genet**. 124: 835–849

Paterson, H.A. 2008. Genomics of sorghum, 6 pages. *In* Hindawi Publishing Corporation. **Int. J. Plant Genomics**. V 2008, article ID 362451 doi:10.1155/2008/362451.

Paterson, A.H., J.E. Bowers, R.Bruggmann, I. Dubchak, J. Grimwood, H. Gundlach, G. Haberer, U. Hellsten, T. Mitros, A. Poliakov, J. Schmutz, M. Spannagl, H. Tang, X. Wang, T. Wicker, A.K. Bharti, J. Chapman, F.A. Feltus, U. Gowik, I.V. Grigoriev, E. Lyons, C.A. Maher, M. Martis, A. Narechania, R.P. Otillar, B.W. Penning, A.A. Salamov, Y. Wang, L. Zhang, N.C. Carpita, M. Freeling, A.R. Gingle, C.T. Hash, B. Keller, P. Klein, S. Kresovich, M.C. McCann, R. Ming, D.G. Peterson, M-U. Rahman, D. Ware, P. Westhoff, K.F.X. Mayer, J. Messing and D.S. Rokhsar. 2009. The *Sorghum bicolor* genome and the diversification of grasses. **Nature** 457: 551-556.

Patrick, J.B., L.R. William, F. Cleve and K. Stephen. 2008. Efficient mapping of plant height quantitative trait loci in a sorghum: Association population with introgressed dwarfing genes. **Genetics** 180: 629–637.

Peng, L., G. Ji, Y. Han, S. Hou, S. Li, X. Ma, R. Du and G. Liu. 2013. Association analysis of sugar yield-related traits in sorghum [*Sorghum bicolor* (L.)]. **Euphytica** 193: 419-431.

Poehlman, J. M. 1987. Breeding sorghum and millet. **Breeding Field Crops**. 3rd. AVI Publishing Co. Ltd. Westport, Connecticut, U.S.A.

Pollock, J. D. 2002. Gene expression profiling: methodology challenges, results, and prospects for addiction research. **Chem. Physics Lipids.** 121: 241- 256.

Punnuri, S, Y. Huang, J. Street and Y. Wu. 2013. Developing new markers and QTL mapping for greenbug resistance in sorghum [*Sorghum bicolor*(L.) Moench]. **Euphytica** 191: 191-203.

Qin, L., H. Overmars, J. Helder, H. Popeijus, J. Rouppe van der Voort, W. Groenink, P. van Koert, A. Schots, J. Bakker and G. Smant. 2000. An efficient cDNA AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. **Mol. Plant Microbe Interact**. 13: 830–836.

Que, Y-Xi., J-W. Lin, X-X. Song, L-P. Xu and R-K. Chen. 2011. Differential gene expression in sugarcane in response to challenge by fungal pathogen *Ustilago scitaminea* revealed by cDNA-AFLP. 10 pages *In* Hindawi Publishing Corporation. **J. Biomedicine Biotechnol.** Vol. 2011, Article ID 160934, doi:10.1155/2011/160934

Quinby, J.R. 1974. **Sorghum Improvement and the Genetics of Growth**. College Station, TX: Texas A&M University Press.

Quinby, J.R. and R.E. Karper. 1954. Inheritance of height in sorghum. **Agron. J**. 46: 211- 216.

Rajewski, J.A., X. Ni, J.P. Wilson, I. Dweika and G.D. Buntin. 2009. Evaluation of resistance to chinch bug in pearl millet in temperate and subtropical environments. Online. **Plant Health Progress** doi:10.1094/PHP-2009-0112-01-RS.

Rami, J.-F., P. Dufour, G. Trouche, G. Fliedel, C. Mestres, F. Davrieux, P. Blanchard and P. Hamon. 1998. Quantitative trait loci for grain quality, productivity, morphological and agronomical traits in sorghum (*Sorghum bicolor* L. Moench). **Theor. Appl. Genet**. 97: 605-616.

Ray, M., T.A.D. Monte, E. Hernandez, P.H. Moore, J.E. Irvine and A.H. Paterson. 2002. Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploidy genomes. **Genome** 45: 794-803.

Reddy, R.N., R. Madhusudhana, S.M. Mohan, D.V.N. Chakravarthi, S.P. Mehtre, N. Seetharama and J.V Patil. 2013. Mapping QTL for grain yield and other agronomic traits in post-rainy sorghum [*Sorghum bicolor* (L.) Moench]. **Theor. Appl. Genet**. 126: 1921-1939.

Rooney, W.R. 2000. **"Genetic and Cytogenetics."Sorghum: Origin, History, Technology, and Production.** C. W. Smith and R. A. Frederiksen eds. New York: Wiley. 261-307.

Salvi, S., S. Corneti, M. Bellotti, N. Carraro, M.C. Sanguineti, S. Castelletti and R. Tuberosa. 2011. Genetic dissection of maize phenology using an intraspecific introgression library. **BMC Plant Biol.** 11: 4.

Sanguinetti, C.J., N.E. Dias and A.J. Simpson. 1994. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels**. Biotechniques** 17 (5): 914-921.

Schloss, S.J., S.E. Mitchell, G.M. White, R. Kukatla, J.E. Bowers, A.H. Paterson, and S. Kresovich. 2002. Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. **Theor. Appl. Genet**. 105: 912-920.

Sharma, H.C. 1993. Host-plant resistance to insects in sorghum and its role in integrated pest management**. Crop Prot**. 12: 11-34.

Smith, G.A. and D.R. Buxton. 1993. Temperate zone sweet sorghum ethanol production potential. **Bioresource Tech**. 43(1): 71-75.

Somashekhar, P., Y. Huang, J. Steels and Y. Wu. 2013. Developing new markers and QTL mapping for greenbug resistance in sorghum [*Sorghum bicolor* (L.) Moench]. **Euphyica** 191: 191-203.

Srinivas, G., K. Satish, S.M. Mohan, R.N. Reddy, R. Madhusudhana, D. Balakrishna, B.V. Bhat, C.J. Howarth and N. Seetharama. 2008. Development of genic-microsatellite markers for sorghum staygreen QTL using comparative genomic approach with rice. **Theor. Appl. Genet**. 117: 283-296.

Srinivas, G., K. Satish, R. Madhusudhana and N. Seetharama. 2009. Exploration and mapping of microsatellite markers from subtracted drought stress ESTs in *Sorghum bicolor* (L.) Moench. **Theor. Appl. Genet.** 118: 703-717.

Srinivas, G., K. Satish, R. Madhusudhana., R.N. Reddy, S.M. Mohan and N. Seetharama. 2009. Identification of quantitative trait loci for agronomically important traits and their association with genic-microsatellite markers in sorghum. **Theor. Appl. Genet**.118:1439–1454.

Stanke, M., M. Diekhans, B. Robert and D. Haussler. 2008. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. **Bioinformatics** 24: 637-644.

Subramanian, R. 1995. **Studies on plant resistance to sorghum to chinch bug, Blissus *leucopterus leucopterus* (Say) (Hemiptera: Lygaeidae).** Ph.D. dissertation, University of Nebraska, Lincoln, NE.

Swapna, M. and P.K. Singh. 2008. Shoot apex development at various stages of flowering in sugarcane (*Saccharum* spp. hybrid). **Cytologia** 73(2): 173-177.

Tao, Y.Z., D.R. Jordan, R.G. Henzell and C.L. Mclntyre. 1998. Identification of genomic regions for rust resistance in sorghum. **Euphytica** 103: 287-292.

Taramino, G., R. Tarchini, S. Ferrario, M. Lee, and M.E. Pe′. 1997. Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. **Theor. Appl. Genet.** 95: 66 -72.

Tautz, D. and M. Renz. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. **Nucle. Acids Res**. 12 (10): 4127-4138.

Teshome, A., D. Patterson, Z. Asfew, J.K. Torrance and J.T. Arnason. 2007. Changes of *Sorghum bicolor* landrace diversity and farmers selection criteria over space and time. Ethiopia. **Genet. Res. Crop. Evol**. 54: 1219-1233.

Theissen, G., A. Becker., A. Di Rosa., A. Kanno., J.T. Munster., K.U. Winter and H. Saedler. 2000. A short history of MADS-box genes in plants. **Plant Mol. Biol.** 42 (1): 115-49.

Ungerer, M.C., S.S. Halldorsdottir, J.L. Modliszekwski, T.F.C. Mackay, and M.D. Purugganan. 2002. Quantitative trait loci for inflorescence developmemt in *Arabidopsis thaliana*. **Genetics** 160: 1133-1152

Van, Ooijen J.W. and R.E. Voorrips. 2001. Joinmap 3.0 software for the calculation of genetic linkage maps. **Plant Research International**, Wageningen, The Netherlands.

Voorrips, R.E., 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. **J. Hered**. 93 (1): 77-78.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van De Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman and M. Kuiper. 1995. AFLP: a new technique for DNA fingerprinting. **Nucl. Acids Res**. 23 (21): 4407–4414.

Vuylsteke, M., J.D. Peleman and J.T. van E. Michiel. 2007. AFLP-based transcript profiling (cDNA-AFLP) for genome-wide expression analysis. **Nature Protocols** 2: 1399-1413.

Wang, S., C.J. Basten, and Z.B. Zeng. 2007. **Windows QTL Cartographer 2.5**. Department of Statistics, North Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm).

Wang, J., B. Roe, S. Macmil, Q. Yu , J.E. Murray, H. Tang, C. Chen, F. Najar, G. Wiley, J. Bowers, M-A.V. Sluys, D.S. Rokhsar, M.E. Hudson, S.P. Moose, A.H. Paterson and R. Ming. 2010. Microcollinearity between autopolyploid sugarcane and diploid sorghum genomes. **Genomic**. 11: 261.

Weber, J.L. and P.E. May. 1989. Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. **Am. J. Hum. Genet**. 44: 388-396.

Weigel, D., J. Alvarez, D.R. Smyth and M.F. Yanofsky and E.M. Meyerowitz. 1992. *LEAFY* controls floral meristem identity in Arabidopsis. **Cell** 69: 843–859.

Williams, J.G.K, A.R. Kublelik, K.J. Livak, J.A Rafalski and S.V. Tingey. 1990. DNA polymorphism's amplified by arbitrary primers are useful as genetic markers. **Nucl. Acids Res**. 18: 6531-6535.

Wu, K., W. Burnquist, M. Sorrels, T. Tew, P. Moore and S. Tanksley. 1992. The detection and estimation of linkage in polyploids using single-dose restriction fragments. **Theor. Appl. Genet**. 83: 294–300.

Xie, H., D. Ding, Z. Cui, X. Wu, Y. Hu, Z. Liu, Y. Li and J. Tang. 2010. Genetic analysis of the related traits of flowering and silk for hybrid seed production in maize. **Gene & Genome** 32: 55-61.

Yamaguchi, N., M.F. Wu, C.M. Winter, M.C. Berns, S.Nole-Wilson, A. Yamaguchi, G.Coupland, B.A. Krizek and D. Wagner. 2013. A molecular framework for auxin-mediated initiation of flower primordial. **Dev. Cell. 24 (3): 271-82.**

Yang, Z., S. Gu, X. Wang, Z. Tang and C. Xu. 2008. Molecular evolution of the CPP-like gene family in plants: insights from comparative genomics of Arabidopsis and rice. **J. Mol. Evol. 67 (3): 266-77.**

Yanofsky, M.F. and F. Martin. 1995. Floral meristems to floral organs: genes controlling early events in *Arabidopsis* flower development. **Ann. Rev. Plant Physiol. Plant Mol. Biol.** 46: 167-88.

Yousra, E.M., T. Shehzard and K. Okuno. 2012. Mapping of QTLs underlying time in sorghum [*Sorghum bicolor* (L.) Moench]. **Breed. Sci.** 62: 151-159.

Yuan, J.S., K.H. Tiller, H. Al-Ahmad, N.R. Stewart and C.N. Jr. Stewart. 2008. Plants to power: bioenergy to fuel the future. **Trends Plant Sci**. 13: 421–429.

Zeller, F.J. 2000. Sorghum (*Sorghum bicolor L*. Moench): utilization, genetics, breeding. **Bodenkultur**. 51 (1): 71-85.

APPENDIX

Appendix A.1 Construction of AFLP adapter

Strategy	Solution
1.This protocol is designed to use *Eco*RI/*Mse*I combination: Dilute the oligonucleotides to 100 µM in ddH_2O	-
2. Recipe for making the 200 µL adapter	**5 µM *Eco*RI adapter** 10 µL 100 µM *Eco* RI.I 10 µL 100 µM *Eco* RI.II Adjust volume with 180 µL ddH_2O **50 µM *Mse*I adapter** 100 µL 100 µM *Mse*I.I 100 µL 100 µM *Mse*I.II
3. Boil the mixture at 95 °C for 10 min, and then slowly cool down to the room temperature. So that, the two oligonucleotides in the mixture will bind together with hydrogen bound and form adapter.	-
4. Store at -20 °C	

Appendix A.2 Buffers and Solutions

Name of buffers and solutions	Components and concentration
1. DNA extraction buffer	50 mM Tris-HCl 25 mM EDTA 1.4 M NaCl % CTAB 1 mM 1,10-phenathroline 1% β-mercaptoethanol (added just before use)
2. TE buffer	10 mM Tris-HCl, pH 8.0 1 mM EDTA, pH 8.0
3. DGGE loading buffer	50 mL glycerol 2.5 mL 20X DGGE buffer 0.05 mg of bromophenol blue
4. 12% non-denatured polyacrylamide	116.8 g acrylamide 3.2 g bis-acrylamide 50 mL 20X DGGE buffer Bring to 1 L with dH_2O
5. 20X DGGE buffer	40 mM Tris-HCl, 1 mM EDTA 20 mM sodium acetate adjust pH 7.7 with glacial acetic acid
6. RNA extraction buffer	2% CTAB, , 2% polyvinylpyrrolidinone K30 (PVP) 100 mM Tris-HCl pH 8.0, 2 M NaCl 2% β-mercaptoethanol (added just before use)
7. SSTE Buffer	1.0 M NaCl 0.5% SDS 10 mM Tris-HCl pH 8.0 1 mM EDTA pH 8.0 Note: the buffer have to be warmed at 65 °C before use

Appendix A.2 (Continued)

Name of buffers and solutions	Components and concentration
8. 20X NBC Buffer	1 M Boric acid, , 20 mM Sodium citrate 100 mM NaOH pH 7.5
9. Glass bond solution	500 µL glacial acetic acid 99.5 mL 95%EtOH 300 µL glass bond Bring to 100 mL with dH_2O
10. 5% denatured polyacylamind gel	125 mL 37% acrylamind solution 450 g urea 200 mL 10X TBE buffer Bring to 1 L with dH_2O
11. 10% APS	0.5 g Ammonium persulfate Bring to 5 mL with dH_2O
12. Sequencing loading buffer	3 g xylene cyanol 3 g bromophenol blue 0.2 mL 0.5 M EDTA 10 mL formamide, Bring to 100 mL with dH_2O
13. 10% acetic acid	250 mL glacial acetic acid Bring to 2.5 L with dH_2O
14. silver solution	2.5 g A_gNO_3 3.75 mL 37% formaldehyde Bring to 2.5 L with dH_2O
15. Developer solution	30 g Na_2CO_3 1.5 mL 37% formaldehyde 200 µL 10 mg/mL $Na_2S_2O_3$.5H2O Bring to 1 L with dH_2O

Appendix A.2 (Continued)

Name of buffers and solutions	Components and concentration
16. 5X TBE buffer	54 g Tris base 27.5 g Boric acid 0.5 M EDTA pH 8.0 Bring to 1 L with dH_2O
17. 10 mg/mL $Na_2S_2O_3.5H2O$	0.1 g Sodium thiosulfate Bring to 10 mL with dH_2O

Appendix Table 1 Sequences of adaptors and primers used in cDNA-AFLP

Adaptors and primers	Sequence **(5'-3')**
Adaptors	
*Eco*RI adapter 1	CTCGTAGACTGCGTACC
*Eco*RI adapter 2	AATTGGTACGCAGTCTAC
*Mse*I adapter 1	GACGATGAGTCCTGAG
*Mse*I adapter 2	TACTCAGGACTCAT
*Eco*RI primers **(5'-3')**	*Mse*RI primers **(5'-3')**
GACTGCGTACCAATTC**ACA**	GATGAGTCCTGAGTAA**AAA**
GACTGCGTACCAATTC**ACG**	GATGAGTCCTGAGTAA**AAC**
GACTGCGTACCAATTC**ACT**	GATGAGTCCTGAGTAA**AAG**
GACTGCGTACCAATTC**AGG**	GATGAGTCCTGAGTAA**AAT**
GACTGCGTACCAATTC**GCA**	GATGAGTCCTGAGTAA**ACA**
GACTGCGTACCAATTC**GCC**	GATGAGTCCTGAGTAA**ACC**
GACTGCGTACCAATTC**GTC**	GATGAGTCCTGAGTAA**ACG**
GACTGCGTACCAATTC**TAC**	GATGAGTCCTGAGTAA**ACT**
GACTGCGTACCAATTC**TAG**	GATGAGTCCTGAGTAA**AGC**
GACTGCGTACCAATTC**TCG**	GATGAGTCCTGAGTAA**AGG**
	GATGAGTCCTGAGTAA**ATA**
	GATGAGTCCTGAGTAA**ATC**
	GATGAGTCCTGAGTAA**ATG**
	GATGAGTCCTGAGTAA**ATT**
	GATGAGTCCTGAGTAA**CAA**
	GATGAGTCCTGAGTAA**CAC**
	GATGAGTCCTGAGTAA**CAG**
	GATGAGTCCTGAGTAA**CAT**
	GATGAGTCCTGAGTAA**CTA**
	GATGAGTCCTGAGTAA**CTC**
	GATGAGTCCTGAGTAA**CTG**
	GATGAGTCCTGAGTAA**GCC**
	GATGAGTCCTGAGTAA**GTA**
	GATGAGTCCTGAGTAA**GTC**

