SORGHUM BIOENERGY GENOTYPES, GENES AND PATHWAYS

A Thesis

by

IAN KENNETH PLEWS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2007

Major Subject: Biochemistry

SORGHUM BIOENERGY GENOTYPES, GENES AND PATHWAYS

A Thesis

by

IAN KENNETH PLEWS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	John Mullet
Committee Members,	William Rooney
	William Park
Head of Department,	Gregory Reinhart

December 2007

Major Subject: Biochemistry

ABSTRACT

Sorghum Bioenergy Genotypes, Genes and Pathways. (December 2007)

Ian Kenneth Plews, B.S., University of Winnipeg

Chair of Advisory Committee: Dr. John Mullet

Sorghum (*Sorghum bicolor* [L.] Moench) is the fifth most economically important cereal grown worldwide and is a source of food, feed, fiber and fuel. Sorghum, a C_4 grass and a close relative to sugarcane, is adapted to hot, dry adverse environments and this plant is a potentially important bioenergy crop for Texas. The diversity of the twelve high biomass sorghum genotypes was analyzed using 50 simple sequence repeats (SSR) markers with genome coverage. The accumulation of biomass during sorghum development was studied in BTx623, an elite grain sorghum genotype.

Genetic similarity analysis showed that the twelve high biomass genotypes were quite diverse and different from most current grain sorghum genotypes. The ratio of leaf/stem biomass accumulation was higher early in the vegetative phase during rapid canopy development and lower later in this phase when stem growth rate increased. This resulted in an increasing ratio of stem to leaf dry weight during development. Numerous cellulose sythase genes have been putatively identified in the sorghum genome. The relative level of Ces5 RNA in leaves decreased during vegetative phase of development by ~32 fold. There was no change in the relative abundance of Ces5 RNA in stems. Also there was no change in the relative abundance of Ces3 RNA in either stem or

leaves during the vegetative stage. The knowledge gained in this study may contribute to the development of sorghum bioenergy hybrids that accumulate more biomass and that are modified in composition to make them more amenable to biofuels production.

NOMENCLATURE

AFLP	Amplified fragment length polymorphism
BAC	Bacterial artificial chromosome
EST	Expressed sequence tag
FAM	Fluorescein
HEX	Hexachlorofluorescein
IRD	Infrared dye
LG	Linkage group
LOD	Logarithm of the odds
NIL	Near isogenic line
QTL	Quantitative trait loci
RI	Recombinant inbred
RIL	Recombinant inbred line
RFLP	Restriction fragment length polymorphism
SSR	Simple sequence repeat
TET	Tetrachlorofluorescein
UPGMA	Unweighted pair group method with arithmetic average

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Sorghum is a C4 grass and the fifth most economically important cereal grown in the world (Doggett 1988). In the United States, sorghum has an annual value of two billion dollars, which makes it the second most important feed grain after maize (Dahlberg *et al.* 1995). In parts of Africa and Asia, sorghum is a staple food consumed in breads and porridges (Mann *et al.* 1983).

Sorghum consists of three species, *S. halepense* and *S. propinquum* from the rhizomatous taxa, and *S. bicolor* from the wild, weedy cultivated taxa. Sorghum bicolor is classified into five races using Harlan and deWet's system which is based on spikelet morphology. These races are Bicolor, Guinea, Caudatum, Kafir, and Durra (Figure 1.1). Due to the variability found in each race and the existence of race intermediates an additional classification scheme was developed. This new classification scheme integrates the Harlan and deWet's classification with working groups (sub-races) based on "head opening" and has resulted in the classification of seventy working groups (Dahlberg *et al.* 2004; Murty and Govil 1967).

_____

This thesis follows the style of Genetics.

Early bicolor sorghum is believed to have arisen from the subspecies verticilliforum in central Africa (Dahlberg 1995). When this early Bicolor introgressed with the wild forms of sorghum, the races Caudatum, Kafir, Guinea, and Durra were created. The Guinea race is believed to have evolved when the Bicolors moved west and came into contact with the wild *S. arudinaceum*. The Caudatum race is also believed to arise from an introgressed cross between an early domesticated Bicolor and wild sorghum (Dahlberg 2000). The Kafir race is thought to be derived from crosses between Bicolor in northern Africa with wild verticilliflorum which was carried east and south by the Bantu speakers of Africa (Dahlberg 1995). The Durra race is believed to have originated in Ethiopia when early Bicolor introgressed with wild *S. aethiopicum* which allowed it to adapt to drier conditions (Dahlberg 1995).

Sorghum is a potentially important biomass crop for Texas in part because this plant is highly productive and relatively drought resistant. Moreover, it may be possible to develop sorghum bioenergy hybrids that can accumulate even more biomass and that are modified in composition to make them more amenable to biofuels production.

Biomass includes plant materials and agricultural, industrial, and municipal wastes or residues derived from plant materials. Biomass is accumulated by plants through the process of photosynthesis. Plant biomass can be enriched in starch (grain), soluble sugars (sugarcane extracts) or lignin/cellulose. Cellulosic biomass contains three primary constituents, cellulose, hemicellulose, and lignin plus various amounts of starch, soluble sugars, proteins, and other minor components (Service 2007).

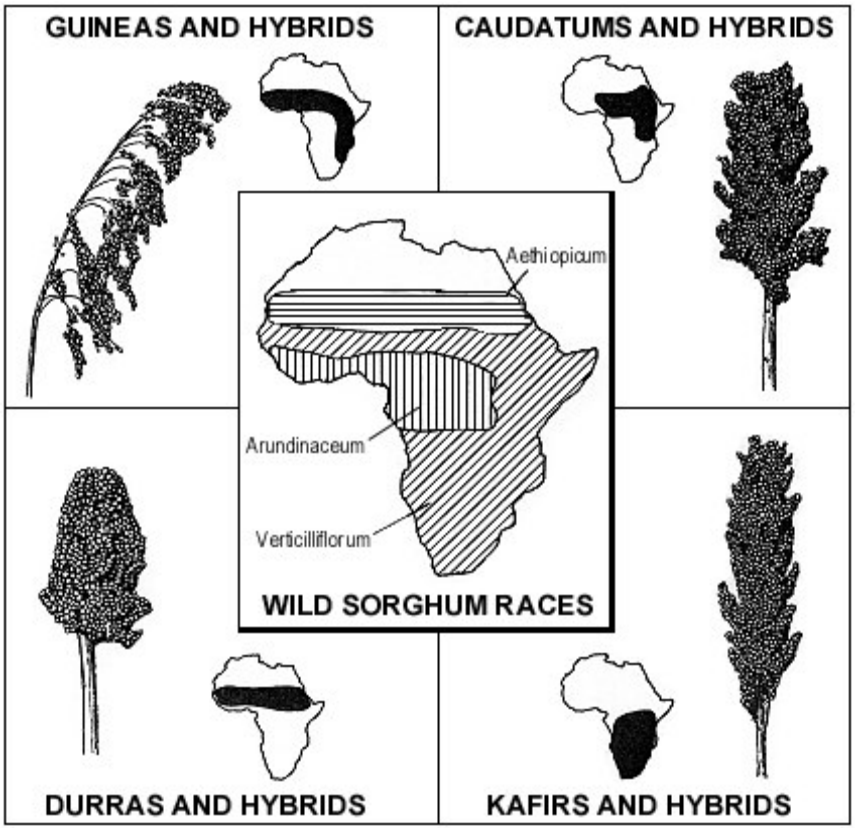

Figure 1.1 Races of sorghum based on spikelet morphology. Courtesy of J.F. Hancock. (2004).