Appendix Table 2. Sequences of flanking QTL primers that related flowering trait and flower-specific sugarcane TDFs primers

Name	Sequences of primer (5'-3')		Reference
	Forward	Reverse	
Bmc0067	AACGTACGAGCTCTTTTTCTA	ATGCCAACTGCTTGTTTAG	Ivandic *et al*., 2002
Bmag0211	ATTCATCGATCTTGTATTAGTCC	ACATCATGTCGATCAAAGC	Ivandic *et al*., 2002
bnlg1329	ATAGAATGGGATGTGGGCAA	TCCGATCATATCGGGAGATC	Xie *et al*., 2010
bnlg1185	CGGTCCAGGCAGGTTAATTA	GACTCGAGGACACCGATTTC	Xie *et al*., 2010
bnlg1784	GCAACGATCTGTCAGACGAA	TTGGCATTGGTAATGGGTCT	Xie *et al*., 2010
bnlg1808	CTTTTCTCTTCTAGTAATGAACAGTCA	GCATGATCGAACGAAGGC	Xie *et al*., 2010
bnlg2046	TTGGTGAAACGGTGAAATGA	CTGGTGAGCTTCACCCTCTC	Xie *et al*., 2010
bnlg2144	TCTGGGTGTGCTTGCTCTC	TGTTCTCAGCATTCCCAACA	Xie *et al*., 2010
dupssr11	AGGCAAGGCTTTCTTCATAC	AGGCAAGGCTTTCTTCATAC	Xie *et al*., 2010
phi339017	ACTGCTGTTGGGGTAGGG	GCAGCTTGAGCAGGAAGC	Xie *et al*., 2010
umc1025	GCTCCACTTCCACCCTGATATG	GCTAATGTCCCCATTGATGAT	Xie *et al*., 2010
umc1044	CACCAACGCCAATTAGCATCC	GTGGGCGTGTTCTCCTACTACTCA	Xie *et al*., 2010
umc1174	GCTAGTAGCTCTAGTTGTCCGCGA	GCTCAAGGTTGTTTTCTGCCAGT	Xie *et al*., 2010
umc1396	TTCGATTATTCCATTGAGCCTCTG	CTCCTAACGCAGGAGACAAGAGAG	Xie *et al*., 2010
umc1640	ACTACACGGTGTGAGATGTGATCG	GTCGTCGCAAGAACAACAAGG	Xie *et al*., 2010
umc1974	ACAAGGAGACCCTCCTCAGCTAGT	GTAAGCTGTGGCCATACTACCACC	Xie *et al*., 2010
umc1987	ACCCTCCGAAAAGCAAGCTC	CGTGGGCTCCTCCTTCTTGT	Xie *et al*., 2010

Appendix Table 2. (Continued)

Name	Sequences of primer (5'-3')		Reference
	Forward	Reverse	
umc2052	GTACCCAACAAGCCCTACACCTCT	CTTCCTCACGCCCCTGTAGTG	Xie *et al*., 2010
bnlg1007	GATGCAATAAAGGTTGCCGT	ATGTGCTGTGCCTGCCTC	Salvi *et al*., 2011
umc1167	CCTGCATGCATTAGGTATACGAAG	GTTTCTTCCAAGTTTTTGGCTTGA	Salvi *et al*., 2011
umc1271	CTCTCCTCGTCCGGTAATTAAGC	GCTTCTTCTTCTTGCGCTTCTCT	Salvi *et al*., 2011
umc1395	TGAATGAGTGGCATTCAAAATCTG	CAGATTGCATGTGTGAGTGTGTGT	Salvi *et al*., 2011
umc1528	AGTTCAACTGCTTAAGATCCGGTG	GTCTGTCGTTGTGTGCCAGTG	Salvi *et al*., 2011
umc1771	CATCAGGAAGGAAGACGACTAGGA	GTGAAATGTTGTTTCCAATGCAAG	Salvi *et al*., 2011
umc1846	ATTATTGGTCACAGGCCCTACCTT	TTAGGCCCTCGTCTTGTAGACTTG	Salvi *et al*., 2011
FLO/LFY	CCAACGACGCCTTCTCGG	GGCACTGCTCGTACAGATGG	Kiesten and Doebley, 2005
WAP1	ATCAGACTCAGCCTCAAACA	TAGAGACGGGTATCATGGAA	Murai *et al*., 2003
26 flower-specific sugarcane TDFs			
sfw1S_1C	GCACTCAACCACTTGGGCTA	TGTATCAATTAGATCCAGCA	Present study
sfw6S_1E	TCATAAAGAGATGGTTGTCATGG	CGGTTGTCAAAAGTTATTTCTGC	Present study
sfw3F_1M	TTCAGCTGAGGAGGGCTTAC	AGGGAACATCCCAACACAAG	Present study
sfw4S_1S	GTGCTTCAGTGAGCTCTCCA	CACGAGATCTCCGTCCTCTG	Present study

Appendix Table 2. (Continued)

Name	Sequences of primer (5'-3')		Reference
	Forward	Reverse	
sfw5A_1S	AAATGGGGGTCCTACTGTCC	CGAATAGCCGTGACACTCAG	Present study
sfw4DS_1X	TTGAGCCTTACATGCACCAG	CAGAAAAGCAATGCAGCAGA	Present study
sfw6F_1Y	CATTCCCGGATCAACAATTC	AGGGTTGGTCCCTGCTTAAT	Present study
sfw1DS_2D	CGGCATAAGGATCGGAGTAA	CTGCGAATTGGTAGCAGTCA	Present study
sfw4DS_2D	GGAAGCGAGAAAAATCAACG	AACGTTTTTCCAGGATGTGG	Present study
sfw5DS_2D	CCCCAACAATGAATTTCGAC	AATTCCGCAGTAACAACCTG	Present study
sfw3E_2D	AAAGTGGGCCAGAAGAGACA	CACCAATGGGTATGGTGACA	Present study
sfw3E_2R	GCCCTCGTCTGAAAGACTTG	GCTTCTTCTGCCCAAGTGAC	Present study
sfw1A_3C	AGATCCATCACACCTGTCAC	GTGTTCAAAGTTTGGTAGCC	Present study
sfw1.1A_3C	ACCTGGAAAGCTCAAAGCAA	GATGCTGCGTCATCTGCTAA	Present study
sfw2A_3C	GGATCAAGTGGTGTGGCTCT	GCTCGTCGAGAACAGTTTGA	Present study
sfw2E-2_3D	GCTTCCTTCACGAACTCGAT	CGCCGGTAAGCCATCTAAT	Present study
sfw3A_3E	GGAACGATGAACGCACTAGG	TTGTCCCATTTGTTCTTGAGC	Present study
sfw1E-1_3E	AGAACAAGTTCGGGCTGAGA	GGCGGACCAGTCACATAGAT	Present study
sfw2DS_3E	CCAAGTTGCGGCTAAGTTTC	AACACGGGTGCCTCAAAG	Present study
sfw1.1E_3G	GATTTGGCTCTGGCTACTGC	TCCGACCAAATGCTAAAAGG	Present study
sfw1.2E_3G	ACGGTGTGTTGCTTCACTCA	TATTCCTCTCGCATGGCTGC	Present study

Appendix Table 2. (Continued)

Name	Sequences of primer (5'-3')		Reference
	Forward	Reverse	
sfw5E_3G	CCCTGCTCAGATTAGCTTTA	TATAGTGAAAGCAGGACGTG	Present study
sfw2E_3H	TGCTCGTCAAATAAAGATGCAC	CTGGATGGGCACAAGGTAAT	Present study
sfw3DS_3H	CATGCAGATGCGGCTAGG	CCTCTGGAGCCTGAAGGAA	Present study
sfw1A-1_3M	GGCAAGAACACAAAGCCATT	CAGGATTCCCTTCTCCTTCC	Present study
sfw3E_3F	GAGGGAGATCATACGGGTCG	GTTAACGCCAAAGGGCAAGG	Present study

Appendix Table 3. Putative conserved domains analysis on of differentially expressed of flower-specific TDFs sequences

TDFs Name	**Length (bp)**	Putative conserved domain **Similarity**	Domain hit	Description of domain	***E*-value**
5C_2C	137	PKc_like superfamily	cl09925	Protein Kinases, catalytic domain. The protein kinase superfamily is mainly composed of the catalytic domains of serine/threonine-specific and tyrosine-specific protein kinases. It also includes RIO kinases, which are atypical serine protein kinases, aminoglycoside phosphotransferases, and choline kinases. These proteins catalyze the transfer of the gamma-phosphoryl group from ATP to hydroxyl groups in specific substrates such as serine, threonine, or tyrosine residues of proteins.	3.56e-03
2E_3H	200	RdRP	pfam05183	RNA dependent RNA polymerase ;This family of proteins are eukaryotic RNA dependent RNA polymerases. These proteins are involved in post transcriptional gene silencing where they are thought to amplify dsRNA templates.	3.64e-03

Appendix Table 4 The sequences of sugarcane flowering TDFs

TDF name	Sequence from 5'-3'
2D-1C	GGGGNANTCTCGCCCAAACCAGTATAAATCGTCTGTGTTCTTCTTGCACA CCATCCGGACTCCCGACGCGCAGACACTCATTACTCGTTGGTGTTAGGAC CGCCGTTTCTTACACCGACATATATATATACATTTTACTCAGGACTCATCA
6S-1E	CAAACTTCNATGTAATTCACAATCATAAAGAGATGGTTGTCATGGGAACC TTGTCCAATGTGTTTCAACAAGGNGATATTACGAGTCTAAAATTCTGAGG CAATTCTGAGGCTGTTTTACATCATGGTGATCTCATGTAGAGGCGCAAAG AGTTGCAGAAATAACTTTTGACAACCGATGGTTTACTCAGGACTCATCA
10A-1E	NGNGGNTCNNNCANNAGGGACGTGAGCTGGGTTTAGACCGTCGTGAGAC AGGTTAGTTTTACCCTACTGATGACCGTGCCGCGATAGTAATTCAACCTA GTACGAGAGGAACCGTTGATTCACACAATTGGTCATCGCGCTTGGTTTAC TCAGGACTCATCAG
1S-1I	GNAANAAAGANNTTGAGCGATTGCCTTCATTGTTTGTTTTGGCCCAACAT GCTTGGGACCATAGCATTTTGTGTTGTGGAGNNGTGGTGACTCACATCAG TTTGTTGCATTTAGACATTTAGTGATGCTACGAGCCATGCATTTTGTGTTA TATACTTGTGCGTAAGGAGTGAGCAAACATCTTGGAAGGTTTTTGCTACT CTTGACTTGTTTGTGATATATGAACTTGCCCATTTCTTTTCATTGGTATCTG CTATGGACATAATTGATGCATTTGATAATATAAACAATTCTTTTTAGCCA ATTACTTACTCAGGACTCATCA
2A-1I	CTGGTTGTTGANATTGGCATTGGCCAACAACGAATATTATTGTTTCCGTCC TTGCCTTCGGACTAGGATTGTTCCTTCTTTCCAAAACCTGCATCCTGTTTG CTTTATCTCGCTGCTTCTTCTGCTTATCCTTACTTTCGGTCAATAAGGCAT ACTTACTCAGGACTCATCA
2A-1K	CANNNAGNNNTCATTATAATCAAGCTGTACGCCTAATCCTACCATGATGA TCAAGACCTCAATGCCGTTAGCTTCAATGACTTACTCAGGACTCATCACA GGGGGNGGGGNNGNN
5S-1K	TGNANGNNNANNANTTATGATAAAGGCTGTTGCACAGGCTATCCCTACTT ATGCAATGGGGTGCTTTGACTTACTCAGGACTCATCA

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
3F-1M	TCGNNTTATTTCTGGGCATGCCCATCCTGGTAGTCCTTATGTATTTGGTCT TGCTGAAGTAGACCATGATGTAGAGTTCCCTGATCCGATGCCTGTTGTTG GGATTTCACATTCTGCAAAAGGTTATTGCTTGATATCTGTGCTAGAGACA ATGAAAACTTATTCAGCTGAGGAGGGCTTACTCAGGACTCATCAACTGAA TAAGTTTTCATTGTCTCTAGCACAGATATCAAGCAATAACCTTTTGCAGA ATGTGAAATCCCAACAACAGGCATCGGATCAGGGAACTCTACATCATGG TCTACTTCAGCAAGACCAAATACATAGGGACTACCGGGATGGGCTTGCCC AGAAATAAATCGACTTTGCG
6B-1M	AATANNACGGCAAAGCTCTTATAACCTCTAGGAATGTGATCTTGAGCTGA TTTTTGATGGCTTACTCAGGACTCATCAAAAATNGCTN
1S-1C	GGNGNGCTTNNCTGGGGACGACTCAGGCCCCCAGGGGTGCCTCGGGAGC ACTCAACCACTTGGGCTAGCGTCCGTTGGCATTTTGTAAAATAGGATAAA TAAGAAAAGTTATAGATTTTTTTTTTGCTGGATCTAATTGATACATGTGTG TGATCTATTGAATATGGTAAATATTTTACTCAGGACTCATCA
3D-1C	GGNNTTGNNGGNCGAACCAGTATAAATCGTCTGTGTTCTTCTTGCACACC ATCCGGACTCCCGACGCGCAGACACTCATTACTCGTTGGGTTAGGACCGC CGTTTCTTACACCGACATATATATATACATTTTACTCAGGACTCATCACAT NTCAATNTGTGGCGTGGCGTTGGGAGNTGTTTGCATGTTGGAGGAGATGT GCCA
7S-1E	TAGCAANTGGCGCGCTTTTCCCTCTTGTTGTTCTGGAAAGCAAAGGCAAT TGGCCCTCTCATCATGAGAGTCATGTACGAATTTTGCAGATTCTACTCTCA CAACAAATTGTGTACGAATTACTGTGTATATTAGAGCAAGGAACATCAAC GGTTACACCAACACCAAAAGAAAGGGACAGGTTTACTCAGGACTCATCA
12A-1E	TGGNNNATTTGCACCGTTCCAGCTCTGAGATCCATGAAGGATTCGGAGTT TGTAGGATAAAACATAGTGGTTTACTCAGGACTCATCA
5B-1M	TGAGAATCAAATTCCCGATAAGGCTTGANGGTATTGCATACTGTTCATGA CNGTNTTCAACAGCTATGAACCATAGCTCATCGCTGGGTACTCGAGACAN
3A-1R	ACTGANCTGGTGCACTTCATTATCTACAAGCTTATTAGCTCAGTGGGNAT TGGTATTGACAACAATATCCAGTTNNTTCCCGATCGTATTCCTCACCCTGC TTCTCTGACAAGGCATATGGAACACAACCAAATCCTCCACCTATATATTC CTGAACCTCTTGTTGCCAAAATCGATCGCTGCATCTGTTCCTTTACAGGAC

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
4S-1R	CGNTCANNTCGTTACGCCATTCGTGCAGGTCGGAACTTACCCGACAAGGA ATTTCGCTACCTTAGGACCGTTATAGTTACGGCCGCCGTTTACCGGGGCTT CAGTTCGAGGCTTCACCTTGCGGCTGACCTCTCCCTTTACTCAGGACTCAT CA
2D-1U	GCTTCTCTTTCACCGATGCCATGGAGATGTACACGGAAAACACATGGGTT TCCGTATTCTCACTCTCAGTAGTCTGTGCGTTCATTATTCTTTCCAGCTCA GGTGGACCTCTTCCAGGCACATGATTTCCAGAGTTGCCGAACTATATACT CAATTTTTTCGCACCAAAGGCATTCTACTATCGGAAGGCATATGACACAA CACTTATGAATGGGGGAAGAAGAGAACCCCACAAATGTAGATAGGTGCA AGGGAGATTAGCAGCTGCTTCCTCGCACCATTTCTANNGACTCATCAN
3D-1U	AGCANGNCCTAAGGANTATCCGCTTGATTCACTGACAAGATGGGAGGTTT TGGATGCAACTATATTTGCATTCTGGGCGAAGACATCAGTGGATGTTGAA CCAAAGAGAATTAGGCTGAAGTCAAGCAGTTATACTTCCAATACTATGCT TGACACTGTGACAGCAGCAACAGTGCATTTACTCAGGACT
4DS-1U	GCNGNCCTAAGGANTATCCGCTTGATTCACTGACAAGATGGGAGGTTTTG GATGCAACTATATTTGCATTCTGGGCGAAGACATCAGTGGATGTTGAACC AAAGAGAATTAGGCTGAAGTCAAGCAGTTATACTTCCAATACTATGCTTG ACACTGTGACAGCAGCAACAGTGCATTTACTCAGGACTCATCA
1DS-1X	CATTGAGTGCGTGGTGCCTTANAGCTNCCTCCCAGGGGGTTCCCGCTATC AAGGATCCCGCAAACCCCATTGGTGACGAGGAGAAGTACGACCGGTGGA ATGCAAAGGCCAAAAATGCACTCTACCGGGGCTTAGGCAAAGATATTTTC AATCGTGTGCGTAATGCAAAGAACGCTCATGATTTGTGGGAAAATCTTTG TGCTCTCCATGAGGGAACTAAGAGTGAGCGTGAGGAACGCTATCATATTG CTTTACTCAGGACTCATCAATGGGG
2DS-1X	TGTGGTNNTGCACCATGAGCCTGGCCAGGACGGGCTTATCACATTGGAAT GCTTCACTGTGCACTAAAGGATCATTCCTCTCCTCCTTCTCAAGGGTTGAT TTACTAAAAACGATGGTCCCAGAAACAACTTCATTATCCTTTTGACGACT TGACGCAACATCCTCAGCATCACGTAACTTAGGTGGCTCATATTTAGCTG GTGGAGCACTTATTGTGACATGAGTGCCAAATTTTTCAGGTAGCAACCCA ACAGGGTGCGCATTACTCAAGACTCATCA