Cellulose synthase catalytic subunit (*CesA*) genes have been identified as encoding the probable catalytic subunits of the cellulose synthase. At least 10 CesA isoforms exist in *Arabidopsis* and have been shown by mutant analyses to play distinct role/s in the cellulose synthesis (Doblin *et al*. 2002). Cellulose synthesis is regulated through tissue specific and differential gene expression during development and,

possibly, regulation of catalytic function. The cellulose synthase enzyme complex can be visualized in the plasma membrane as rosettes as seen in Figure 1.2

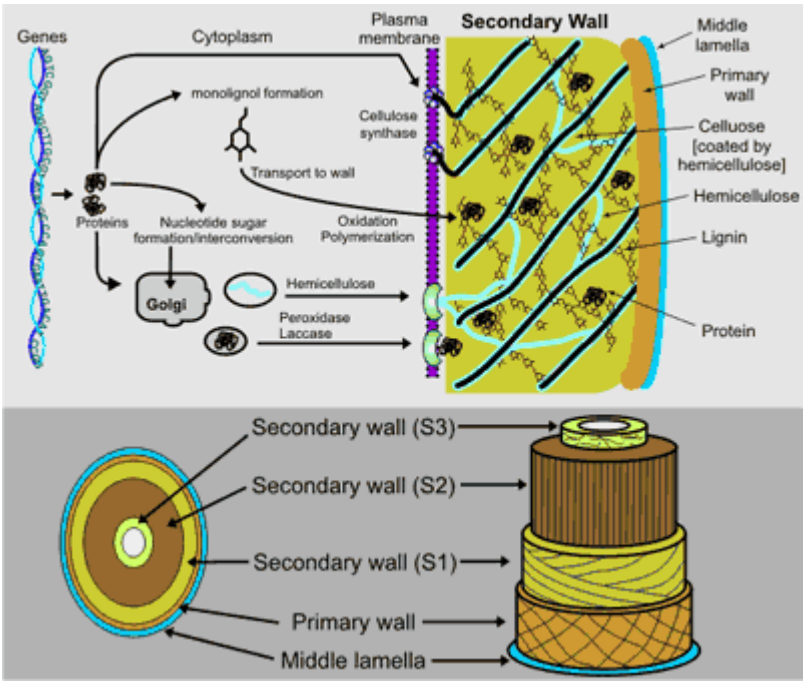

Figure 1.2: Current model for how a plant accumulates biomass in cell walls (www.ccrc.uga.edu/~mao/intro/Secwall.gif).

Cellulose is the most abundance polymer in plants. Cellulose can be hydrolyzed by acid treatment or cellulases to release simple sugars that can be fermented to produce ethanol. The process of cell wall decomposition to sugars is complex and still expensive relative to sugar derived from starch or soluble sugar extracts. This is why ethanol production today is primarily based on corn grain or sugarcane where the rest of the plant is either used as an animal feedstock or to produce energy. Biotechnology

companies are working on developing enzymes that are capable of turning cellulose into ethanol. Once this is achieved it would then lower the cost of the ethanol (or other biofuels) production.

Ethanol is now being used as an alternative transportation fuel in several parts of the world and in the United States. Ethanol derived from biomass is a renewable fuel source and one that contributes less than oil to increases in atmospheric carbon dioxide concentrations. One disadvantage of ethanol is that it has a lower energy content/unit volume than petroleum.

Sorghum belongs to the Poaceae family (tribe Andropogoneae) which includes rice, maize, barley, oats, rye, millet, and wheat. Despite the separation of sorghum from maize and rice approximately ten to twenty million years ago and fifty million years ago, respectively, there is significant conservation of the gene order among these plant genomes (Bennetzen 2000). Sorghum has a diverse germplasm and a relatively small diploid genome of 760 – 810 Mbp (Arumuganathan and Earle 1991) making it well suited for genomic approaches. A high-density integrated genetic and physical map of sorghum has been created (Klein *et al.* 2000). A dense genetic map of sorghum was obtained by scoring 2454 AFLPs, 203 RFLPs, and 136 SSRs in a recombinant inbred (RI) population, consisting of 137 lines. The RI lines were derived by crossing the *S. bicolor* genotypes BTx623 and IS3620C (Menz *et al.* 2002). A physical map was then generated using three different BAC libraries with BAC contigs anchored every 1.5 cM on all ten chromosomes. The integration of genetic and physical maps permits map based cloning of important genes (Childs *et al.* 1997; Klein *et al.* 2005).

The sorghum genome of the genotype BTx623 has recently been sequenced by the Department of Energy Joint Genome Institute (Sorghum Genomics Workshop Planning Participants 2005). An 8X shotgun sequence was combined with the paired-end sequencing from sorghum BACs as well as other publicly available sorghum sequences such as the 550,000 methyl filtered sequences (Bedell *et al.* 2005) to form pseudomolecules that cover most of the chromosomes (Paterson *et al.* 2006).

.

CHAPTER II

MATERIALS AND METHODS

Sorghum seeds were treated with the fungicide Captan 400 (Gustafson LLC, Plano, TX, USA) at a 1:26 ratio of Captan to water and allowed to dry. The treated seeds then were grown on Whatman chromatography paper (Whatman Inc, Florham Park, New Jersey, USA) in distilled water under growth chamber conditions. Once the seeds had germinated they were transferred to hydroponic growth conditions as described (Buchanan *et al.* 2004). All plants were fertilized with 1x Hoaglands containing .005M Ca(NO3)2 x 4H2O, 5 mM KNO3, 5 mM MgSO4 x 7H2O, 5 mM MgSO4, pH 6.5 with KOH, 1.4μM ZnSO4 x 7H2O,46 μM H3BO3, 14.1 μM MnCl2 x 4H2O, .48μM CuSO4 x 5H2O, 82.6 ηM NaMoO4 x 2H2O, and .055g/L of Sprint 330 containing 10% Iron.

The genomic DNA of etiolated samples were then extracted using a FastDNA kit (Qbiogene, Irvine, CA, USA) with the FastPrep FP120 instrument to homogenize sorghum seedling leaf tissue according to the Manufacturer's protocol.

SSRs were amplified and analyzed using fluorescent infrared dye (IRD)-labeled primers as described (Klein *et al.* 2000) such as 5' Hexachlorofluorescein (HEX) or Tetrachlorofluorescein (TET) or Fluorescein (FAM) forward labeled primers (XXIDT, Coralville, IA, USA). For reactions containing HEX/TET/FAM labeled primers the PCR conditions were 2.5 pmole/μl of the forward and reverse primer. Following amplification, the 10μl PCR products from the HEX/TET/FAM labeled

primers were diluted with 20µl of reverse osmosis (RO) water. Five µl of formamide:size standard mix at 1ml of HiDi formamide:25µl of 400 HD ROX (Applied Biosystems, Branchburg, NJ, USA) was added to a 96-well Perkin Elmer plate and 1 µl of dilute PCR product was then added.