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
3DS-1X	ATAGCCGGNAGANATCCCAATACTTCCCACGAAAGCCACCATCTTTGCCT CAGTAAACAAACGTAACCCGTTTTTACAGCCAAATATTGTCCATGGAGAA GGTGCAATAACATCAGATGCCATTGGATCGTTATGAGCAGATGAAAATA AAGTCCAATCTATTGAGTGCATGCGGCTATTGCGGCGGCTCAACCTGAAA GCCTGCCATCTTCTGTGTGAACAAGCAACAATGTGCTCATCTTTTCTAGG GGGCTTTACTCNNAACTCATCA
4DS-1X	GCACTTCTCCTACTACAGTACAGTATACAAAATGATGGTGGATTATGTTT AGACAGACAGTTGAGCCTTACATGCACCAGGTTCTGATGTATGAAGATCC AGTGGGACTAGAGGCAGCACGTAATACAGTTCCAATATGCGAACTTGAG GAGAAGGCGTTGGTTTCACTGGCCAAGGAAGGGAATTTCAATCCTTCCAA AGATGAAGAGAAGCATGCCTTTCTGCTGCATTGCTTTTCTGCTTTACTCCG
5DS-1X	GCNNNCCTGGNTCATTTATTATGCATCCTGTGTCATGGTGAGTGGACAAT GCAGGCCCAAGGCGNCATCANGGCTTGACAAATATAGGCCAAATTGAGT GGAACTGTGGCAGGTCGTCCTCTGCCATCAGCGACACCTTGTCTTACAAT TCACCACATCATTACTCAAGACTCATCAA
6DS-1X	ACATCTGTCAGAATCTGCCTTATGGATTGGCTGTCTTGTGCCGACCAAAC TNCGGTGGCATCNCTGGTGTTGGGGGATCTTTTTAAACTCGTCTCCATCAT GGCTGCTCGAAGAAAAAGCTCGACATAAAGGTGGCTTTTCCAGGACTCN
4S_1X	TGGGCTTCCGTTAGCAGCTTCCTGCTGTTGCTCGCTAGTGACAGAAATGG GGGTCCTACTGTCCCAGACCTGCTTTTGGACCTGCTGCTGTTTCGACCTCT TATCCTGCATCAACTTCTGGCGGCTCTCATAGAAGGCAAAGTCATCCAGG ATGGACGTCTTGCTGACATGATCCTTGAAGATCTTGAGCATCTGAAGGCC TTGTTCAAGGTGTATCTCCTGAGTGTCACGGCTATTCGTGACAGGCTTGTT CTCATTGTTCTCTATCCGCTTGAGCTA
6F-1Y	AGTCCCTGGGCACCGGATCCCAATCTTCCCTAGACCAAATTGCAAGTCCT TTCATGAGAATTCCCAATGACCGGGGAACAAATTCCAAGTGCACGGTTAA CAAATTCCCAAGCCTTTGATCCCAGATTCCCC
9A-1Y	CTACTTTGTGGAGTTTTGAGAACCAGGTCATTCGTCTCATCCATATGTTGA GATTTGATGAAAAGCTACTTTCTAGAGCACAGGATTCTGGAAAGGAAGT ATCACTTGGGAGTGATAATGAATATCATGAGGATTCAAGGTTTGAGACTG CTGAATTTACTCAGGACTCATCA

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
1S2-1Z	TCACCGTGTANNATCCGNACAGAACCAATCGGCGAAGGAGACGCCGCTG AAGATCTACCCGTAGTGCCCGGACCACAGGCCTTCCACCTTTCTCAAAGA TGACATGCTTTATCAGGGATTCTGCAAGACGTGTCTCCACCCTGCGGGTG ATGATGACTAGATCACAAGGGCGTGATCGTCTTTATGTACTTTCCTCATG CTTTTGCGGGATGTTGATCATCCAGACTGGCAGTATTTACTCAGGACTCA
2S-1Z	CTAAGCCACTGATTATTCATCTATGTGCACCTATGTGAAGCCGAGGCTCAAT CAACTAACTTTGAGCTGAGAACGATGTCCAATGCTTGGCTGTCAGTTGAAA AATATCCCAGACATGAAGTTATTTTGCTGGTAGCTTCAGGGTGTTTTCAGAT TGGTTTACAAACTTATTAGCACTGTCCAGCTGAAATTTCTTAGGTCATGCAA TGCCCATTTTGGTTGTTGGTTTATTTACTCAGGACTCATCA
3S-1Z	CTATNTCCGGCNAGGAATTATCAAGTCAAGTAACTCGGCTATGTTTTTGCAT TTGGCAAGCCAGCTAGAAGCATGCGACTCTGTATGCATGTTCGGTTCTGGGA GCATGGTCATTTGGTGGCAATGAATTTTGCCACACCTCATTCTGCTTTGCCA CCAGAATAATGTCACCAAACTGACAATATTTACTCAAGACTCAACA
4S-1Z	GGTGCACGCGNGTTCAACGTCAGTCTGATGCTAGTGTTCGCCGACTCTCCCA GAAGATCCATTCCTTGTTGTTTAGTTTCTAGACCTAAACTTTACATCTATCAC ATCGGATATTTACTCAGGACTCATN
1D-2C	TCTCACTGNCTGCGTACCAATTCCGGGTGACTGCGTACCACTTCACACGTAA GAAAGTGCCTCCCGACTCATCACTGACTGCGTACCAATTCACGGATAACTGT GTTCCACTCCACTCATAATGGGCTGCGTCCCACTTCACAGATGGCTGTGTAC CACTTCACACATAATTGGGTCCCACTCACCTCTTAGTGGGTTGCACCCCACT TCTCAGAGAGTTCCTCTCCTCTCAACACTTAGTGCGTTCCAACTCACAGATC ACAGAGTTCCTTTCCACTTGACAGGGAGTGCCTTTTCCCTCACCGTTTACTCT ACTTCTAAGCTAACAAGAAAGTGCTACTCACCCTGCTCTAGNTACGTGGNTC TTCAAATCGGAGAGNTTTTGGCGGACTCCGGGTGGAAGTGTCTACAACATG
3E-2D	AAAGTGGGCCAGAAGAGACATCAAATAAGTAGATCAAGAATATTCTACAG AAGATGCCCCNATGAAATGGGAACCGTATGGAAAGATCGTTTACTAGAAGC ACTTTGGGCTTATCAGACTGCCTACAAGACACCATTGGGTATGTCACCATAC CCATTGGTGTATGGAAAGACCTGTCATCTACCTGTGGAGCTAGAACACAAG GCTCATTGGGCTATCAAACGTTGGAACATGGATTTTACTCAAGACTCATCAA AAGCCTGATGAGTGAGCTAACTGATATTGTTT

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
10D-2D	TANAGCTGAATTCNGAGGTCCGGCAGCCGTGGCGAGCGATATTATANAATC CATGNNGTGAAGATGGCTACCGGGGGGCAATTGTTATCTTTGCGNCGGGGC TGCAACAACTTAGCCGGAAACATNANAGNAATGCCTGGGGAGCCTAATGAG TGACAAAACTCTTATTANTTGNGATCCGCTCACTGCCCGCTTTCTAGTTGTG ATACCTGGCTTGCCACCTGCCCTCCTGGATCTACAAAAGCTCTTGGAGAGGC GGTTTGATTCTTGGGCGCTCTTCCGCTTCCTCGCTTGCTGACTTTCTGCGCTC GGTCGTTCGGCTGCAGCGAGCGGTATCAGCCCACTCAAAGGCGGTACCACC
1A_3C	AGGAGTGTGANGGAGGAAGAACGCTTTGTCCAGCCTTCAGATCCATCACAC CTGTCACAGGCGGGAGCGCCAGCGCAGGCTGCTGCTTGTTTAGCTTCTCAAT GATGTAAGAGAAAGTCCCAAAGGTCAGGCAGCTCTGAATCAGAGTTGGTGG TGCACCTGGAAAGCTCAAAGCAAGGCCAGTACAACAGCCTGCGACTCCTGC ATTGATTCCATCATCTTTTCCACGAAGCTTTCTCAAAGAGCAGGCTACCAAA CTTTGAACACCAGCAAGAACAGCAAATATCTTAGCAGATGACGCAGCATCA GGAAATGATCCTTTACTC
2A_3C	CCCGGGCAAATTTGTAGGGTTGGATCAAGTGGTGTGGCTCTTTTGATATTAG TAACGTAAGTTTCGTGTCACTAAATGCGGTTGCGCTGCTCTGTATCGACAAT GCGTTGCGGTTCAGCATCGAGAAGGTTATGGAGAGGGAGAACCATCATATC CAAAGGGAAAGGGTTTATGAAGCAGAATGTAACCCTTTACTCAGGACTCAT CAAACTGTTCTCGACGAGCTAATCTTAGGACACCTACGTTATCCTTTACTCA GGACTCATCAAA
3S_3C	TAGTCNNGANTGAAGAAGAACGCTTGGGCCTCTCTTCATATTCATCACACCA AGTTACGTGTCAGNAATGCGGTTGCACTGCTCGGGACTCCCGATGTGTTTCC ATCATCTTTCCCGAAGGCTTGGAGAGGGAGAACCNNCCATCCAAAGGGAAA GGGTTTATGAAGCAAAATGTAACCCTTAACACGCGACTCATCAAAATGATCC TTTACTCAGGACTCATCA
4DS_3C	CTGNTTGCACCATATATAACATGCTTGGAAGGCTGTGAGTGTTGCTCTAGCT CCATCAAGCTCTTGTACTCAATGTAATTTCCACCACCGATCATGAAAACAAC AGCTTCTCTGAAGGGTCCTTTACTCAGGACTCATCAG
5E_3C	TAACCCTGCTGCGCAAAAATGGGATATGCGTTTGAAATAGATGTGCGAGTTA TTACGTATATCATGATCGATACAGAAATCGATCGAATCATCTGTTCCTTTACT CAGGACTCATCAA

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
6E_3C	ACTCAAGGACTATGTGTCTTGATCAAAGACATCTGAGCGCTACCATCTTCCA TAGTTATGACAACCTTACTTTCTTCTTTTGACACCTTTACTCAGGACTCATCA
7S_3C	TAAGATCCGAGAAACGAAAGGTTCTAAGTCGTACACAAATCTGCTAGATAT GTGTTGTGAAGATGTGTTCCAACAACCTTTACTCAGGACTCATCAGTTGTGT
3DS_3D	GGGAAAATGTAAGGCTTCGTTACTTACATACGGGAAACAAAGCAACTTTTCC CGTTCCTCCGCCATAACAATTTTTCCGTCCATTAGAACTGCCTTTTCTTGTAT TTCAGAACAGTAACCTAAAAAGCGTACTATATTTTGATGCCGAACCTTGATC ATGCTATCAACCTCACGAAGAAACATATTTTGATCAATTATATGGTTTTTACT CAGGACTCATCAA
4E_3D	GTTGGGACTACTGATGAATACATGCATAAGGTTGTTATTTCATCTCAATCTTC AGTTCTTACCAGACTCCAAGAAACACCATATGTTTTTACTCAGGACTCATCA
4A_3E	CCCCCAGGCCCCATCATTGTATAACTNGAATGGTTGGATGCTTGTATTCAAT GCACTATTGGTAGCTAATGCAGCTTATGCGCCTGGGGCTGGAGTTTACTCAG GACTCATCA
5E_3E	GTGGAACGGCGGGGGCACATACACCAAATACCATGNAATAACATTTTTCGA AGCACTACATTCCATTATTCTATAAACCTGCAGTTTACTCAGGACTCATCAA
4DS_3F	GGGGANTTTTATGGTTCGAAACATCTTATTTGGGTCGCAATTTTTACAAATA CCTCAGCTGTGAGCTCCCCAAGGTAGTATTATGTGACTATGTTTTCATGTGGT ATTCTTGTTAGAGGTCATTTACTCAGGACTCATCA
4D_3G	ANANTTTGCCGATGAGCATGAGGTGATACGGAGATCACCGGAGTCCCGCGG CAGCGAGCAGGAGAAACCACAAGGGCGGTCTGAAGTGCCAAGCCTCGGGA AGCAACCGAAGAACTAAGCGCGGTGGCCATGTGGGCTTTACTCAGGACTCA TCNCGACGAATTNNTACGTAGTACATTGGTNGTTTGATGGTTCTCTGGATCC
2E_3I	TCTACANGGGANATTTTCTATTCTTCACAAGAACTGACGGCTTGAGCGGAGG CCTGGAAGACCGCTCGGTGTGCCATTATACCAACCTCCTTGTGGTGTCGCTA ATGATCACACAGAACCAGTCCCCCTGGTGCTGCAGGCCGTCTTAGACCACCT TGTTCATTGGCGTTTACGCAGGACTCATCAGCCGATTGTTGACTAATTACGG GCACTTCTAAGACAAGACCTTGAATCAGCTTTTGAAGAGGAACTTGATAGCA TCTTTGATGTCACTCAGTTGCGCCAGTCACTTGGGCAGAAGAAGCGAGAGCT GGAGATTGAACTGAAGCGGATCAAGCGTTTACTCAGGACTCATCA

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
3B_3I	TTCTNGGTNANGGTGGTTNCGNGTAATGTAGTTNGCGGTGATAGGTCNNTCT TTTGTTTTTGTTCGAAGTGNNGANATTCTAAATGTTATGTCTGAAGTCTATAA AATGGNTTTTAANCNGGACTTGTCATTGTGNTGGTTTTTGATGNTGAATTTTG
4DS_3I	ACTCGGTATCCTAGTTTCAGTAGTTTCCGCAATCACATTATCACTTGCTGTAC TTGTGGAATCAAATAGATGTTAGAATCACAATTTGTTGGCGTTTACTCAGGA
6E_2M	CAAACCTCTTCCCTTGATTCAGTCCAAACATATCCATGGTAGATTGCCGTTTA GATTGTCTCTGGTTGAGCCATTTCCCATATTTACTCAGGACTCATCAAAGCG GCGGCTCGCGATAAGAAGGGAGAAAATTCTAATTTACTACACGAGAGATTT CCCAAACTAGGGACTCCTCCATGGCAGTTTTGGTGTGCTTTACCCATGGGGT AGGTACGTATATATAGGGAGAAAAACTCTCCACCTCCATGTCAACTGGCGAT
8E_2M	ACTTTGGGGAATATCCGACCTGCTAGCTACTTGCAGCATTTGTTCATATCTAC TCACTCACGCATCATCACNNNNNCCN
3E_2N	TTGCGAACCGGGGGGCTCTGGATGATACATGCATAAGGTTGTTATTTCATCT CAATCTTCAGTTCTTACCAGACTCCAAGAAACACCATATGATTTTACTCAGG ACTCATCAA
6E_2R	CTATCTGACTTCCCATAGAATTGAGAGCTGCATGTAGCCTTTTGTTTCGTTGC TGCATCTCTTTATTCAAGGCGTGCACTTGGTCACAAGCGGCGTCCATTTGGG CGTTTACTCAGGACTCATCAN
7D_2R	TACTCGGTATCCTAGTTTCAGTAGTTTCCGCAATCACATTATCACTTGCTGTA CTTGTGGAATCAAATAGATGTTAGAATCACAATTTGTTGGCGTTTACTCAGG ACTCATCA
5DS_2B	TTTNGGGGGANCACNTNCGNAAATGTACGGATCATCGTGAGAATGTGTCAA AATTACTATGCAGTAATCGTATTCAAGCCACGGACCATTTACTAGTCACCGT ACAGTCATTTGCTCGAGCAGTTGTCTTCACGCGATTAGTTTTACATGTGCTTG TAGGAAGATAAGCCTTCAACATATATTGGAAGAGTCAAAGTGCAAGAGTTC ATAATATTTGATGAACT
2C_2C	ATGNACTCTGTACGCACATTCACAAACACCGCTTGATTACTGCTGCTGGGCA CGGGCTGCAGCTGTGCCGACGCACTTCACCCACTCCAGACCTTGTTTTTCTTT CCACCATGGCATTATCTGGAAGATGAGTACAGGAAGCAGCAGTAAGAAGCT GATTAACTCCACAATGCATATGCAACCAGCAGTGGAGTTGTATGTACAAACC GGTTCAACCCAAAAACTTATGTTACTCAGGACTCATCAN