Reactions were mixed, centrifuged briefly to remove bubbles and denatured at 95C for 5 minutes before loading the plate on an ABI 3700 Sequencer. Genotyping was performed on the sequencer by using the SUP2_POP5 module. Collected data was analyzed with Gene Scan version 3.7 Fragment Analysis Software (Applied Biosystems) and peaks were scored manually by using the Genotyper version 3.7 Fragment Analysis Software (Applied Biosystems). SSR primer sequences as well as amplification product sizes are listed at http://sorgblast3.tamu.edu/ and an annealing temperature of 50C was used for all reactions.

For genetic similarity analysis using the SSRs, the presence of a band was represented as a "1" and the absence of a band was represented with a "0". Failed reactions were represented as a "9". The program NTSYS-pc (Rohlf 1994) was used to generate a genetic similarity matrix among all pairs of lines by Dice coefficient of similarity (Nei and Li 1979). A dendrogram was created from the similarity matrix by the unweighted pair group method with arithmetic average (UPGMA). For bootstrapping, the program FreeTree (Hampl *et al.* 2001) was used and 500 bootstrap repetitions were used.

BTx623 seed was first bubbled in distilled water over night in growth chamber conditions. The seeds then were grown on Whatman chromatography paper (Whatman Inc, Florham Park, New Jersey, USA) in distilled water for three days until they germinated. The germinated seeds were then transferred to fifty five gallon pots and grown on soil under green house conditions. Each five gallon pot contained three plants. Once a week the pots were watered as needed and rotated to different locations to reduce spatial effects. On day ten of plant growth and every 14 days after that, 6 BTx623 plants were harvested for analysis. Plant height, leaf number, and stem thickness (at the base) were measured. The leaves and stems were separated and weighed. The leaves and stems were then dried at 65 degrees in an oven for a couple of days until dry and then weighed again to determine dry weight. On day 10 and every 14 days thereafter three BTx623 plants were harvested, frozen in liquid nitrogen and ground.. RNA was extracted using Tri Reagent and phenol. The RNA was then purified using a Qiagen RNA Clean up Kit. cDNA was synthesized from total RNA obtained from at least three independent biological replicates per sample using random hexamers and TAQMAN reverse transcription reagents (Applied Biosystems, Branchburg, NJ). Quantitative real time PCR was performed on an Applied Biosystems 7900HT machine using SYBR green detection of product accumulation. Primers and probes were designed using Primer Express software (Applied Biosystems) to allow for amplification of 100-bp products of similar GC and Tm characteristics. Thermal cycling conditions were 2 m at 50 _C and 10 m at 95 _C followed by 47 cycles at 95 _C for 15 s and 60 _C for 1 m. Samples were assayed in triplicate and data was analyzed using the ABI PRISM 7900HT

SDS software (Applied 702 Biosystems). Quantification was achieved using the comparative CT method (Bieche et al., 1999), which consists of normalizing the abundance of target gene copies to an endogenous reference gene (18S rRNA, detected using the ribosomal TAQMAN kit supplied by Applied Biosystems). Fold inductions were calculated as 2^ðDCTcontrol_ DCTtreatmentÞ. Variability of RT-PCR results among biological replicates was determined by analyzing the mRNA levels of two genes in RNA derived from three biological replicates. Sorghum genes encoding cellulose synthase were identified using maize and rice cellulose synthase gene sequences to identify the corresponding sorghum genes (Blast analysis). Real Time Polymerase Chain Reaction (RT-PCR) primers were developed for two of the putative sorghum genes so that expression of the cellulose synthase genes at different stages in the plant growth could be determined.

CHAPTER III

DIVERSITY OF SOURCES OF BIOMASS IN THE SORGHUM GERMPLASM COLLECTION

Introduction

Characterize the genetic diversity of high biomass sorghum genotypes and determine if they are similar or different from grain sorghum genotypes. Twelve high biomass sorghum genotypes identified by Dr. Bill Rooney were analyzed to determine the genetic similarity of the bioenergy sorghum lines, and their similarity to other grain sorghum genotypes in the sorghum-breeding program. A set of 50 SSR markers were multiplexed and used to analyze DNA from the twelve high biomass sorghum genotypes. The 50 SSR markers used in this study had been previously been used to analyze ~50 sorghum breeding lines and the progenitors of the stay-green trait. The SSR markers identified 2 to more than 24 alleles within a collection of 81 sorghum genotypes analyzed.

Information from the SSR assays was used to examine the relationship of the sorghum bioenergy genotypes relative to each other and other sorghum breeding lines. The results are displayed in a dendrogram which was created by first calculating the genetic similarity between all pairs of lines of sorghum and then clustering these data (Figure 3.1).