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
1D_3A	GCAAATTGGTTCGGATTGCGTTTGGCAGGAAAACCTGACTGTTTAGTTTTCC ATTAGCTTGTTCATTTGTTTCCCTGTGATTAGCGGAAGCAATATTGAAATATT GACAGTGGAAGAACTACACATTTTTTTATTCAACCTTTCTATCCTATAAATTA CATCTAGATATGCAATTCATATGGCAACAAATTCCATATGATATATGAATTA TATTCATGCTCTCATCATTTGTAACTGTATTTACTCAGGACTCATCAA
2D_3A	CCTTTTCCCTATGATTTCCGATTGATTCATTAGGCTCTGGTTTGGAGGTTGTG ACTATCATGGGTTGGAAGGAATTACTTGATTGGTGGTGTACAATTCTTGGAT AGGATAAAGGTTTCAGTTTTTGGTCTATTAGTTCAGTATTGTTTCATTTACAT GTCTATTCAGTACACTCGTATAATATATATGTATTTACTCACGACTCATCACA
5E_3G	ACTCCCGACGCGGAGGGAATTGACCTCATTCCCATCGACTACGGCTTTCGCC CTCGCCTTAGGGGCCGGCTAACCCTGCTCAGATTAGCTTTACTCAGGACTCA TCACGGGCGACAGAGAGGAGTCAACACCACCCGAAGCATCAGCCCGAGAAC AACAGGGCCCAGCAGGAGCGCAAGACCATCAGC
1A-1-3M	GCTCCTGAGACAATAAGGCAAGAACACAAAGCCATTCCCAGATAAAGAAAT GAGGCGAAAACAGTGCAAGTTCAAGTCTTAGGCATGGCCATTACCTGAGGT ATGTGATAATTCCTGTGCGCTACATGTTCCTTATTTGAATTGTCAAGTGAAG GATCAATATGTTTCCTCTTCCTTGATGTGGAAGGAGAAGGGAATCCTGTAGC TCCAATAGCACCATTTACAGCTGCATTTGTNGNCTGCTGCATGTTACTCAGG
1A.2-3M	TGATGATCNNAGCTTTCGACCAAATTTATGTCCTACTTAGTTCAATGCTTGA GCCATGACGAATCAGTCCGCATGCCTGATGGTGGTGGCTACAGGGTGTTCCT ATCCCACCGGATTGGGGAATACGTCAACATGGAGGATGAGGAGCAATTTGG AGAGTGGGAGTTTCATGATGCCCTGGAAATTTTAGAGAGTGACATTGCTGAT GTTGACCANNCTAATGTTACTCAGGACTCATCA
3A-3M	AGNAGTNNNANNTATGGCTAGAAACACATTTGAGCACTCCACACCTTTTAG CCCATCTTCCATGAGAACTCTAAATAGCTTGTCGACACCATCAAGCATGCTG CTAACTCGCACCTCCCAGGACTACATGCCTGCAGATGTTACTCAGGACTCAT CCANGACTGCTNACCANTTCNGGGACAGCCGGAGTCCANAGGGGGGTGGCA NTCGNANAGAATCCGAANTTGATGAAGGGN
4F-3M	TTGATANTNCNANACGGTCCTCATCTGTGCTTGGGTTGCTGTTATGCTAACTC GCAACTCTCTGTACCCCTCTCTTCAAAATCAGTTGCAAATGTGTNACTATGTT ACTCANNACACNNCAN

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
5F-3M	AGCTCATGATATACTCNGCGGATTCACCCNCCGAGATGTGTTTTCTTGGACN TCATGACCTGGGCTTTAATGTTACTCAGGACTCNANACANCTANACNATGCT CCATATCATGAATGCGTGGTTGTGCTCGCGTTTGCGTTTACATTTTCGTTTGT CTGCTGAACCTTGTAATTCCGTTACGTTCGCAGTGAATT
1DS-2D	AATTCTATATGAGGGANAAACCTGCTTATATTCAGCCCTTTGATATGGAAAC ACTGGGGCAAGCGTTCCAACTTCGAGAAACAGCTCCTGTGGATTTGCCCTCT ACTGAAAAGGGTATACCGACTATTTCGGGTAAACCAAAAAGTGAGTCCAAG GACAAAGAGAAGAAGCATAAAAGGCACAAAGACAAAGACAGGGACAAAG ACAAGGAACATAAGAAGCACAAACATCGGCATAAGGATCGGAGTAAGGAC
4DS_2D	CATTGAGTAGTCGCAAGATTTCCCAGAAATGGAAGCGAGAAAAATCAACGG CGGTAAACAGACGAAGCTGAAAACAACCCCCTTAGTGTTCTAATTCATTCCT TTTTAGTCGATTCTCAGTTGAGTTTAGGCGCCATGTGGTGCTTTGTATTAGTT TAGTTTGGCACTTGAATGGAATGTGAGCTGGTCTTTAGGTTATGCCACATCC TGGAAAAACGTTGCATCAGGAGCATCAGGCGTGTCCAATAAGATACTGAGG
5C-2C	GNGCNTGATCTGGTTTTACTTATATGCAGGCTGTTCCTCCAGTAGTCCACCGT GACATCAAATCTCCTAACATTTTGCTGGACCAGGCAATGCATGCTAGGGTAT GTTACTCAGGACTCANNNNNGACTGCNTCCCA
7A-2D	TGCNCTCANATGCTGCATCATAAGAGGAGTATATCTTCGATGACCAGGGTTG GTTCAGTTGTGTACTGTATCTCCTAATGCCCTAATGTTTACTAGGTTCATCTT GCTGAGAACCTGGTGAATAGTCATTTTTTAGTTACTCAGGACTCANCATTGA CTGCTACCAATTCGCAAAATAGGTGTTGTTACTCAGGACTCATCATTGACTG CCACCCATTCGCAAAATTGGTTTTGTTTCCCCGGANTTAATAATTGCTGCCTA
11A-2D	AGNGGTNNTTCATCTATTAGTACTGGCTAAGAGCAACGGTACGCACAGTAA AAGTAGTTACTCAGGACTCATCANTNNTCCGNNNNCGGGNGGCCNTGTAGT TNCTCAGGNNCANTCNNTGGCTGCCCNNGAGANTGTTNTTNANAGGCCCAA
1D_2Z	AATGAGATCNNNNNGAAACTCNNANNAGGAATAGATTATGTGCACCGCAAT TGACTCAGCTGCTGGCGACCTAGAGTTCCAATAAAATGTTGAGGAGTTTGGC TCAGATCGTTGAAGTATACCTCTNAGGNATACAACTTTCTCTTGTCACATGG ATGATAATTGCTAACACATCTTATCCATTCTATGAGAATAACAAACATATTG ATCTACGCCATAGATNGATGTATACTCAGGACTCATCATTGACTGCGTCCCA TTTCTGTTGTTTCTGGTACACTGGTCNCATCCTAGNCTGCTTTCNTNTGACGG

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
1E-1_3E	ACGNTGAACGCACTAGGCTTGTNTNANGAACTCTTTGAAGAACAAGTTC GGGCTGAGAACAGGCCAAACACGCACCTGAACAACATTGGTTACCGCC AAGTTGCGGCTAAGTTTCAGCAGAGGACACAGCTTCTCTATACAAAATT GCAGCTCAAGAACAAATGGGACAAGTTTACTCAGGACTCANNNNNTNC GGTGGATCTATGTGACTGGTCCGCCCATATNGGCGTCNGTAGCCTTATC AGNNTNNACTTGCTTNNNANNACCCGTGTTCATCATCTCATAGTTTACT
1E-2_3D	TCGAGATGTCTGCNCTTCCTTTGCCATTTCTTTTGCGGTATCTTGCAAAA ATTTTGCATACTCCATTCTTGCNTTCAATTTCCTTTTTACTCAGGACTCA ACAGGGACCATTAGACATACTTCGGGCNTACGATANGTTNANACGNCN NAATNNNTNANNAACATCNTCNCCANTNGGCNGNNTAATNAGNTTCGN
1E-2_3E	TATCGATTGTTCGANAACTCTTTGGAGAACTCTTGGGGATGNGAAAAGG ACAAACACGAACATGAACAAATTGGTTACCGCCAAGTTGCGGCTAAGT TTCAGCAGAGGACACAGCTTCTCTATACAAAATTGCAGCTCAAGAACAA ATGGGACAAGTTTACTCAGGACTCANNNNGGTACTGTGGATCTATGTNN ATGGTCTGCNCATGANGACGTNAGTAGNCNTATAAGTNACTACTTGCTT CCCGTGTTCATCATCTCATAGTTTACTCAGGACTCA
2DS-1_3E	TATNCAATAAACCTTTGGAGAACTCTTGGGGNTGNGAAANGCCAAACA CGCACATGAACAACATTGGTTACCGCCAAGTTGCGGCTAAGTTTCAGCA GAGGACACAGCTTCTCTATACAAAATTGCAGNNCAAGAACAAATGGGA CAAGTTTACTCAGGACTCANCANNTGCTCTACAAATTATGTCATATTAT AGTTGATTTTGTAAGTNNNCTCAGGACTCNNNNNNNCTTGCTTTGAGGC ACCCGTGTTCATCATCTCATAGTTTACTCAGGACTCA
2DS-2_3E	GAGTNNNANGNTGAACGCACTAGGCTTGANNTNANGAACTCTTTGAAG AACAAGTTCGGGCTGAGAACNGGCCAAACACGCACCATGAACAACATT GGTTACCGCCAAGTTGCGGCTAAGTTTCAGCAGAGGACACAGCTTCTCT ATACAAAATTGCAGCTCAAGAACAAATGGGACAAGTTTACTCAGGACT CANNNNNTGCTCTACGAATTATGTCATATTATAGTTGATTTTGTAAGTTT
2E-2_3D	GANTACNGTAGGTGCATCGTTTAGAGGAAGATTTCCCTTTCCACGGNNN NTGCTTCCTTCACGAACTCGATGATTTCGTCGCGCTGGAANGNACTACG GAGTGAAGCGTAGGGCCTTTCTTCACGTTGAGAGCTACCATGGCAGGAT AACCGTATCCACCCCCCTCCAAAAAGGTTCTCAAGATTAGATGGCTTAC

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
3DS_2Z	AACTATTANGGNGNTNNCCTGTTCATCTAACTCCGGNTTGACTACGTAA TCTATTTCATCGAGTATCTCCCGTATCTCTGCAACATGTTCCGGCAGCAT CATTGCAGTCCACTTTATATNACCTCTATAGCGTTTCATGTTTACTCAGG ACTCNNCCTTGACTGCGTACCAAANTCGGNGNGTCTGGTTTNNGGANTT ACTGGNGNCANATNANAGGNTGCNGNCAATNNNGGGGATTTCCGGGCN
5A-2_2Z	ATNCNGTGCCGCGCCTGAGAGATGGACCCGATGGACATCGTGGGCAAG TCCAAGGAGGACGTCTCCCTCCCCAAATCAACAATGTTTACTCAGGAAT CATCAGTGACTGCNACCAATTCTCGCNGGAANNNGTAATGTTTACTCAG GACTCATCAGTGACTGCTACCAATTCTCGCACGAACNCGNTAATGTTTA CTCAGGACTCATCAGTGACNGCAACCNATTCCCGCNAGAAANNGCTNN
5DS_2D	GTGCGANNANTGGCTCACTCCCTCCATANNNNCCACGNCCTCTAGATTT GGCACGCTTGCNTTGNAGCCACCCTTGTGAATGTAGTCTGGTACCCCCG GGGCGTTGTATGTATGTCTAGGTCCCCAACAATGAATTTCGACGGAACT TTTATCTGAGCTCCAGTCCATGGTGCAGAAAGCTCCCAACTCCTGTCTAT GCATCTGTAGTAGTTACTCAAGACTCATCNNTGACTGCTACCAATTTGC NAAATNCAGGTTGTTACTGCGGAATTAACAATGACCGNCNACCATTNCC
5S_2K	TCCTCTATAGTTCTAGTCAGCAATATTTTCTCCTCTATACTCAATTCTCCC GCACTCATCCATCCTCTATATTATCCTCTATCCACTCTATCCATTATGGG ACCCACTTGTTACTTTATCTAAATATATCCATGTGTGATCTAGTTTTGAA GGATTTATTGAGACGAATTTACTCAGGACTCANCATTGACTGCGAACCA ATCCAATGGTTGGGANCNATTGAANTGNCAANGAANTANTC
7A_2K	ATGTGNNCTNAGTTNGGAATCGTGGCTANNNTTCCGACGTCTCTGATTG GAGTACTAATGCATCCNGCCCGATAACAANTTNTNCNTCCCTGGAGNAN GTCGACTCAAGCCGTCAACNAGTGGGAGCAAGGCTTCTTTTACTCGGAG CTNGTCAATNTTTACTCAGGACTCATCAATGACTGCGTCCCATTTCTCCC TGCCCGCGACCCATTCTTGCCTGCCTCCNATTCCTCCGAGAGTGTCCCCC
1DS-2_3F	TGGTGATGANAAAGAGTGACTTCTAGCCAATTGGATGTCCGGACACTTG CTTCNAGGATAAANCNANTTNNGNTAATACACATTCATNACTACTATGA GGTGANTCTCAGTNTTTNGGTNTATTTTGTTGTNGCCAAGACTGTCATG GNCTATAGCATGATCTTCTNTCATCCGANANAAACAATCTATATGGTGA GTATCCGTCTGTGCATCCNTAGACANANNACCANGACTTGGNNNNTGNT

Appendix Table 4. (Continued)

TDF name	Sequence from 5'-3'
1E_3G	CAATACTTGATCATCTTGAGATTTGGCTCTGGCTACTGCTTTGCCATCAGTC CTTTANCAGATGCAGGTAGAGGGGTTCTCATGAAACTCCCCCCTCCCTGGC CTCAGCTCTTTGGGCTCCCCCATAAATCCTTTCTGAACCATGTGTACTTTTG AGGAGAATGCATCCCATGACATGTTTCCTCTGTTCAAAAATAAGGGATAGA GCCTTTTAGCATTTGGTCGGATCGTTCTCTTGTGCCCAAAGTATCTCCTTCG GCCAGCCAAAATTTTAGCCATGTTGCCATTGTTGTTAGCCACAAAAGCATC ACTTTCATCACAGACAATGAAATCAAGTGCAGCCATGCGAGAGGAATATTT AGAGAATGGAGCCAGCTCCTCTTTGCTAGATATTGTTTCTTTTGTGTGGAG
1E_3I	GCTCGCNTGNNCTGGTGGNTTNNNCNGCTCTAANNGCTCCTCCTGCTTCTT CGGTAACCTTCCCAACTTGGATGGAGTCCAGCTGGCAGNCCGGACTTGCCC CGCCCAAGGCGAATCTCTNTCTTGACCTGTCTCTCATCACCACATAAGAAA TAGGCATANAGAAGATGCTTCGCATTNCGTTTACTCAGGGACTCANCANNN ANGTGTTGTTTACNCGGGANTCNNNNNGCTCTATCNNNNNTGTCGGCGTCT ACATCNACGATCTCATCATCATCGGCGCGGAGGAACGTGAAGTGGAGGCG
2A-2_3F	TGTGTGCAGGAACAGCCTGCGCACATGATCGTTCTTTTTTCTCTTTTGGTTT GGCCCAGGGTCTCTGTCTTCATCTGGTTCCCGGGCCGCCTGTCAGCGGTAT GACTCTATATACTCAACACTATGAATGATGGACCCACTTTATGTCCTCTCG ACCGATGTGGAAAGGAAAATCACGCTTTAACCTTCATCTNCCCNCAGTGCA AAATTTTNNGGAANNAAGANCATTGCNTTCCCAATCTTGATGNCGGCTTAG AAGCNGCCATCANATACTCNATACTGCTNNNCGGATGCCATGGTTCCCTGC
2A_3L	ACGTGNGNTNCNATTTTGTTAGTGCTTAGTTGGTTTGGTTGCATGGATGAGC TGGTTTTTGGGAATAATAGTTTCTTGATTCGTCATGTTTCTGTGTTACTCAG GACTCANCNNTGACTNCNTCCNANTCCNTCAATCTTGNTACCNGNNTTGAT ACACNAGAATCATCANNGGGTGAGTCCGGTTTCNNCAAGTNCTGNAAAAA AGTNCNGGGGTNGTAGNANGGAATGNNTAGGGGANNGGATAGGAGAGAA
2DS_3G	GATTGAGAGGAATAGANTGCTTACTTGTNCCAGGTTTNNCATTTATTGTTTC TGCGCCCATCTCATAAGAACCCTCCTTCGATAGCTGAGTTGCTTTATTCCTG ACTTATCANCAGTGGGGNNNGGAGGCTTAACGGACGACTCAACACATTTC GGGTGACGTGGCCNTGGGCTCCTTCGAANANTCAGGCNNNCTGTGTGGNTT TACTCANTNCTCANCANATCGCTNGACCAGTGAGCTGTNACGCTNGCTTTA

Appendix Table 4 (Continued)

TDF name	Sequence from 5'-3'
2E_3H	ACTTGGGCCTGCTGCCCCCCAAAAAAATTACTCAGGACTCATCGGNTGCTC GTCAAATAAAGATGCACTTGTCTGTNTGTTCATCNGNTCACCTTGTTTACTC AGGACTCATCANNGTCTCATCGTAGATTCTACCNAAAATAGATGTAGAATG GTAGGAATCAAATCCTTTCTTCTCCATGAAGTGTGGATATGATTGTACTCTC AGTTTACGATCCACAATTACCTTGTGCCCATCCAGAGCCAAGTAATAAAGG TCAACTAATTTACTCAGGACTCA
3DS_3H	AGAGGTNCATGCAGATGCGGCTAGGCCCGTGCGGAGNNTTGGTTAGAAAG AAAAATGNNCCCCCTCATCTGAAGTAGAGAGTCATTTGGGTAAATCAANN GGTCTCCCCAGTAAGATGCANTGTACANNCGACTCANNNTCCATGTTCCAC NTTNCGGAAGATGATGATCTNTTTCCTTCAGGCTCCAGAGGAGGCTCTNGG AAAATTTACCCNGAATTTA
3E_3F	CCACTACTGTNNNNCGCTANNTCTGGGTCGTTCACTTGTTTTATGCTCCGTA CGGGTCGCNATTGCNATTTTTNAAGCTAGCTTCGACGATCTGAATGTGCAG GAATGACTAGCCTGATATAAGTGCTAATGGTATCTGGCCTGATTACCNNTT GTNGCTACCTCNACNCTNAAC
3E_3G	ACCTTTNANTTCATTGAGAGGACTAGCGTGCCTACTTGTTGTAGTGGCTTCA CATTTTCCTGGATTCTACGCCCATATAATAAGAACCCTCCTTCGATCGGTTT GTTGCTTTACTCAGGACTCATCANCAGTGCTGAGGGATGGCTTTACTCAGG ACTCANCNNNNGTCGGGTGAGATGGCCATGGNCGCCTTCGAAGAATCAGG CCATCTTTGTGGCTTTACTCAGGACTCANCANNGCGCTTNNNCNGGGATNN NNAANNCTTGCTTTACTCAGGACTCATCAN
4E_3H	ATAATTTCNATGCAGATGNCGACCAAGGACGCTTGGGTCTCTGAGGTTAGA ATGCGGGTAGTGCAGCTATCATCTGAAGAAGAGTGACATTTAAGGCATTCA CATCGNTCTTGGTTGAGAGNTGTTTGTGCNNNAATATTCAGCANCCATGTA ACCACAAGTTCCTACAACCGCATCCGTTTCTTTCTCATCTTCATCAGAATTT ACT