Results

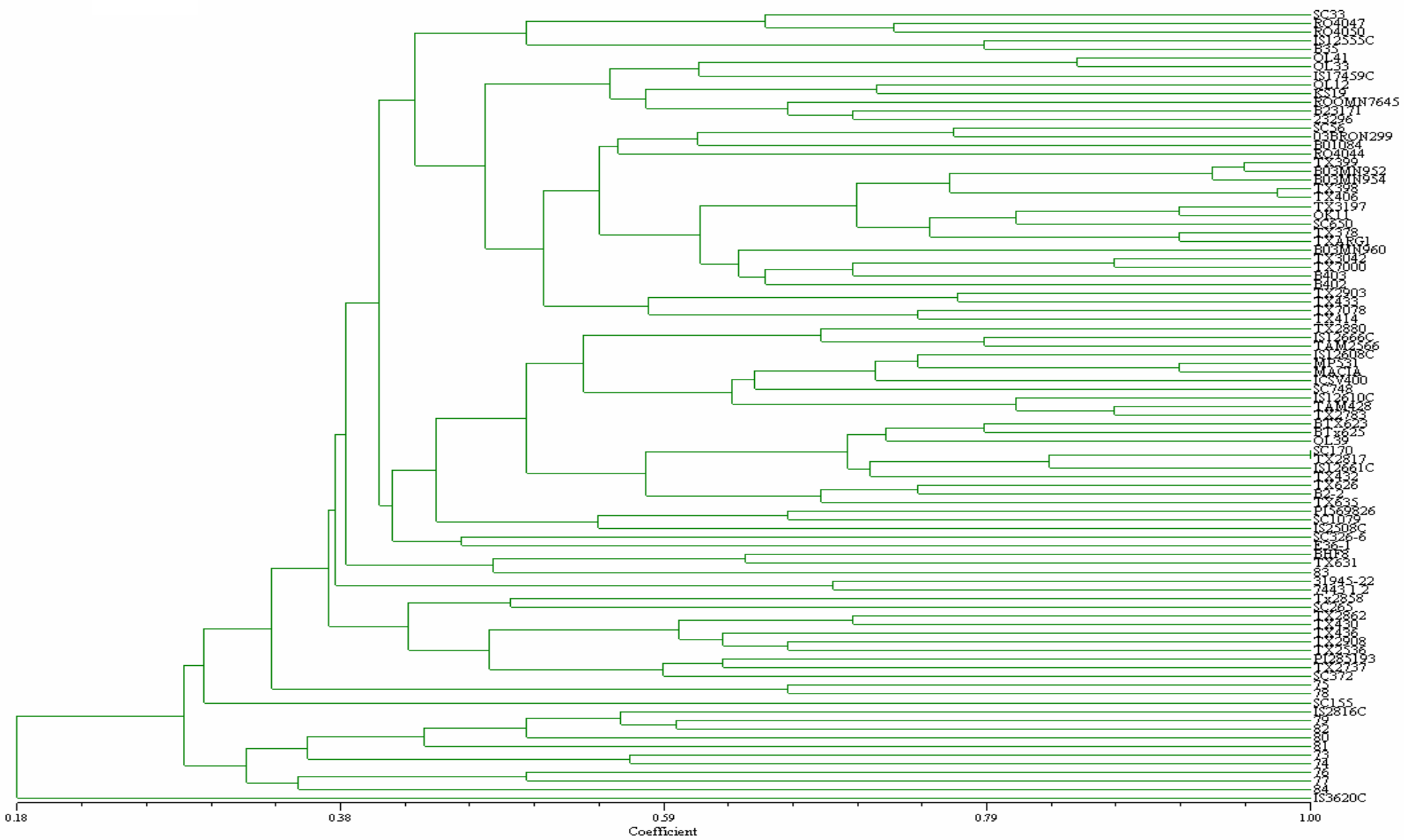

Figure 3.1: Dendrogram of 81 sorghum genotypes revealed by cluster analysis of genetic similarity using data collected from 50 genome-wide SSR markers. The twelve sources of the high biomass lines are numbered (list here) . The thick black vertical line denotes the cutoff value for the clusters. Group I includes genotypes of the race Durra, Group II includes genotypes of the race Caudatum, Group III includes genotypes that are related to QL33/RTx430/RTx436, Group IV includes genotypes that belong to the race Kafir, Group V includes genotypes of the working group Caudatum-kafir of the race Caudatum, and Group VI includes genotypes from the working group Caudatum-kaura of the race Caudatum. 500 bootstrap repetitions were performed and shown.

Discussion

The information from SSR analysis was used to assess the genetic relatedness of the 81 genotypes analyzed in this study (Rohlf 1994; Nei and Li 1979).). The genotypes are distributed across the dendrogram and most clustered with other genotypes of the same race or working groups as previously described (Menz *et al.* 2004).

The twelve sorghum bioenergy genotypes for the most part did not cluster with sorghum genotypes that have been used in breeding grain sorghum in the U.S. over the past 50+ years. There was also significant diversity among the twelve high biomass genotypes that may be useful in subsequent breeding activity.

CHAPTER IV

QUANTIFICATION OF LEAF AND STEM BIOMASS ACCUMULATION DURING VEGETATIVE PHASE

Introduction

Differences in biomass accumulation among genotypes could be due to variation in overall growth rates, differential growth during a specific phase of development or the duration of vegetative growth. Grain sorghum is generally photoperiod insensitive and will flower in less than 100 days in College Station. In contrast, high biomass sorghum is photoperiod sensitive and many genotypes will not flower for at least 175 days. The composition of biomass in leaves and stems is different and changes during plant development. For example, leaves contain high levels of chlorophyll, lipids and starch whereas most stems have higher levels of cellulose and lignin. Therefore an understanding of biomass accumulation during plant development may be important in order to optimize biomass composition for downstream conversion to biofuels.

Differences in gene expression often are the basis of differences in phenotype. Prior studies showed that plants contain a large number of genes for cellulose synthase and that different members of the gene family are differentially expressed during development and in specific tissues (Doblin 2002). Therefore, members of the sorghum gene family that encode cellulose synthase were identified and the expression of two of these genes characterized during vegetative phase development.

Results

BTx623 plants were grown and harvested at different time points through out the vegetative phase of plant development. The accumulation of leaf and stem fresh and dry weight during development is shown in Figures 4.1 and 4.2. The ratio of leaf and stem fresh and dry weight was calculated and displayed in Figure 4.3. The change in height and width of the stem is shown in Figures 4.4 and 4.5. While the changes in the Ces 3 and Ces 5 during development are displayed in Table 4.1.

Discussion

The results show that the biomass in sorghum accumulates throughout the vegetative phase but it is allocated differently between the leaves and the stem during development. During the early part of the vegetative phase the ratio of biomass accumulation in leaves is more rapid than stems whereas during the latter portion of this developmental stage stem biomass accumulation is more rapid than leaf biomass accumulation. This is consistent with the need for rapid leaf development early in plant growth in order to create a large photosynthetic apparatus for carbon fixation . As canopy closure occurs, there is a transition to greater relative stem growth as the plant increases in height. This is seen in that twice as much biomass is partitioned to the stem as in the leaves by the end of the vegetative phase.

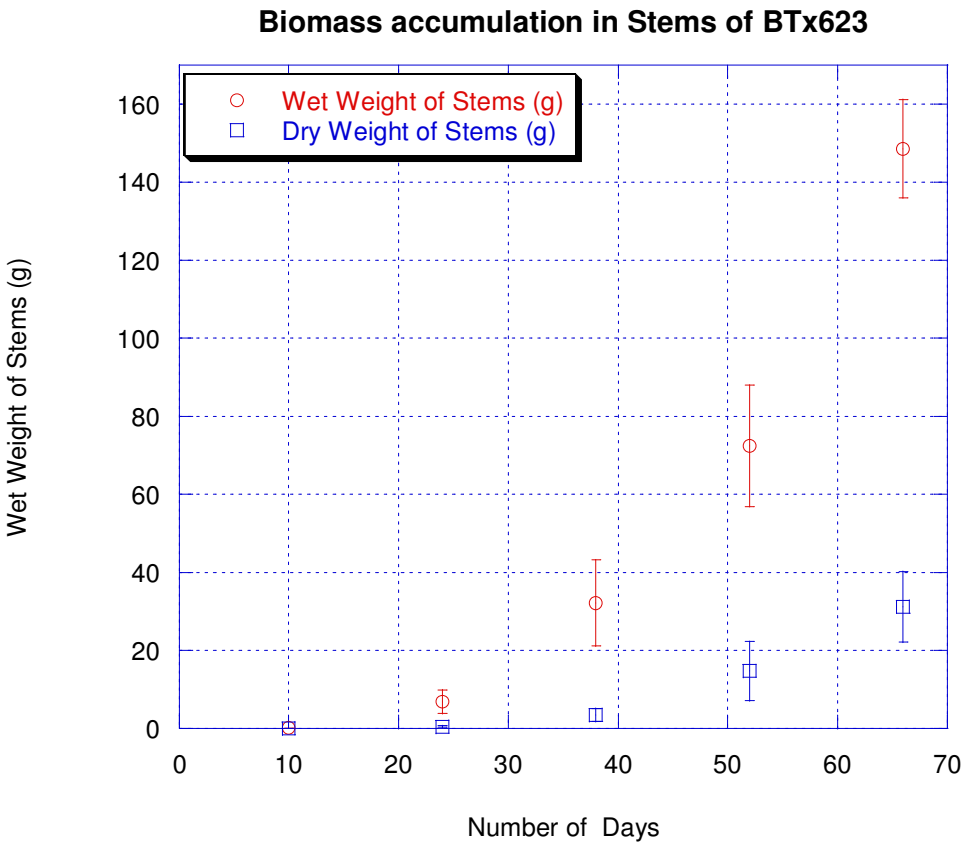

Figure 4.1: Accumulation of biomass in stems of BTx623 over a period of 66 days post germination. Plants were harvested at the time points shown, leaves and stems separated and stems weighed (fresh weight) then dried and reweighed (dry weight). Each point on the graph represents an average of six plants.

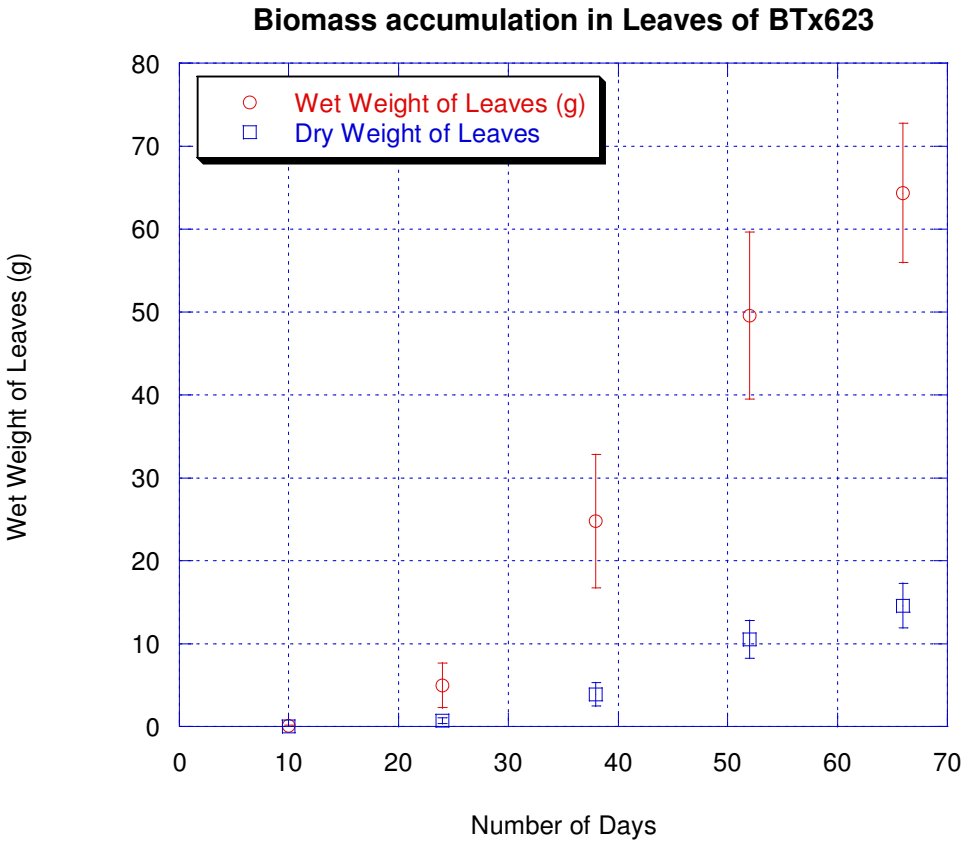

Figure 4.2: Accumulation of biomass in leaves of BTx623 over a period of 66 days post germination. Plants were harvested at the time points shown, leaves and stems separated and leavess weighed (fresh weight) then dried and reweighed (dry weight). Each point on the graph represents an average of six plants.

.

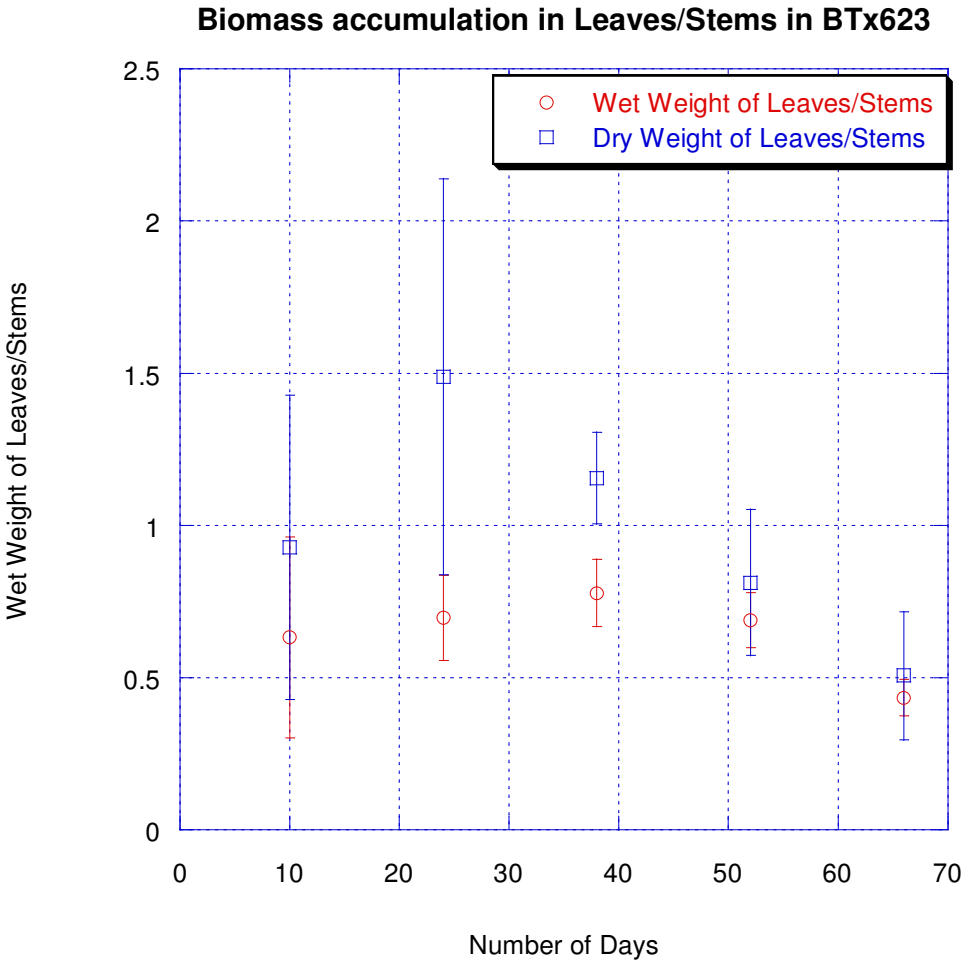

Figure 4.3: Ratio of biomass accumulation in leaves/stems of BTx623 over a period of 66 days post germination. The graph was created by dividing the weight of each plant's leaves by the weight of the plant's stem at each time. The points on the graph represent the average of the six ratios at each time.