Appendix Table 5 Individual genotype frequency base on F_5 population DDYM x Mapila, analyzed by JoinMap® 3.0 program

Nr	Individual	a	h	b	c	d	-
1	301(1)	69	2	129	5	0	1
2	302(1)	82	2	121	1	0	0
3	304(1)	78	2	121	4	1	0
4	306(1)	108	2	94	1	1	0
5	307(1)	112	1	84	5	2	2
6	308(1)	133	0	66	6	1	0
7	309(1)	132	0	63	5	3	3
8	312B(1)	102	2	97	3	0	2
9	313(1)	142	0	58	3	2	1
10	314(1)	84	10	102	3	0	7
11	317(1)	99	15	87	2	2	1
12	319(2)	77	0	120	4	2	3
13	321(1)	82	0	116	5	2	1
14	323A(1)	100	4	95	5	2	0
15	323A(3)(2)	111	6	84	3	2	0
16	323B(1)	118	0	82	3	2	1
17	323B(2)	111	2	88	4	1	0
18	325(1)	86	2	114	4	0	0
19	329(1)	107	2	92	4	1	0
20	332A(2)	79	1	117	5	2	2
21	332B(1)	75	1	122	6	1	1
22	334B(1)	97	4	103	1	0	1
23	335(1)	96	0	103	4	3	0
24	339(1)	120	10	70	3	1	2
25	340A(1)	99	5	93	3	2	4
26	340B(1)	78	0	122	2	1	3
27	341(1)	132	2	64	3	3	2
28	342(1)	85	33	78	2	2	6
29	345(1)	129	0	70	2	2	3
30	347(1)	62	18	119	5	1	1
31	348A(1)	83	1	116	5	1	0
32	348B(1)	99	0	101	5	1	0
33	350(1)	92	4	102	4	1	3
34	352(1)	117	8	78	2	1	0
35	353A(1)	82	0	120	2	2	0
36	356(1)	103	0	98	3	2	0
37	357(1)	104	0	97	3	2	0
38	358(1)	115	0	83	2	2	4
39	361(1)	86	1	115	2	0	2

Appendix Table 5 (Continued)

Nr	Individual	a	h	b	c	d	-
40	363(1)	63	13	119	1	2	8
41	365(1)	93	6	101	2	2	2
42	366(1)	72	13	116	3	2	0
43	370(1)	96	3	97	6	1	3
44	372(1)	71	12	117	5	1	0
45	374(1)	118	2	83	2	1	0
46	377A(1)	108	0	92	3	3	0
47	378(1)	108	9	86	3	0	0
48	380(1)	79	27	92	4	2	2
49	385(1)	110	4	88	1	1	2
50	387(1)	84	1	112	6	1	2
51	388(1)	119	0	82	4	1	0
52	395(1)	87	3	109	5	2	0
53	396(1)	105	0	96	4	1	0
54	399(1)	96	11	92	3	3	1
55	402(1)	98	31	69	4	2	2
56	408(1)	96	9	95	4	2	0
57	410(1)	117	1	82	2	3	1
58	410(2)	116	3	82	2	3	0
59	411(1)	114	7	76	5	0	4
60	412(1)	89	8	99	4	1	5
61	417(2)	67	16	116	6	1	0
62	424(1)	80	4	115	3	1	3
63	428(1)	100	9	91	4	1	1
64	430(1)	87	2	109	3	0	5
65	436(1)	106	0	92	6	2	0
66	438(2)	109	2	86	6	2	1
67	441(1)	64	0	138	3	1	0
68	442A(1)	132	1	67	2	1	3
69	443(2)	76	1	125	4	0	0
70	444(1)	51	18	128	6	1	2
71	445(1)	113	3	85	2	2	1
72	446(3)	105	5	89	5	1	1
73	447(1)	87	1	107	6	0	5
74	449(1)	77	20	102	4	1	2
75	450(1)	113	12	73	5	2	1
76	451(1)	73	2	124	6	1	0
77	452(1)	110	10	79	3	2	2
78	455(1)	84	13	104	5	0	0
79	456(1)	90	8	98	5	0	5
80	459(1)	71	1	129	2	2	1

Appendix Table 5 (Continued)

Nr	Individual	a	h	b	c	d	-
81	463(1)	83	9	106	5	0	3
82	465A(1)	99	19	79	5	1	3
83	465A(2)	120	5	72	4	1	4
84	465A(3)	127	0	72	4	1	2
85	465B(1)	86	24	91	5	0	0
86	466(1)	97	23	80	4	2	0
87	466(2)	109	13	78	5	0	1
88	474(1)	111	3	81	3	2	6
89	475(1)	66	23	112	4	1	0
90	478(1)	127	0	76	3	0	0
91	479(1)	127	1	74	1	2	1
92	483(1)	82	1	118	5	0	0
93	484(1)	99	1	100	3	1	2
94	486(1)	84	9	111	2	0	0
95	487(1)	74	4	120	4	0	4
96	490(1)	65	11	126	3	0	1
97	492(1)	90	1	110	4	1	0
98	493(1)	82	10	107	5	0	2
99	494(1)	124	1	73	4	3	1
100	495(1)	83	5	110	6	2	0
101	495(2)	98	2	98	6	1	1
102	497(1)	92	10	95	5	1	3
103	498(1)	123	0	79	4	0	0
104	500(1)	87	28	84	6	1	0
105	501(1)	99	1	101	3	0	2
106	502(2)	92	7	99	4	1	3
107	503(2)	104	4	93	3	1	1
108	504(1)	104	4	89	5	2	2
109	505(1)	116	0	85	4	1	0
110	508(1)	110	7	85	3	1	0
111	509(1)	102	0	100	3	0	1
112	510(2)	81	3	117	3	0	2
113	511(1)	116	0	86	3	1	0
114	512(2)	106	11	83	4	0	2
115	513(1)	102	21	79	3	1	0
116	514(2)	102	3	91	3	1	6
117	515(2)	92	3	103	3	2	3
118	524(1)	103	0	96	6	0	1
119	526(1)	114	6	81	4	1	0
120	528(2)	87	5	111	1	0	2
121	531(1)	124	12	60	5	3	2

Appendix Table 5 (Continued)

Nr	Individual	a	h	b	c	d	-
122	532(1)	108	4	88	3	2	1
123	537(1)	63	16	120	3	1	3
124	538(1)	84	28	86	6	0	2
125	540(1)	94	9	100	1	1	1
126	540(2)	93	13	96	3	1	0
127	540(5)	86	16	100	3	0	1
128	545(1)	113	0	80	4	2	7
129	545(3)	89	0	111	4	0	2
130	560(1)	63	0	140	3	0	0
131	565(1)	76	37	85	5	1	2
132	586(1)	82	4	114	3	0	3
133	587(2)	110	5	87	4	0	0
134	75166(3)	5	1	194	6	0	0
135	75166(3)(4)	6	0	195	5	0	0
136	336	93	7	99	4	2	1
137	354	77	0	123	4	2	0
138	373	94	0	104	4	2	2
139	393(1)	128	5	68	3	2	0
140	429	67	16	111	4	2	6
141	453	95	6	95	2	2	6
142	11GHCL4A	113	2	84	5	2	0
143	11GHCL5C	81	0	121	2	1	1
144	11GHCL11C	101	9	93	2	0	1
145	11GHCL12B	95	16	92	1	2	0
146	11GHCL5A	86	6	105	2	1	6
147	11GHCL7A	89	21	93	3	0	0
148	11GHCL8B	101	4	95	5	1	0
149	11GHCL8C	101	5	93	2	2	3
150	11GHCL10A	72	11	116	5	2	0
151	11GHCL4B	82	18	102	3	1	0
152	11GHCL6A	68	1	133	4	0	0
153	11GHCL6B	79	28	94	3	2	0
154	11GHCL7B	95	28	77	4	2	0
155	11GHCL9B	81	1	117	6	1	0
156	11GHCL10B	102	5	92	5	1	1
157	11GHCL11A	108	9	86	2	0	1
158	383(1)	88	10	105	3	0	0
159	400(2)	84	29	87	3	2	1
160	409(1)	62	28	105	4	2	5
161	415(1)	100	4	97	2	2	1
162	421(3)	88	34	78	3	3	0

Appendix Table 5 (Continued)

Nr	Individual	a	h	b	c	d	-
163	423(1)	104	7	88	5	2	0
164	476(1)	125	0	76	2	1	2
165	384(1)	98	3	98	2	0	5
166	11GHD4unknow	75	17	103	2	1	8
167	371(1)	108	4	92	1	1	0
168	523(2)	114	9	78	3	1	1
169	11GHCL1A(1)	116	0	85	3	2	0
170	11GHCL1A(2)	104	13	83	3	2	1

Appendix Table 6 Locus genotype frequency base on F_5 population DDYM x Mapilaanalyzed using JoinMap® 3.0 program

Nr	Locus	a	h	b	c	d	-	X2	Df	Signif.	Classifi-cation
1	TX152	68	3	98	0	0	1	5.42	1	**	[a:b]
2	Drenhsbm63	117	0	53	0	0	0	24.1	1	*******	[a:b]
3	Drenhsbm64	88	1	81	0	0	0	0.29	1	-	[a:b]
4	XsBarlbk1.01	95	0	75	0	0	0	2.35	1	-	[a:b]
5	XsBarlbk1.39	85	1	84	0	0	0	0.01	1	-	[a:b]
6	XsBarlbk1.40	81	1	88	0	0	0	0.29	1	-	[a:b]
7	XsBarlbk1.61	94	5	70	0	0	1	3.51	1	*	[a:b]
8	XsBarlbk1.69	89	7	73	0	0	1	1.58	1	-	[a:b]
9	XsBarlbk1.72	105	3	62	0	0	0	11.1	1	*****	[a:b]
10	XsBarlbk2.74	46	0	123	0	0	1	35.1	1	*******	[a:b]
11	XsBarlbk2.75	45	0	124	0	0	1	36.9	1	*******	[a:b]
12	TX126	77	7	86	0	0	0	0.5	1	-	[a:b]
13	XsBarlbk3.62	81	9	80	0	0	0	0.01	1	-	[a:b]
14	XsBarlbk4.00	83	5	75	0	0	7	0.41	1	-	[a:b]
15	XsBarlbk4.50	88	10	71	0	0	1	1.82	1	-	[a:b]
16	XsBarlbk4.54	82	10	77	0	0	1	0.16	1	-	[a:b]
17	XsBarlbk5.08	96	8	63	0	0	3	6.85	1	***	[a:b]
18	XsBarlbk5.56	79	5	84	0	0	2	0.15	1	-	[a:b]
19	XsBarlbk5.16	84	5	79	0	0	2	0.15	1	-	[a:b]
20	Drenhsbm03	97	9	64	0	0	0	6.76	1	***	[a:b]
21	XCUP57	99	8	63	0	0	0	8	1	****	[a:b]
22	XsBarlbk7.59	85	3	80	0	0	2	0.15	1	-	[a:b]
23	Xtxtp273	94	2	74	0	0	0	2.38	1	-	[a:b]
24	Drenhsbm16	94	1	72	0	0	3	2.92	1	*	[a:b]

Appendix Table 6 (Continued)

Nr	Locus	a	h	b	c	d	-	X2	Df	Signif.	Classification
25	Drenhsbm60	72	10	86	0	0	2	1.24	1	-	[a:b]
26	XsBarlbk8.05	64	10	95	0	0	1	6.04	1	**	[a:b]
27	XsBarlbk8.49	62	5	96	0	0	7	7.32	1	***	[a:b]
28	XsBarlbk9.01	83	8	79	0	0	0	0.1	1	-	[a:b]
29	XsBarlbk9.55	74	9	87	0	0	0	1.05	1	-	[a:b]
30	XsBarlbk9.09	61	6	103	0	0	0	10.8	1	****	[a:b]
31	XsBarlbk9.13	78	2	90	0	0	0	0.86	1	-	[a:b]
32	XsBarlbk9.45	74	9	85	0	0	2	0.76	1	-	[a:b]
33	XsBarlbk9.07	80	9	80	0	0	1	0	1	-	[a:b]
34	XCUP07	77	3	90	0	0	0	1.01	1	-	[a:b]
35	XCUP50	75	7	88	0	0	0	1.04	1	-	[a:b]
36	XsBarlbk10.00	66	6	96	0	0	2	5.56	1	**	[a:b]
37	Xtxtp01	68	2	97	0	0	3	5.1	1	**	[a:b]
38	Xtxtp03	83	1	84	0	0	2	0.01	1	-	[a:b]
39	Xtxtp04	87	0	83	0	0	0	0.09	1	-	[a:b]
40	Xtxtp07	54	10	102	0	0	4	14.8	1	******	[a:b]
41	Xtxtp08	47	8	112	0	0	3	26.6	1	*******	[a:b]
42	Xtxtp09	71	5	93	0	0	1	2.95	1	*	[a:b]
43	Xtxtp10	66	11	93	0	0	0	4.58	1	**	[a:b]
44	Xtxtp11	76	9	84	0	0	1	0.4	1	-	[a:b]
45	Xtxtp12	91	0	79	0	0	0	0.85	1	-	[a:b]
46	Xtxtp15	64	5	99	0	0	2	7.52	1	***	[a:b]
47	Xtxtp18	57	13	100	0	0	0	11.8	1	*****	[a:b]
48	Xtxtp19	81	1	85	0	0	3	0.1	1	-	[a:b]
49	Xtxtp21	50	6	113	0	0	1	24.4	1	*******	[a:b]
50	Xtxtp23	99	0	69	0	0	2	5.36	1	**	[a:b]
51	Xtxtp29	85	13	71	0	0	1	1.26	1	-	[a:b]
52	Xtxtp34	67	1	102	0	0	0	7.25	1	***	[a:b]
53	Xtxtp37	86	8	76	0	0	0	0.62	1	-	[a:b]
54	Xtxtp38	96	0	74	0	0	0	2.85	1	*	[a:b]
55	Xtxtp41	84	11	73	0	0	2	0.77	1	-	[a:b]
56	Xtxtp43	68	4	97	0	0	1	5.1	1	**	[a:b]
57	Drenhsbm04	72	13	82	0	0	3	0.65	1	-	[a:b]
58	Drenhsbm17	64	13	92	0	0	1	5.03	1	**	[a:b]
59	Drenhsbm39	78	8	80	0	0	4	0.03	1	-	[a:b]
60	Drenhsbm48	80	6	83	0	0	1	0.06	1	-	[a:b]
61	Drenhsbm50	79	9	80	0	0	2	0.01	1	-	[a:b]
62	Drenhsbm57	69	7	92	0	0	2	3.29	1	*	[a:b]
63	Drenhsbm65	82	6	81	0	0	1	0.01	1	-	[a:b]
64	Drenhsbm66	70	11	89	0	0	0	2.27	1	-	[a:b]

Appendix Table 6 (Continued)

Nr	Locus	a	h	b	c	d	-	X2	Df	Signif.	Classifi-cation
65	Drenhsbm72	70	1	99	0	0	0	4.98	1	**	[a:b]
66	Drenhsbm76	97	0	71	0	0	2	4.02	1	**	[a:b]
67	Drenhsbm77	92	4	74	0	0	0	1.95	1	-	[a:b]
68	Drenhsbm79	29	6	133	0	0	2	66.8	1	*******	[a:b]
69	Drenhsbm83	80	0	90	0	0	0	0.59	1	-	[a:b]
70	Drenhsbm86	95	6	69	0	0	0	4.12	1	**	[a:b]
71	Drenhsbm96	82	0	88	0	0	0	0.21	1	-	[a:b]
72	XCUP05	94	5	71	0	0	0	3.21	1	*	[a:b]
73	XCUP38	74	10	85	0	0	1	0.76	1	-	[a:b]
74	XCUP48	94	0	65	0	0	11	5.29	1	**	[a:b]
75	XCUP49	74	6	90	0	0	0	1.56	1	-	[a:b]
76	XCUP69	65	11	93	0	0	1	4.96	1	**	[a:b]
77	XCUP70	98	4	68	0	0	0	5.42	1	**	[a:b]
78	XCUP305	95	5	70	0	0	0	3.79	1	*	[a:b]
79	XCUP314	82	6	82	0	0	0	0	1	-	[a:b]
80	XCUP320	98	4	67	0	0	1	5.82	1	**	[a:b]
81	XCUP324	43	0	117	0	0	10	34.2	1	*******	[a:b]
82	XCUP350	80	6	84	0	0	0	0.1	1	-	[a:b]
83	XCUP357	99	9	60	0	0	2	9.57	1	****	[a:b]
84	Flo08	44	7	118	0	0	1	33.8	1	*******	[a:b]
85	Staygreen44	73	0	97	0	0	0	3.39	1	*	[a:b]
86	XsBarlbk1.06	77	0	93	0	0	0	1.51	1	-	[a:b]
87	XsBarlbk1.27	74	14	82	0	0	0	0.41	1	-	[a:b]
88	XsBarlbk1.51	74	8	88	0	0	0	1.21	1	-	[a:b]
89	XsBarlbk1.63	118	1	50	0	0	1	27.5	1	*******	[a:b]
90	XsBarlbk1.66	109	2	59	0	0	0	14.9	1	******	[a:b]
91	XsBarlbk2.48	69	5	96	0	0	0	4.42	1	**	[a:b]
92	XsBarlbk2.62	53	7	110	0	0	0	19.9	1	*******	[a:b]
93	XsBarlbk2.64	62	7	101	0	0	0	9.33	1	****	[a:b]
94	XsBarlbk2.76	58	5	104	0	0	3	13.1	1	******	[a:b]
95	XsBarlbk3.00	84	6	79	0	0	1	0.15	1	-	[a:b]
96	XsBarlbk3.10	72	11	81	0	0	6	0.53	1	-	[a:b]
97	XsBarlbk3.16	81	10	79	0	0	0	0.02	1	-	[a:b]
98	XsBarlbk3.74	79	8	81	0	0	2	0.02	1	-	[a:b]
99	XsBarlbk4.12	89	0	81	0	0	0	0.38	1	-	[a:b]
100	XsBarlbk4.13	80	7	81	0	0	2	0.01	1	-	[a:b]
101	XsBarlbk4.62	70	14	85	0	0	1	1.45	1	-	[a:b]
102	XsBarlbk5.17	91	8	71	0	0	0	2.47	1	-	[a:b]
103	XsBarlbk7.01	68	6	95	0	0	1	4.47	1	**	[a:b]
104	XsBarlbk7.02	64	14	89	0	0	3	4.08	1	**	[a:b]