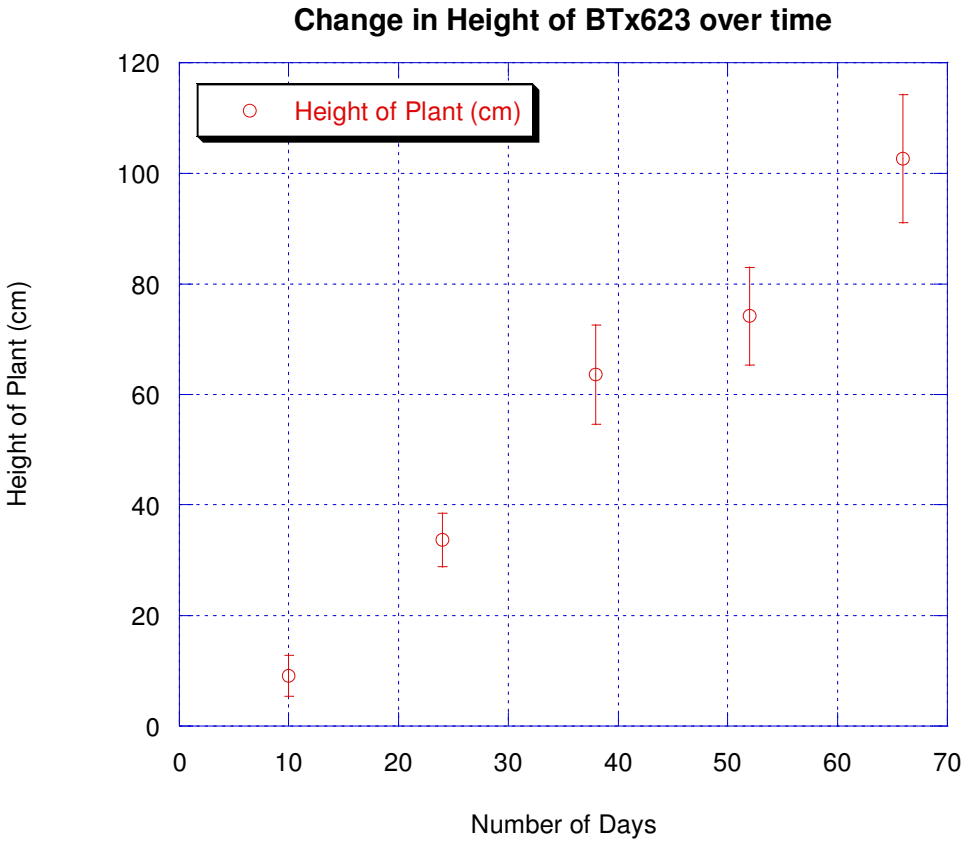

Figure 4.4: The change in the height of BTx623 over a period of 66 days post germination. The points on the graph at each time point show the average height of six plant heights measured in centimeters

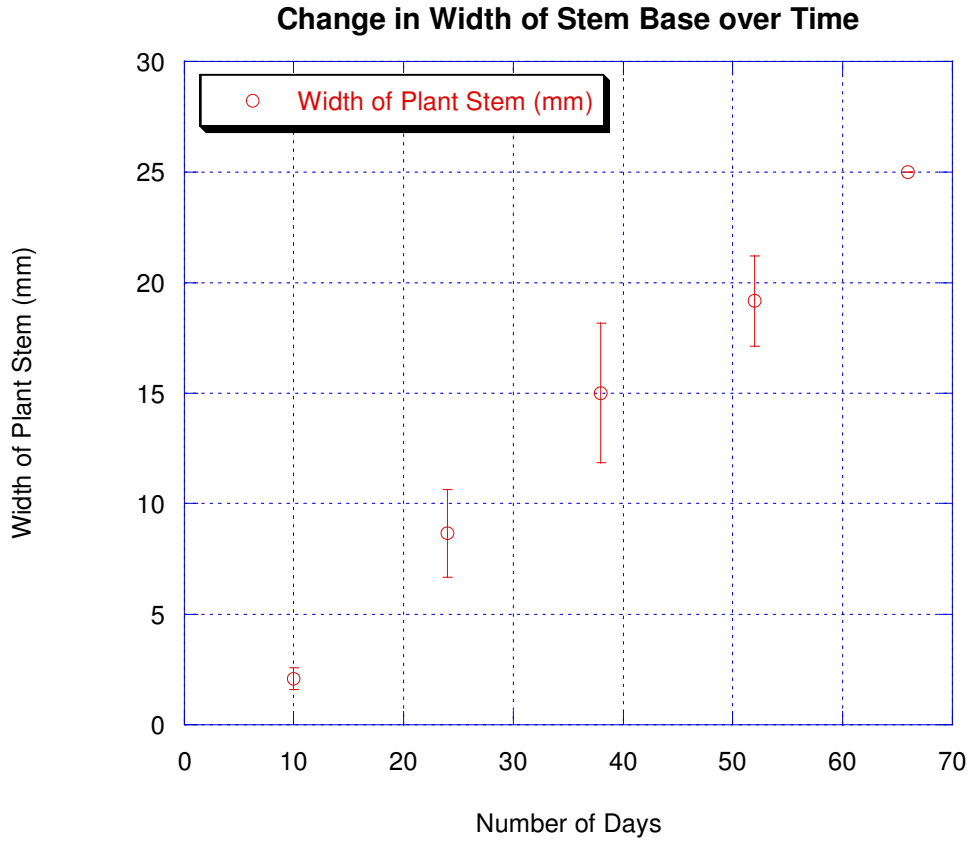

Figure 4.5: The change in the width of stem base of BTx623 over a period of 66 days post germination. The points on the graph at each time point show the average width of six plants stems measured at the base in millimeters.

Table 4.1: Abundance of Ces3 and Ces5 RNA in leaves and stems during vegetative phase development of BTx623.

Ces #3	10 Days	24 Days	38 Days	52 Days	66 Days
Leaves	12.37	13.32	13.82	13.22	13.21
Stems	12.41	12.79	11.44	11.69	12.41
Ces #5	10 Days	24 Days	38 Days	52 Days	66 Days
Leaves	8.87	10.04	10.14	12.19	14.55
Stems	8.87	8.44	7.84	7.80	8.22

The numbers in the table are delta Ct values and were determined by subtracting the 18S Ct values from the Ces Ct values. The delta Ct values 8.870165 and 14.55196 are significant because there is a difference of 5 cycles between them. A sample whose delta C_t is 5 cycles earlier than another's has $2^5 = 32$ times more template. , need to explain what the numbers mean (delta CT…).

The switch in biomass accumulation may be caused in part by the shade avoidance response and partly to a transition to the booting phase of sorghum development. This allows plants to grow in height and out compete its neighbor for sunlight. This information may influence when to harvest the plant because more cellulose is located in the stem than in the leaves.

Results from the RT-PCR assays showed that the relative expression of Ces5 RNA in leaves decreased during vegetative phase development by 32 fold from the beginning to the end. There was no change in the relative expression of Ces5 RNA in stems. Also there was no change in the relative expression of Ces3 RNA found in either stem or leaves. The 32 fold drop in Ces5 RNA level during development is consistent with a decrease in rate of leaf growth that occurs during plant development.

CHAPTER V

CONCLUSIONS

This research showed that the twelve different high biomass lines were genetically different from most of sorghum genotypes that have been used to produce grain sorghum. A subset of the fifty SSRs revealed polymorphisms that distinguish each of the high biomass lines. These SSRs can be used to confirm crosses among these lines and they will be useful in subsequent genetic mapping experiments.

Sorghum accumulates biomass throughout the vegetative phase but it is allocated differently between the leaves and the stem. During the beginning of the vegetative phase the biomass partitions differentially to the leaves because the plant is building its photosynthetic capacity. The biomass then partitions more biomass to the stem at the middle to end of the vegetative phase. Therefore, if the plant is to be used to produce ethanol from lignocellulose it may be better to harvest the plant at the end of the vegetative stage.

The relative expression of Ces5 RNA in leaves decreased during vegetative phase development by 32 fold. The decrease in Ces5 mRNA level may reflect the decrease in overall leaf growth rate during development. There was no change in the relative expression of Ces5 RNA in stems. Also there was no change in the relative expression of Ces3 RNA found in either stem or leaves suggesting that this gene may function to produce a basal level of cellulose synthase during all phases of plant growth.

LITERATURE CITED

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

Bedell, J. A., M. A. Budiman, A. Nunberg, R. W. Citek, D. Robbins *et al*., 2005 Sorghum genome sequencing by methylation filtration. PLoS Biol. **3:** 103-115.