Appendix Table 6 (Continued)

Nr	Locus	a	h	b	c	d	-	X2	Df	Signif.	Classifi-cation
105	XsBarlbk7.03	79	13	74	0	0	4	0.16	1	-	[a:b]
106	XsBarlbk7.04	82	15	72	0	0	1	0.65	1	-	[a:b]
107	TX06	81	4	85	0	0	0	0.1	1	-	[a:b]
108	TX08	66	11	93	0	0	0	4.58	1	**	[a:b]
109	TX11	75	4	91	0	0	0	1.54	1	-	[a:b]
110	TX12	66	7	95	0	0	2	5.22	1	**	[a:b]
111	TX21	64	8	94	0	0	4	5.7	1	**	[a:b]
112	TX25	111	4	54	0	0	1	19.7	1	*******	[a:b]
113	TX33	80	9	77	0	0	4	0.06	1	-	[a:b]
114	TX34	85	2	83	0	0	0	0.02	1	-	[a:b]
115	TX37	87	8	75	0	0	0	0.89	1	-	[a:b]
116	TX43	98	8	64	0	0	0	7.14	1	***	[a:b]
117	TX45	92	10	68	0	0	0	3.6	1	*	[a:b]
118	TX47	48	10	112	0	0	0	25.6	1	*******	[a:b]
119	TX48	97	8	65	0	0	0	6.32	1	**	[a:b]
120	TX54	83	3	84	0	0	0	0.01	1	-	[a:b]
121	TX61	79	5	85	0	0	1	0.22	1	-	[a:b]
122	TX64	82	9	77	0	0	2	0.16	1	-	[a:b]
123	TX71	71	10	87	0	0	2	1.62	1	-	[a:b]
124	TX73	88	8	73	0	0	1	1.4	1	-	[a:b]
125	TX75	96	2	70	0	0	2	4.07	1	**	[a:b]
126	TX80	71	5	94	0	0	0	3.21	1	*	[a:b]
127	TX85	85	0	83	0	0	2	0.02	1	-	[a:b]
128	TX88	55	6	107	0	0	2	16.7	1	*******	[a:b]
129	TX92	76	11	83	0	0	0	0.31	1	-	[a:b]
130	TX98	98	8	63	0	0	1	7.61	1	***	[a:b]
131	TX106	108	0	62	0	0	0	12.5	1	******	[a:b]
132	TX109	83	5	72	0	0	10	0.78	1	-	[a:b]
133	TX110	86	9	73	0	0	2	1.06	1	-	[a:b]
134	TX115	108	7	54	0	0	1	18	1	*******	[a:b]
135	TX117	101	6	63	0	0	0	8.8	1	****	[a:b]
136	TX118	61	6	102	0	0	1	10.3	1	****	[a:b]
137	TX124	78	6	73	0	0	13	0.17	1	-	[a:b]
138	TX127	74	7	88	0	0	1	1.21	1	-	[a:b]
139	TX130	66	10	94	0	0	0	4.9	1	**	[a:b]
140	TX136	95	0	74	0	0	1	2.61	1	-	[a:b]
141	TX137	82	15	70	0	0	3	0.95	1	-	[a:b]
142	TX139	56	13	101	0	0	0	12.9	1	******	[a:b]
143	TX140	84	9	77	0	0	0	0.3	1	-	[a:b]
144	TX144	88	3	78	0	0	1	0.6	1	-	[a:b]

Appendix Table 6 (Continued)

Nr	Locus	a	h	b	c	d	-	X2	Df	Signif.	Classifi-cation
145	TX145	72	12	85	0	0	1	1.08	1	-	[a:b]
146	TX157	78	8	83	0	0	1	0.16	1	-	[a:b]
147	umc1246	80	0	89	0	0	1	0.48	1	-	[a:b]
148	umc1449	75	5	89	0	0	1	1.2	1	-	[a:b]
149	umc1596	74	5	90	0	0	1	1.56	1	-	[a:b]
150	bnlg1131	75	10	85	0	0	0	0.62	1	-	[a:b]
151	SAM04318	92	10	68	0	0	0	3.6	1	*	[a:b]
152	SAM05060	74	0	95	0	0	1	2.61	1	-	[a:b]
153	SAM05934C	87	3	77	0	0	3	0.61	1	-	[a:b]
154	SAM10627	99	4	63	0	0	4	8	1	****	[a:b]
155	SAM11159	73	8	86	0	0	3	1.06	1	-	[a:b]
156	SAM14426	71	11	82	0	0	6	0.79	1	-	[a:b]
157	SAM14927	71	12	86	0	0	1	1.43	1	-	[a:b]
158	SAM16073	71	14	84	0	0	1	1.09	1	-	[a:b]
159	SAM16246	40	8	122	0	0	0	41.5	1	*******	[a:b]
160	SAM17940	62	2	105	0	0	1	11.1	1	*****	[a:b]
161	SAM18581	82	4	82	0	0	2	0	1	-	[a:b]
162	SAM33545	83	7	79	0	0	1	0.1	1	-	[a:b]
163	SAM34164	85	9	73	0	0	3	0.91	1	-	[a:b]
164	SAM36890	90	7	73	0	0	0	1.77	1	-	[a:b]
165	SAM38058	84	5	81	0	0	0	0.05	1	-	[a:b]
166	SAM38304	75	5	90	0	0	0	1.36	1	-	[a:b]
167	SAM38921	95	3	71	0	0	1	3.47	1	*	[a:b]
168	SAM39963	61	9	98	0	0	2	8.61	1	****	[a:b]
169	SAM47801	76	5	89	0	0	0	1.02	1	-	[a:b]
170	SAM50260b	77	6	87	0	0	0	0.61	1	-	[a:b]
171	SAM51932	114	5	50	0	0	1	25	1	*******	[a:b]
172	SAM51941	116	6	48	0	0	0	28.2	1	*******	[a:b]
173	SAM52852	71	2	95	0	0	2	3.47	1	*	[a:b]
174	SAM54806	78	6	82	0	0	4	0.1	1	-	[a:b]
175	SAM54949	84	2	80	0	0	4	0.1	1	-	[a:b]
176	SAM55268b	60	0	110	0	0	0	14.7	1	******	[a:b]
177	SAM56441a	79	8	83	0	0	0	0.1	1	-	[a:b]
178	SAM56359	88	7	74	0	0	1	1.21	1	-	[a:b]
179	SAM57962	75	12	81	0	0	2	0.23	1	-	[a:b]
180	SAM59132a	85	4	81	0	0	0	0.1	1	-	[a:b]
181	SAM60877	63	5	101	0	0	1	8.8	1	****	[a:b]
182	SAM61376	83	4	83	0	0	0	0	1	-	[a:b]
183	SAM61728	70	7	93	0	0	0	3.25	1	*	[a:b]
184	SAM62005	106	7	56	0	0	1	15.4	1	*******	[a:b]

Appendix Table 6 (Continued)

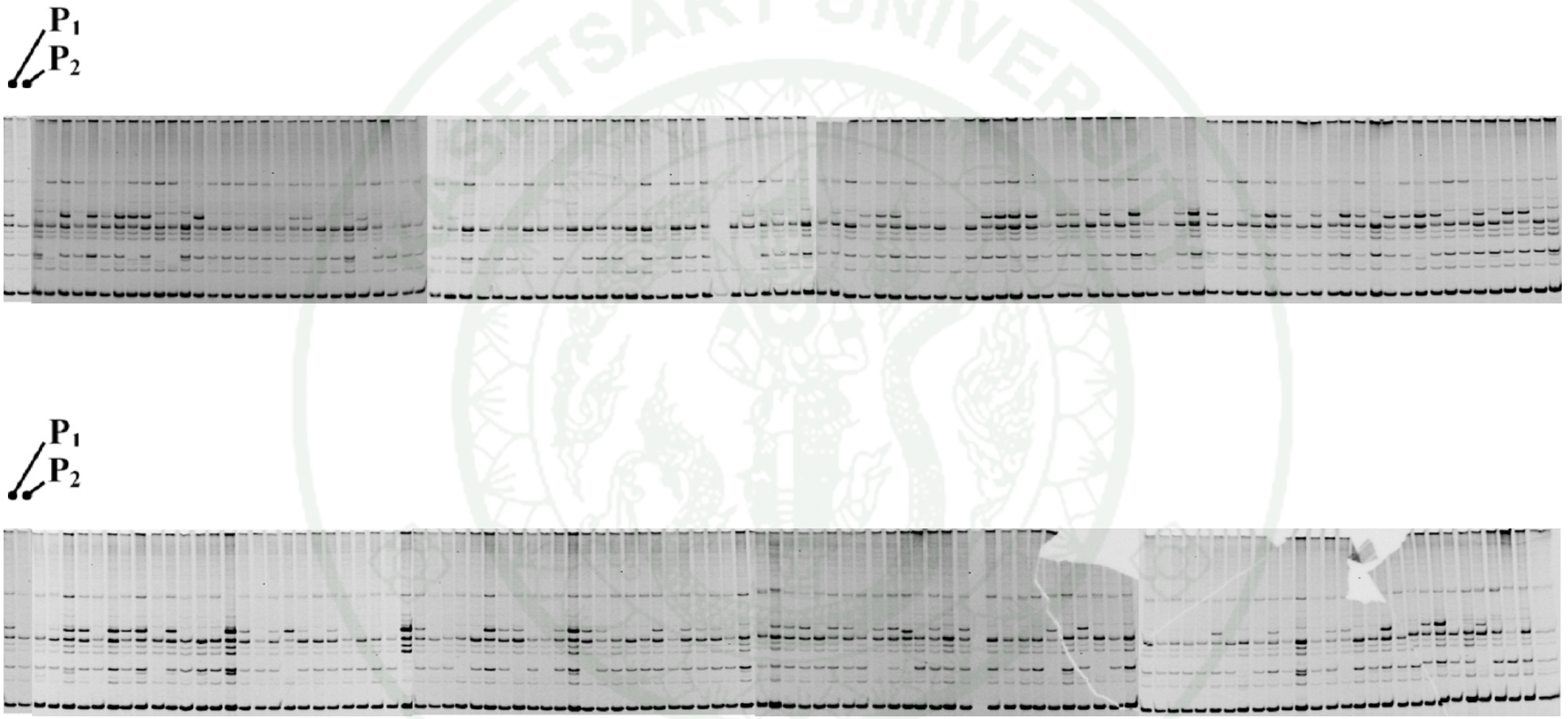
Nr	Locus	a	h	b	c	d	-	X2	Df	Signif.	Classifi-cation
185	SAM63115	70	3	94	0	0	3	3.51	1	*	[a:b]
186	SAM65307a	83	7	79	0	0	1	0.1	1	-	[a:b]
187	SAM68768	98	8	64	0	0	0	7.14	1	***	[a:b]
188	SAM72523	89	12	69	0	0	0	2.53	1	-	[a:b]
189	SAM72965	73	1	96	0	0	0	3.13	1	*	[a:b]
190	SAM73036	84	6	80	0	0	0	0.1	1	-	[a:b]
191	SAM74034	86	8	76	0	0	0	0.62	1	-	[a:b]
192	SAM74922	76	12	82	0	0	0	0.23	1	-	[a:b]
193	SAM75510	83	15	72	0	0	0	0.78	1	-	[a:b]
194	SAM75705	107	6	56	0	0	1	16	1	*******	[a:b]
195	TDF3DS.1.3H	83	3	84	0	0	0	0.01	1	-	[a:b]
196	TDF3DS.2.3H	0	0	91	0	79	0	91	1	*******	[a:b]
197	TDF3DS.3.3H	0	0	86	0	83	1	86	1	*******	[a:b]
198	TDF3F.1M	91	0	0	78	0	1	91	1	*******	[a:b]
199	TDF1.2E.1.3G	104	0	0	66	0	0	104	1	*******	[a:b]
200	TDF1.2E.2.3G	23	0	0	147	0	0	23	1	*******	[a:b]
201	TDF2DS.3E	79	0	0	91	0	0	79	1	*******	[a:b]
202	TDF2A.3C	78	0	0	91	0	1	78	1	*******	[a:b]
203	TDF2E.3H	115	0	0	54	0	1	115	1	*******	[a:b]
204	TDF6F.1Y	80	0	0	90	0	0	80	1	*******	[a:b]
205	TDF3.2E.3F	0	0	129	0	39	2	129	1	*******	[a:b]
206	TDF4DS.1X	72	7	90	0	0	1	2	1	-	[a:b]

Appendix Table 7 Similarity of individual base on F_5 population DDYM x Mapila, analyzed by JoinMap® 3.0 program

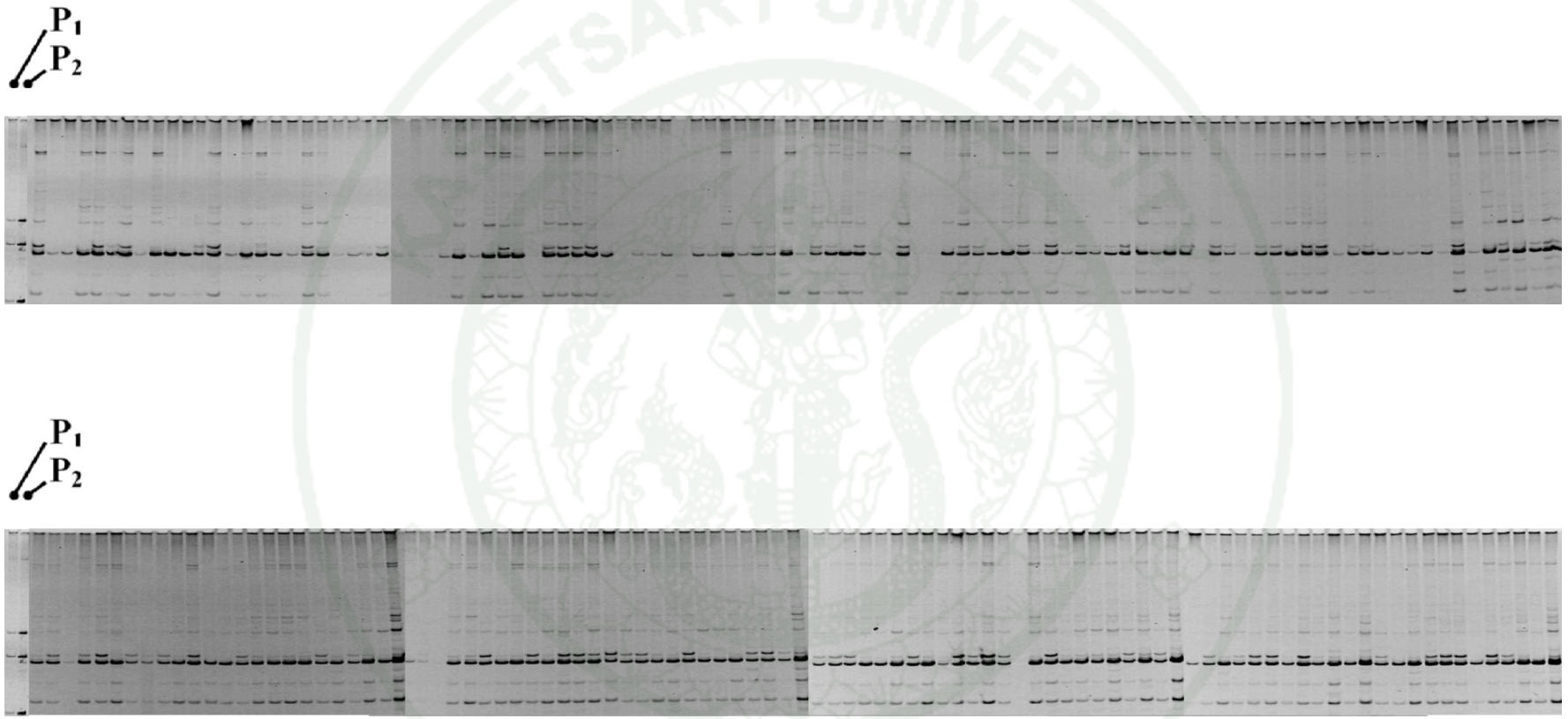
Nr1	Individual1	Nr2	Individual2	Similarity
6	308(1)	7	309(1)	0.951
27	341(1)	29	345(1)	0.951
57	410(1)	58	410(2)	0.971
134	75166(3)	135	75166(3)(4)	0.981

Appendix Table 8 Similarity of loci base on F_5 population DDYM x Mapila, analyzed by JoinMap® 3.0 program

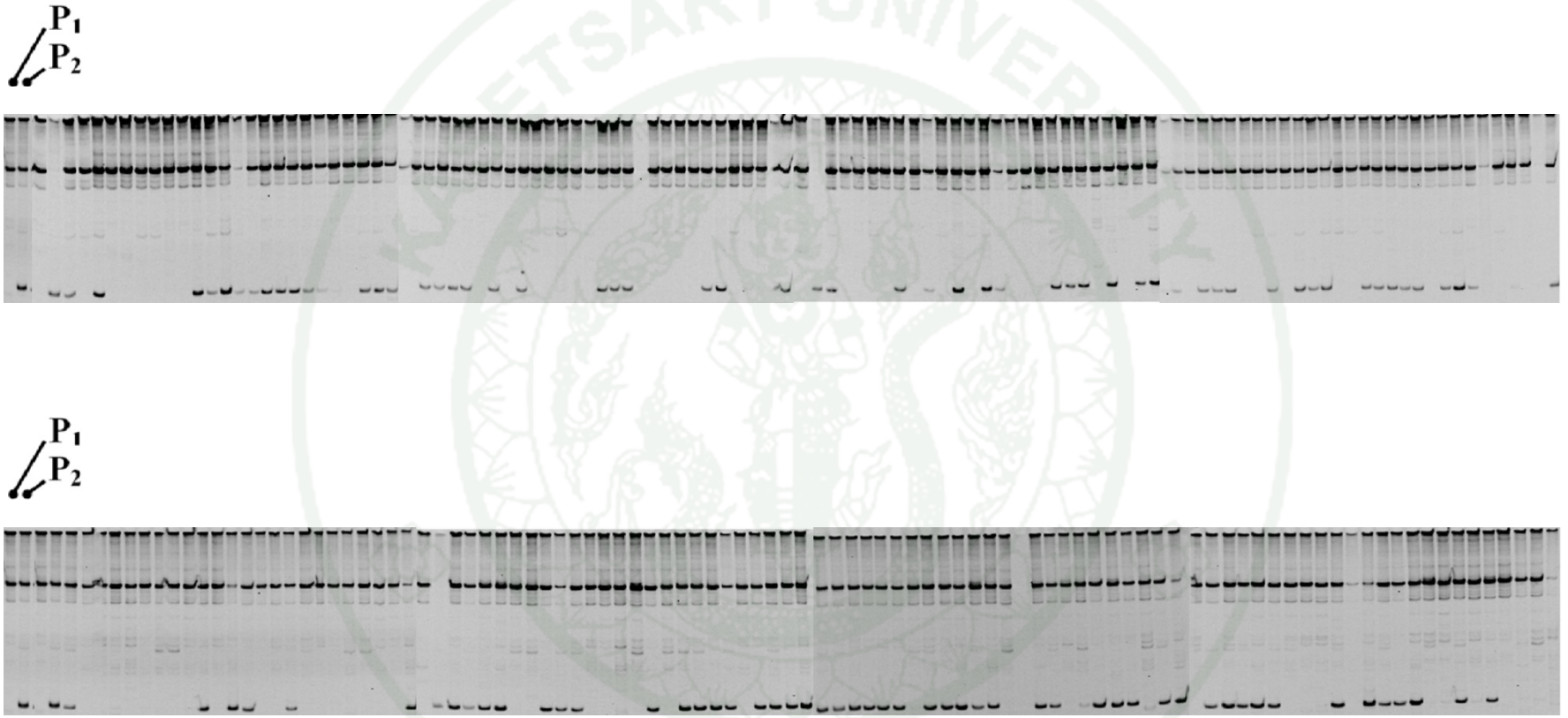
Nr1	Locus1	Nr2	Locus2	Similarity
10	XsBarlbk2.74	11	XsBarlbk2.75	0.965
21	XCUP57	83	XCUP357	0.965
38	Xtxtp03	127	TX85	0.965
57	Drenhsbm04	158	SAM16073	0.953
72	XCUP05	78	XCUP305	0.994
75	XCUP49	169	SAM47801	0.976
82	XCUP350	169	SAM47801	0.971
82	XCUP350	174	SAM54806	0.959
112	TX25	171	SAM51932	0.953
119	TX48	130	TX98	0.988
169	SAM47801	174	SAM54806	0.959



Appendix Figure 1 The Polymorphic pattern of sfw1.2E.3G sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila

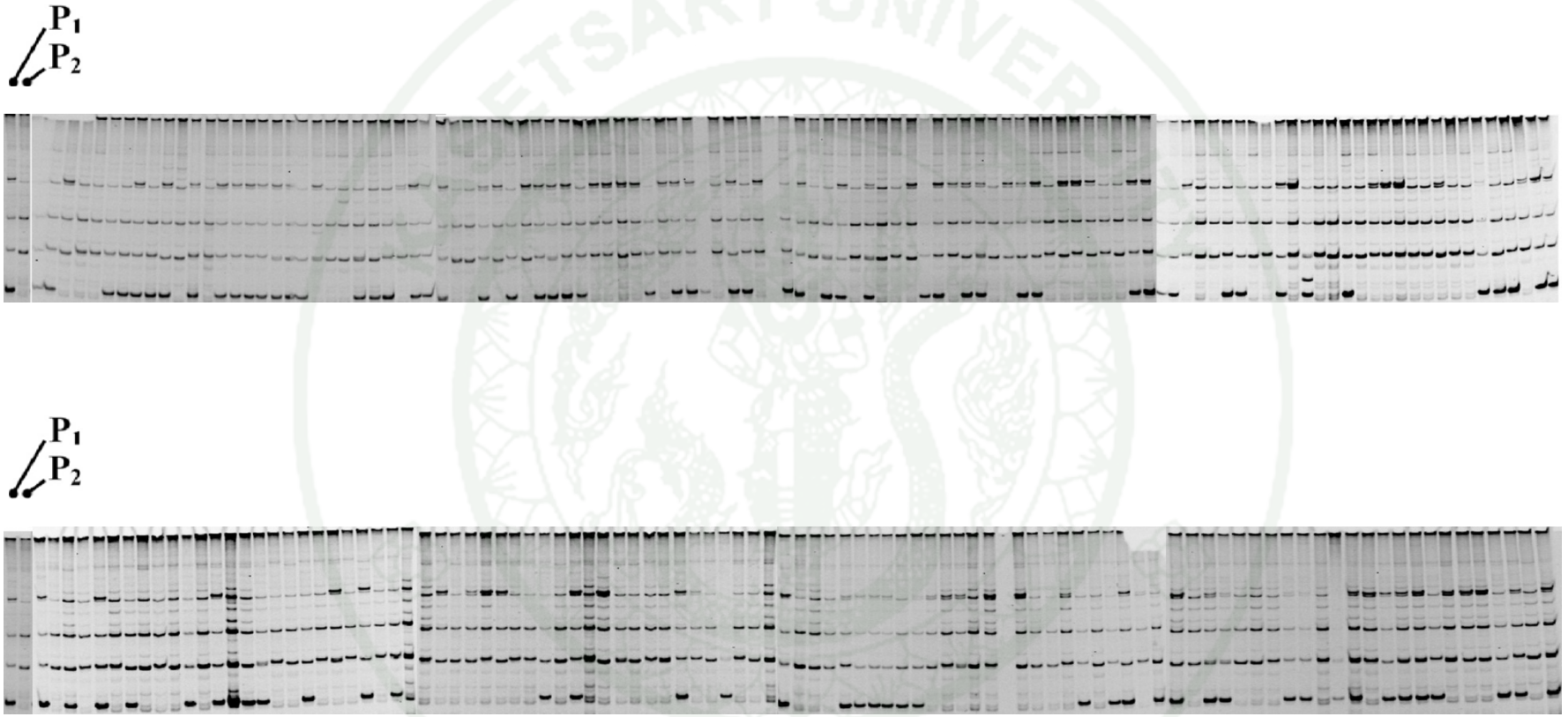


Appendix Figure 2 The Polymorphic pattern of sfw2A.3C sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila

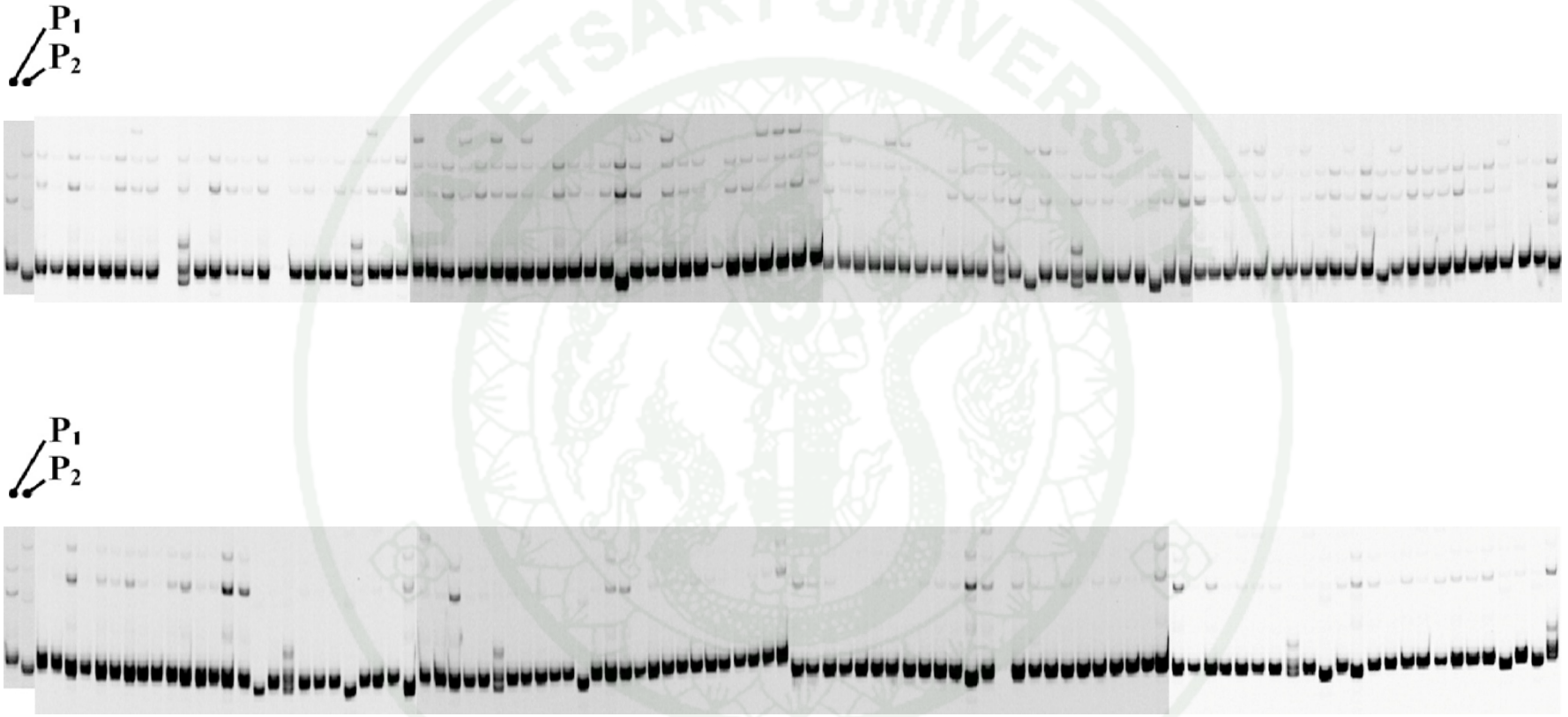


Appendix Figure 3 The Polymorphic pattern of sfw2DS.3E sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila

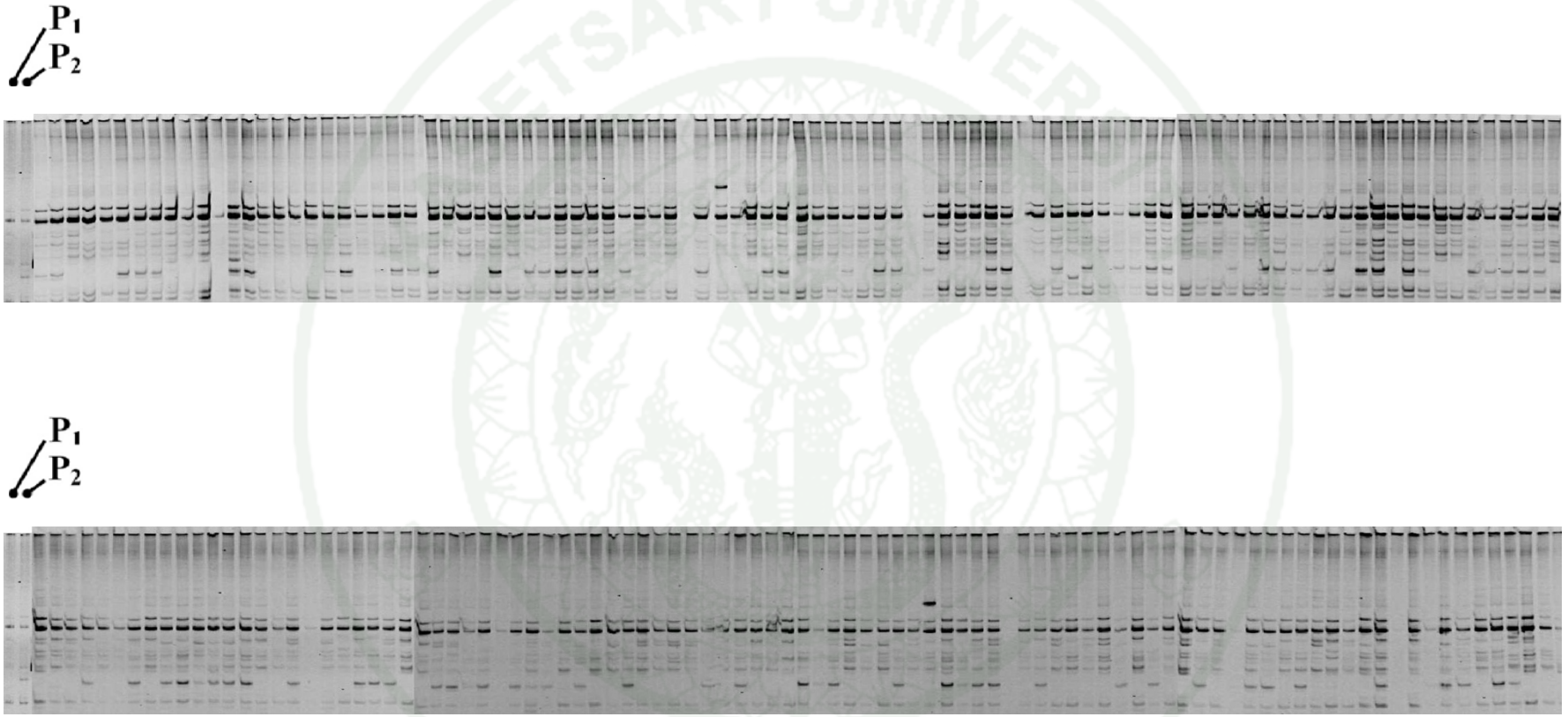
Appendix Figure 4 The Polymorphic pattern of sfw2E.3H sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila



Appendix Figure 5 The Polymorphic pattern of sfw3DS.3H sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila



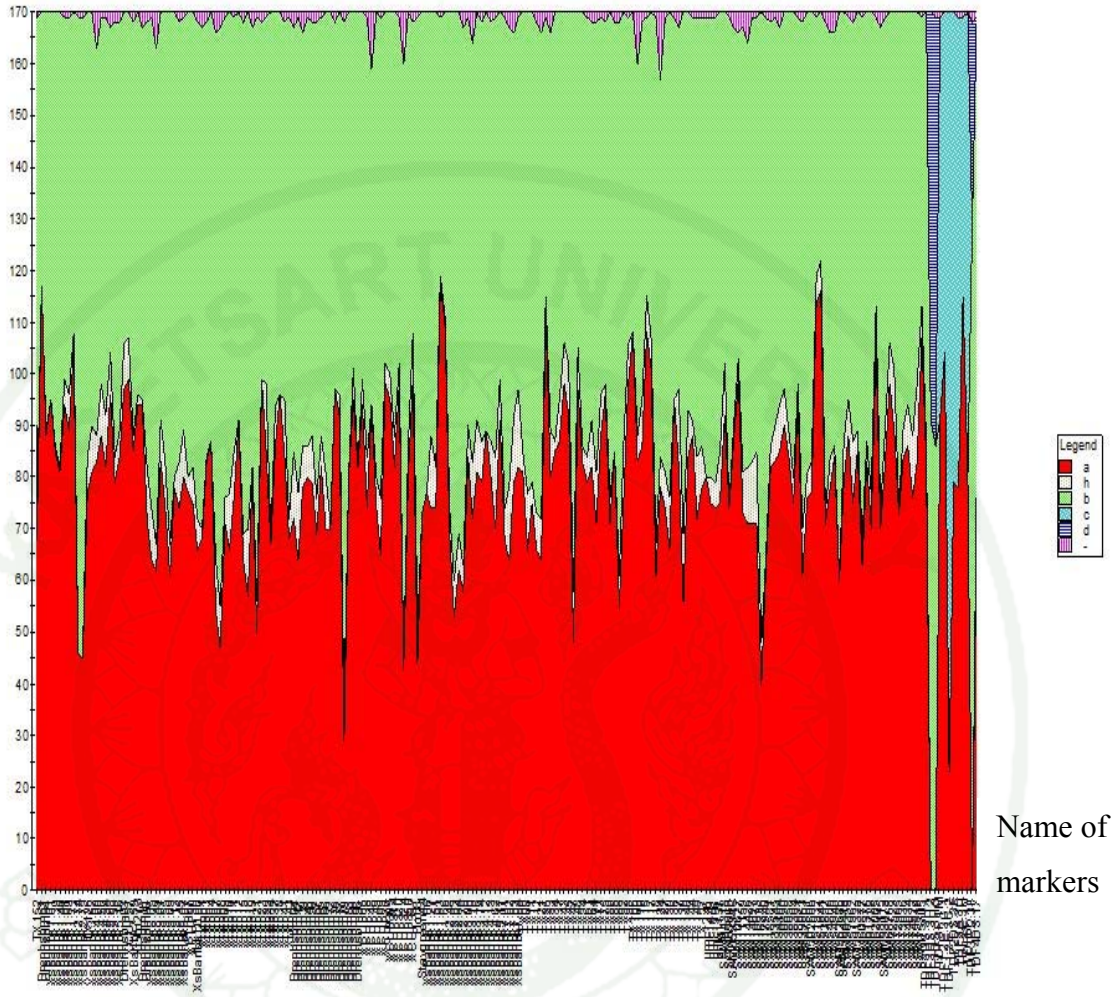
Appendix Figure 6 The Polymorphic pattern of sfw3.2E.3F sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila



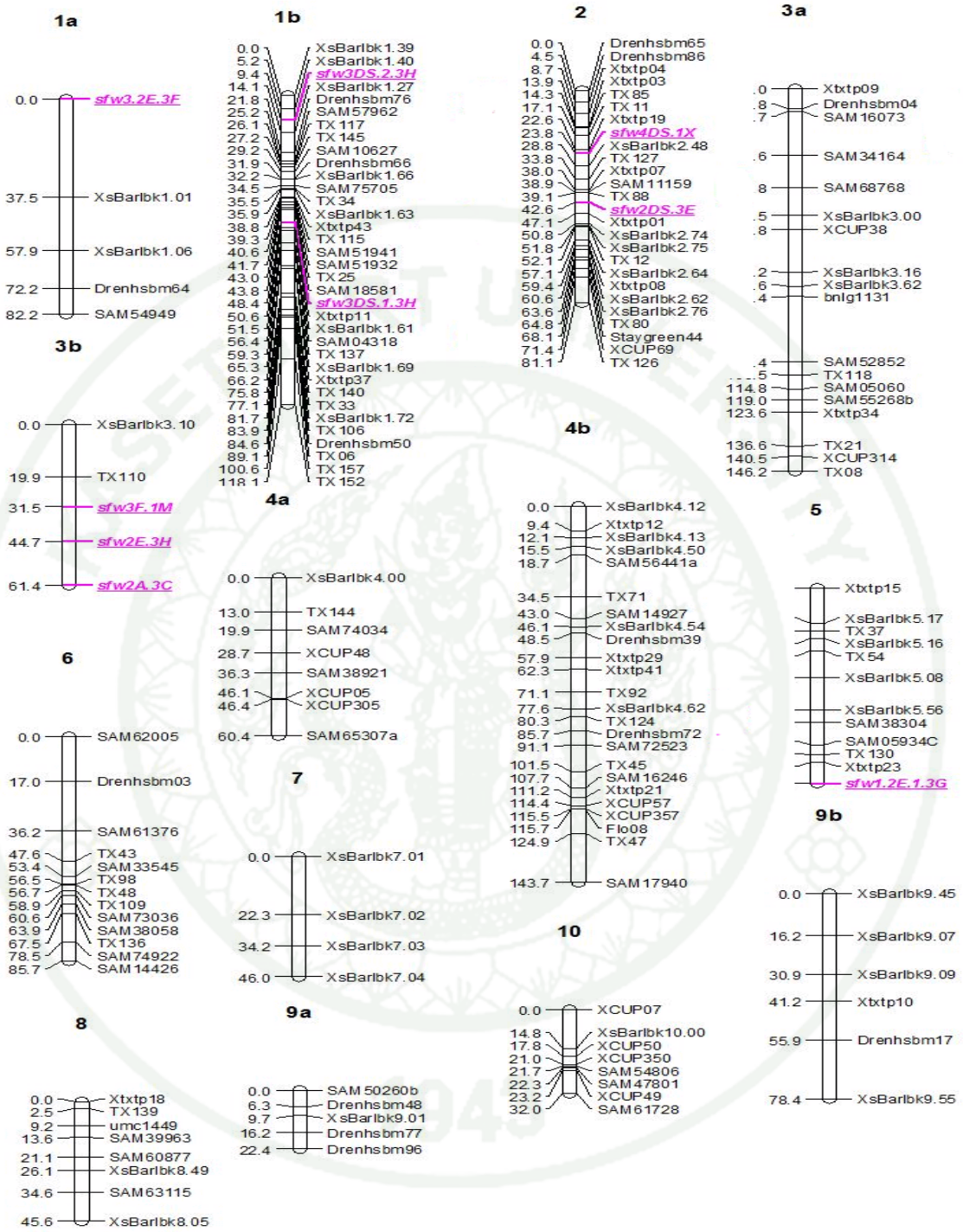
Appendix Figure 7 The Polymorphic pattern of sfw3F.1M sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila

Appendix Figure 8 The Polymorphic pattern of sfw4DS.1X sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila

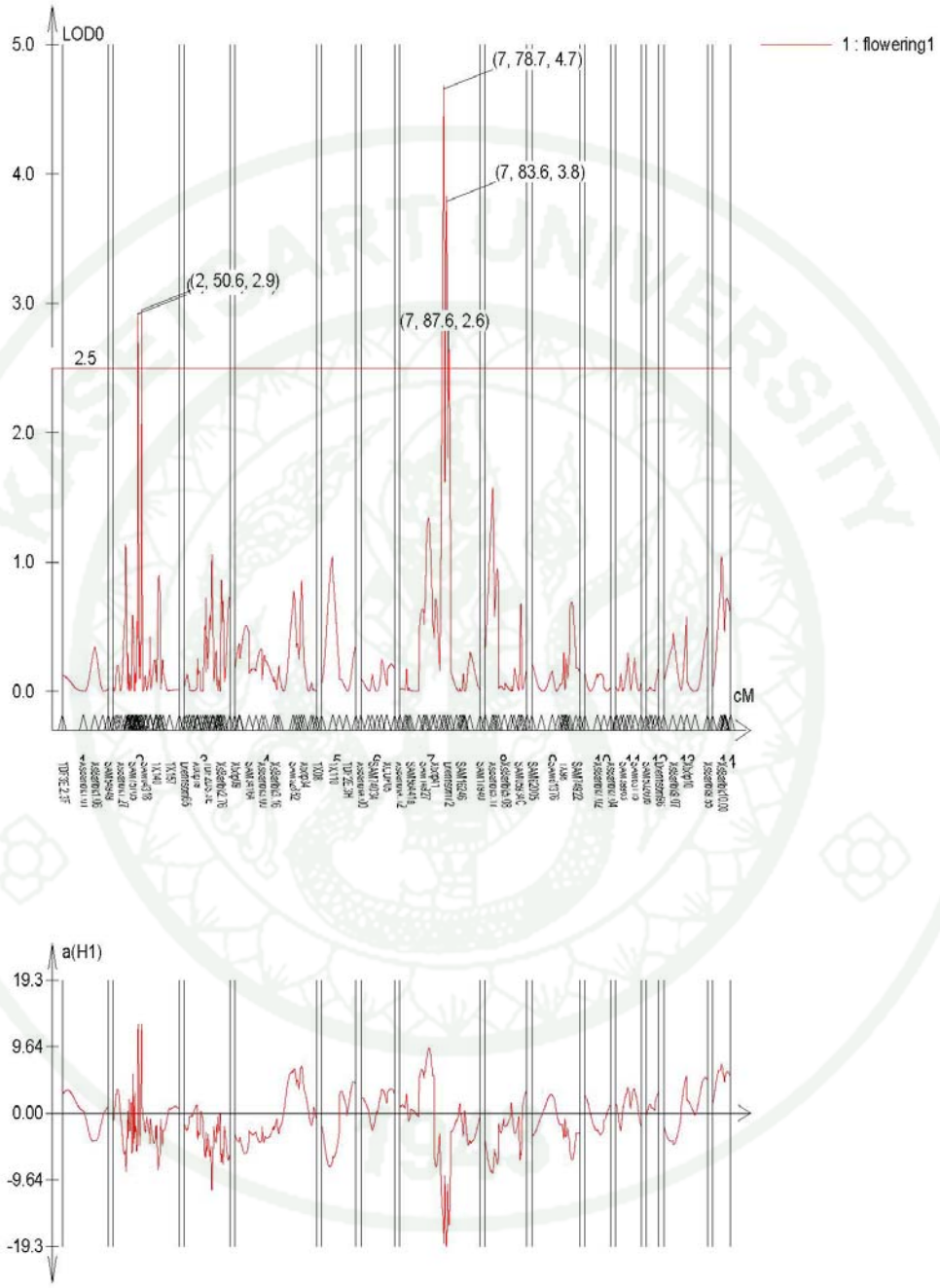
Appendix Figure 9 The Polymorphic pattern of sfw6F.1Y sugarcane TDF with F5 sorghum RIL derived from DDYM x Mapila



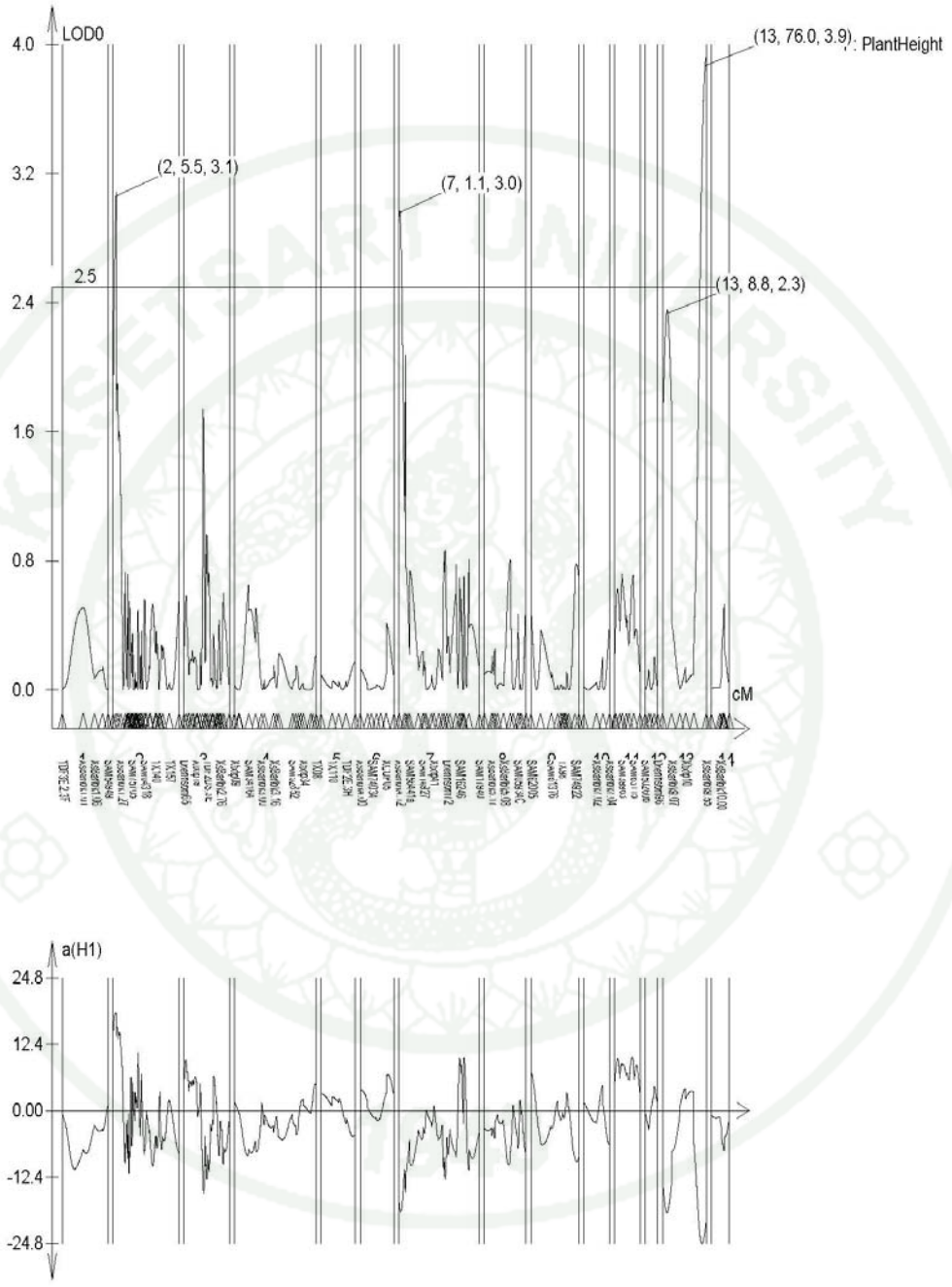
Appendix Figure 10 The proportion between a (red area represented allele A that derived from P1), b (green area represented allele B that derived from P2), h (white area represented heterozygous), c and d (blue and pink area represented allele that derived from P1 or P2, respectively) and "–" or pink area represented missing data.



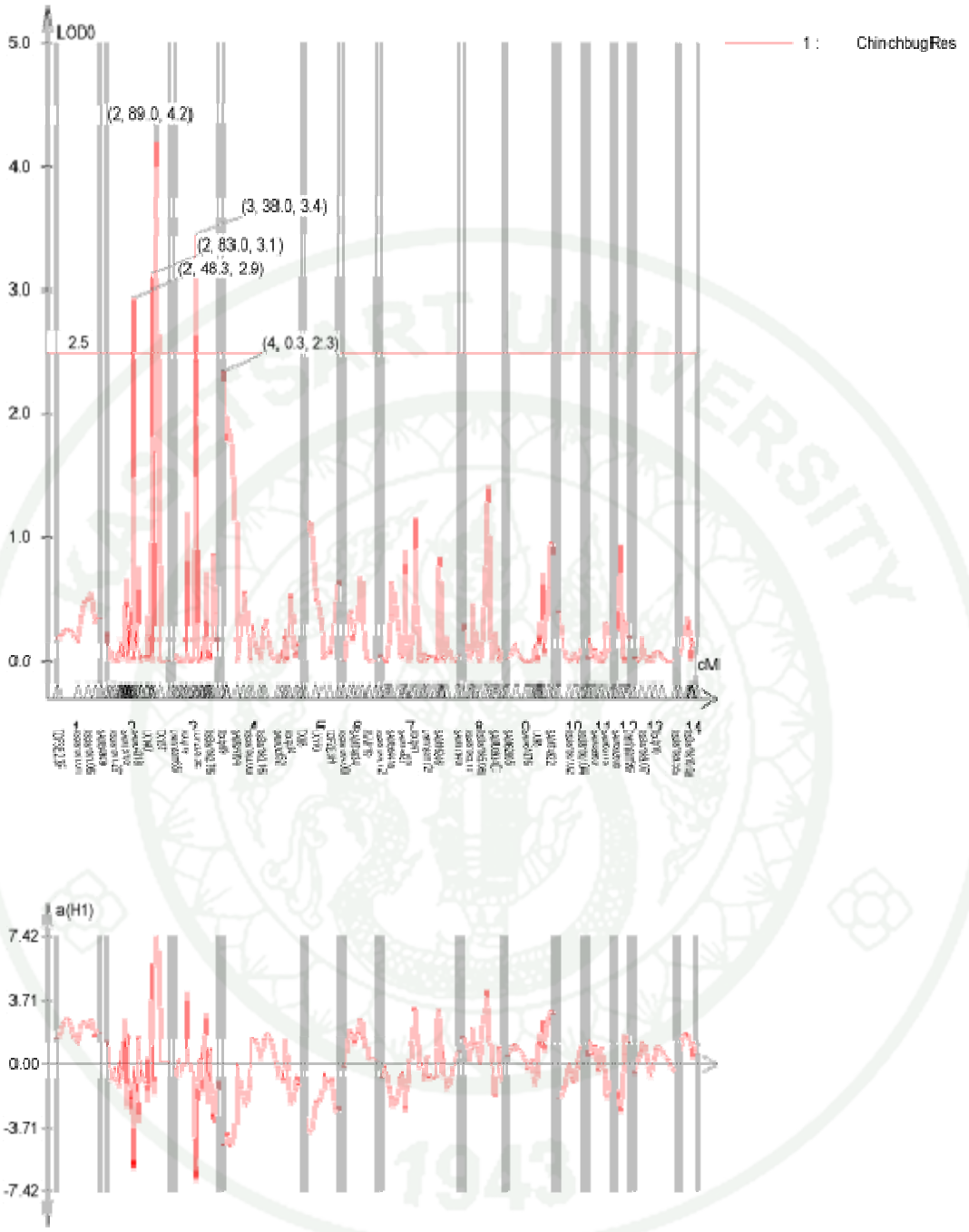
Appendix Figure 11 The sorghum genetic linkage map associated with SSR markers and sugarcane specific- flowering based on F_5 RILs population of DDYM x Mapila.



Appendix Figure 12 Chromatograph of flowering QTLs detected using Window QTL Cartographer



Appendix Figure 13 Chromatograph of plant height QTLs detected using Window QTL Cartographer



Appendix Figure 14 Chromatograph of chinch bug resistance QTLs detected using Window QTL Cartographer

CURRICULUM VITAE

NAME : Miss. Pattama Srinamngoen

BIRTH DATE : September 21, 1977

BIRTH PLACE : Kamphaeng Phet, Thailand

EDUCATION : **YEAR INSTITUTE DEGREE/DIPLOMA**

1999	KMITL.	B.Sc. (Agriculture)
2003	Kasetsart Univ.	M.S. (Agriculture)

POSITION/TITLE : Lecturer

WORK PLACE : Faculty of Science and Art, Burapha University, Chanthaburi Campus

SCHORLARSHIP : National Science and Technology Development Agency, Ministry of Science and Technology 2008

: National Science and Technology Development Agency, Ministry of Science and Technology 1999