Bennetzen, J. L., 2000 Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions. Plant Cell **12:** 1021-1030.

Bieche, I., Laurendeau, I., Tozlu, S., Olivi, M., Vidaud, D., Lidereau, R. and Vidaud, M. 1999. Quantitation of MYC gene expression in sporadic breast tumors with a real-time reverse transcription-PCR assay. Cancer Res. 59: 2759–2765

Buchanan, C. D., P. E. Klein and J. E. Mullet, 2004 Phylogenetic analysis of 5'-noncoding regions from the ABA-responsive rab16/17 gene family of sorghum, maize and rice provides insight into the composition, organization and function of cisregulatory modules. Genetics **168:** 1639-1654.

Childs, K. L., F. R. Miller, M. M. Cordonnier-Pratt, L. H. Pratt, P. W. Morgan *et al.*, 1997 The sorghum photoperiod sensitivity gene, Ma_3, encodes a phytochrome B. Plant Physiol. **113:** 611-619.

Dahlberg, J., 1995 Dispersal of sorghum and the role of genetic drift. African Crop Sci. Journal **3:** 143-151. 119

Dahlberg, J. A., 2000 Classification and characterization of sorghum, p.99-130 in *Sorghum:Origin, History, Technology, and Production,* edited by C. W. Smith and R. A. Frederiksen. John Wiley & Sons, Inc., New York, NY.

Dahlberg, J. A., J. J. Burke and D. T. Rosenow, 2004 Development of a sorghum core collection: refinement and evaluation of a subset from Sudan. Econ. Bot. **58:** 556 567.

Doblin, M.S., I. Kurek, D. Jacob-Wilk and Deborah P. Delmer, 2002, Cellulose biosynthesis in plants: from genes to rosettes, Plant and Cell Physiology **43** No. 12 1407-1420

Doggett, H., 1988 *Sorghum* (2 Edn ed.), John Wiley & Sons Inc., New York.

Hancock JF. 2004. Plant evolution and the origin of crop species. Wallingford: CABI Publishing.

Hampl, V., A. Pavlicek and J. Flegr, 2001 Construction and bootstrap analysis of DNA fingerprinting-based phyologenetic trees with the freeware program FreeTree: application to trichomonad parasites. Int. J. Syst. Evol. Microbiol. **51:** 731-735.

Klein, P. E., R. R. Klein, S. W. Cartinhour, P. E. Ulanch, J. Dong *et al.*, 2000 A highthroughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res. **10**: 789–807.

Klein, R. R., P. E. Klein, J. E. Mullet, P. Minx, W. L. Rooney *et al*., 2005 Fertility restorer locus Rf1 of sorghum (Sorghum bicolor L.) encodes a pentripeptide repeat protein not present in the collinear region of rice chromosome 12. Theor. Appl. Genet. **111:** 994-1012.

Mann, J. A., C. T. Kimber and F. R. Miller, 1983 *The Origin and Early Cultivation of Sorghums in Africa*. Bulletin B Texas Agric. Exp. Stn. 1454, 21pp.\

Menz, M. A., R. R. Klein, J. E. Mullet, J. A. Obert, N. C. Unruh *et al.*, 2002 A highdensity genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. Plant Mol. Biol. **48:** 483-99.

Menz, M. A., R. R. Klein, N. C. Unruh, W. L. Rooney, P. E. Klein *et al.*, 2004 Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. Crop Sci. **44:** 1236-1244.

Murty, B. R., and J. N. Govil, 1967 Description of 70 working groups in genus sorghum based on a modified Snowsen's classification. Indian J. Genet. **27:** 75-91.

Nei, M., and W. H. Li, 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76:** 5269-5273.

Paterson, A. H., J. E. Bowers, A. R. Gingle, D. G. Peterson, S. E. Kresovich *et al.*, 2006 *Sequencing of the Sorghum Genome*. Plant and Animal Genome XIV Conference. January 14-18 2006. San Diego, CA.

Rohlf, F. J., 1994 *NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Ver 1.8.* Exeter Software, Setauket, NY.

Service, R.F.,2007 Cellulosic ethanol. Biofuel researchers prepare to reap a new harvest. Science. 2007 Mar 16;**315** (5818):1488-91

APPENDIX A

Table A-1 Weights and Measurements from six different BTx623 plants taken over a period of 66 days

Day 10	Bag weight	Bag Wet weight	Bag Dry Weight	Length (cm)	Width (mm)	# of Leaves
Stem 1	1.7159	1.7648	1.7206	5.7	2	3
Stem 2	1.7042	1.9588	1.733	14	3	3
Stem3	1.7097	1.7995	1.7209	11.2	2	2
Stem4	1.7032	1.7798	1.7111	7.8	1.5	2
Stem5	1.6939	1.9232	1.7202	11.4	2	3
Stem6	1.6947	1.7425	1.6995	4.6	2	3
Leaves1	1.7469	1.7985	1.7541			
Leaves2	1.7502	1.9174	1.7772			
Leaves3	1.7535	1.7843	1.7586			
Leaves4	1.7514	1.7796	1.7562			
Leaves5	1.7523	1.8375	1.7652			
Leaves6	1.7571	1.8051	1.7645			
Day 24						
Stem1	2.48	9.02	2.91	35.5	10	8
Stem2	2.46	7.24	3.22	31.5	9	6
Stem3	2.48	12.86	2.84	42	10	14
Stem4	2.44	8.52	2.88	28	8	8
Stem5	2.47	5.46	2.71	32	5	5
Stem6	2.47	12.98	3.14	33	10	6
Leaves1	2.5	6.55	3.01			
Leaves2	2.48	6	3.02			
Leaves3	2.43	9.17	3.35			
Leaves4	2.45	7.26	3.13			
Leaves5	2.42	3.91	2.68			
Leaves6	2.48	11.76	3.72			

Table A-1 Continued

Day 38						
Stem1	2.48	34.01	8.49	66.5	15	8
Stem2	2.5	43.66	6.98	71.5	15	10
Stem3	2.48	15.95	3.96	55.5	10	8
Stem4	2.48	33.14	5.14	49.5	20	10
Stem5	2.48	33.5	4.86	68	15	8
Stem6	2.48	47.61	6.44	70.5	15	9
Leaves1	2.49	31.5	8.31			
Leaves2	2.5	31.93	7.23			
Leaves3	2.48	12.69	4.18			
Leaves4	2.47	30.45	6.05			
Leaves5	2.48	23.16	5.65			
Leaves6	2.48	33.9	6.76			
Day 52						
Stem1	2.43	82.58	15.57	77	20	11
Stem2	2.43	74.96	30.6	67	20	11
Stem3	2.42	82.2	14.83	72	20	13
Stem4	2.41	53.7	9.68	73	15	11
Stem5	2.44	59.94	11.5	66	20	13
Stem6	2.43	95.83	20.7	90	20	13
Leaves1	2.5	63.26	14.34			
Leaves2	2.41	57.42	12.9			
Leaves3	2.44	53.02	13.68			
Leaves4	2.43	34.46	9.09			
Leaves5	2.45	47.22	11.95			
Leaves6	2.44	56.58	15.8			
Day 66						
Stem1	2.41	149.69	41.25	100	250	13
Stem2	2.48	156.04	30.66	97	250	13
Stem3	2.46	145.05	23.53	112	250	11
Stem4	2.47	133.6	25.38	84	270	13
Stem5	2.4	171.61	46.45	116	250	12
Stem6	2.49	150.39	34.7	107	200	13
Leaves1	2.5	59.53	17.24			
Leaves2	2.38	74.31	12.69			
Leaves3	2.48	57.97	21.03			
Leaves4	2.44	71.93	16.26			
Leaves5	2.48	76.68	16.95			
Leaves6	2.47	60.46	17.89			

Table A-2.1: RT-PCR values for Ces 3 from triplicate BTx623 plants at 10 days

Ct	Avg Ct	St. Dev.	Delta Ct	
38.76049	38.53312	0.321549	12.17067	
38.30575				
28.7362	28.79464	0.094693	12.27083	Leaf and Stem at 10 Days
28.74382				12.37
28.90389				
27.97306	27.79493	0.210086	12.67017	
27.56325				
27.84848				

Table A-2.2: RT-PCR values for Ces 3 from triplicate BTx623 plants at 24 days

Ct	Avg. Ct	St. Dev.	Delta Ct	
32.60344	32.79635	0.272816	12.89913	
32.98926				
	31.633	0.21313	13.90929	Leaf at 24 Days
31.4823				13.32
31.78371				
28.64419	29.48534	0.837363	13.14431	
30.31887				
29.49297				
28.7362	28.79464	0.094693	12.20717	
28.74382				
28.90389				
	29.29779	0.380104	13.68755	Stem at 24 Days
29.56657				12.79
29.02902				
25.95944	26.18743	0.221674	12.48786	
26.20064				
26.4022				

Table A-2.3: RT-PCR values for Ces 3 from triplicate BTx623 plants at 38 days

Ct	Avg. Ct	St. Dev.	Delta Ct	
	30.2968	0.310771	15.09153	
30.07705				
30.51654				
30.70884	30.84753	1.096632	13.64804	Leaf at 38 Days
29.82684				13.82
32.00691				
30.42575	30.35887	0.094588	12.73325	
30.29198				
27.59546	27.7334	0.195083	11.1527	
27.87135				
28.0245	28.23301	0.276668	10.73151	Stem at 38 Days
28.54688				11.44
28.12766				
27.97306	27.79493	0.210086	12.43335	
27.56325				
27.84848				

Table A-2.4: RT-PCR values for Ces 3 from triplicate BTx623 plants at 52 days

Ct	Avg. Ct	St. Dev.	Delta Ct	
27.90745	28.05921	0.142871	14.43002	
28.07907				
28.19111				
25.21719	25.40443	0.191115	13.99237	Leaf at 52 Days
25.5992				13.22
25.39688				
25.47471	25.74646	0.254087	11.24182	
25.78657				
25.97811				
31.47217	31.3155	0.202502	12.53958	
31.38749				
31.08684				
31.01171	31.26684	0.226197	11.5003	Stem at 52 Days
31.44283				11.69
31.34598				
30.3973	30.21634	0.165661	11.0182	
30.07216				
30.17955				

Table A-2.5: RT-PCR values for Ces 3 from triplicate BTx623 plants at 66 days

Ct	Avg. Ct	St. Dev.	Delta Ct	
28.34384	28.25727	0.202427	13.17421	
28.02596				
28.40202				
28.53643	28.50846	0.039563	13.35544	Leaf at 66 Days
28.48048				13.21
32.90827	33.42832	1.004896	13.10586	
32.79002				
34.58666				
	31.89585	0.540252	12.61529	
31.51383				
32.27786				
31.42477	31.047	0.53424	12.47636	Stem at 66 Days
30.66924				12.41
31.38274	31.32509	0.134923	12.13293	
31.17092				
31.42162				

Table A-3.1: RT-PCR values for Ces 5 from triplicate BTx623 plants at 10 days

Ct	Avg. Ct	St. Dev.	Delta Ct	
25.19776	25.071	0.166128	25.071	
24.88293				
25.1323				
25.57898	25.74966	0.147856	25.74966	Leaf and Stem at 10 Days
25.83854				8.87
25.83145				
25.63966	25.71577	0.11617	25.71577	
25.65817				
25.84949				

Table A-3.2: RT-PCR values for Ces 5 from triplicate BTx623 plants at 24 days

Ct	Avg. Ct	St. Dev	Delta Ct	
30.4018	30.32016	0.254081	10.42294	
30.52338				
30.03529				
	28.50295	0.097998	10.77924	Leaf at 24 Days
28.57225				10.04
28.43366				
25.20262	25.24834	0.076921	8.907306	
25.20526				
25.33715				
24.34168	24.63683	0.259319	8.049368	
24.74073				
24.8281				
24.70826	24.94874	0.223066	9.338493	Stem at 24 Days
25.1489				8.44
24.98906				
21.75757	21.6304	0.179849	7.930826	
21.50323				

Table A-3.3: RT-PCR values for Ces 5 from triplicate BTx623 plants at 38 day

Ct	Avg. Ct	St. Dev	Delta Ct	
27.19028	26.91991	0.234708	11.71464	
26.80094				
26.76851				Leaf at 38 Days
28.31798	28.12541	0.167439	10.92592	10.14
28.04401				
28.01422				
25.20327	25.40028	0.174338	7.7746611	
25.53461				
25.46296				
23.61643	23.77216	0.22024	7.191518	
23.9279				
24.55029	24.49713	0.162253	6.995629	Stemat 38 Days
24.31497				7.84
24.62614				
24.81475	24.7015	0.174501	9.339931	
24.78921				
24.50055				

Table A-3.4: RT-PCR values for Ces 5 from triplicate BTx623 plants at 52 days

Ct	Avg. Ct	St. Dev.	Delta Ct	
31.65584	31.74832	0.113408	14.86283	
31.71426				
31.87485				
31.10072	30.85985	0.222049	13.58537	Leaf at 52 Days
30.66331				12.19
30.81551				
26.99778	27.10624	0.179875	8.131723	
27.31387				
27.00706				
23.29012	23.37079	0.114081	7.851166	
23.45146				
21.54726	21.66071	0.098362	7.756585	Stem at 52 Days
21.71264				7.80
21.72222				
22.05598	22.15749	0.113529	7.429219	
22.13641				
22.28008				

Table A-3.5: RT-PCR values for Ces 5 from triplicate BTx623 plants at 66 day

Ct	Avg. Ct	St. Dev.	Delta Ct	
31.72576	32.08189	0.308478	13.36025	
32.25364				
32.26626				
34.42497	34.00937	0.587749	16.31781	Leaf at 66 Days
				14.55
33.59377				
32.30409	32.19705	0.151367	13.97782	
32.09002				
23.89969	23.9433	0.140293	8.301313	
23.82999				
24.10022				
25.87255	25.94908	0.150537	9.916982	Stem at 66 Days
26.12251				8.22
25.85219				
27.9817	27.8494	0.187098	6.430796	
27.7171				

VITA

Ian Kenneth Plews received a Bachelor of Science degree in chemistry from the University of Winnipeg in 2004. He entered the biochemistry and biophysics program at Texas A&M University in August 2005, and received his Master of Science degree in December 2007. His research interests include plant molecular biology, genetics, and genomics.

Mr. Plews can be reached at the Borlaug Center, c/o John Mullet, Texas A&M University, MS 2123, Room 112, College Station, TX 77843. His email address is ianplews@tamu.edu.